



Information Society  
Technologies

# erpastudies

**pharmaceuticals**

---

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).

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 2003



**Table of Contents**

Executive Summary .....	4
Chapter 1: The ERPANET Project .....	5
Chapter 2: Scope of the case studies .....	6
Chapter 3: Method of Working .....	8
Chapter 4: Introduction to the pharmaceutical sector.....	9
Agency regulations.....	10
Chapter 5: Details of interviews .....	12
Chapter 6: Circumstances .....	15
Chapter 7: Analysis.....	16
Perception and Awareness of Digital Preservation.....	16
Preservation Activity .....	18
Compliance Monitoring.....	22
Digital Preservation Costs.....	23
Future Outlook .....	23
Chapter 8: Conclusions.....	25
Chapter 10: Resources .....	26
Appendix: Interview Instrument.....	29

## **Executive Summary**

Considerable time and effort has already been invested in digital preservation in the pharmaceutical sector, largely due to pressure by governmental agencies who are committed to improving the safety, quality and efficacy of new medical products. Organisations in the sector are under economic and regulatory pressure to move to more efficient and effective digital processes.

To date, efforts in digital preservation focus predominantly on research and development, but solutions are required across these large complex organisations to meet the growing demands of digital asset management. Organisational frameworks and practices have been put in place for paper-based information in response to the strict controls of government regulatory agencies, but these have yet to be transferred into the digital environment. Factors restricting this expansion include the large range of complex data types found in this sector, questions of responsibility for action, and the vast amount of resources required to execute a suitable and effective digital preservation solution.

Activity to date has been focussed in persuading software and system suppliers to create suitable solutions that can be installed at pharmaceutical organisations. Some organisations are more active than others having developed system specifications, while others remain rather more observant encouraging solutions to be provided for them. However, beyond the need for technical solutions to the problem, the sector is in need of a more comprehensive understanding of the problems in managing digital assets, as comprehensive strategies for long-term digital preservation are lacking. Increased co-operation and dissemination could be of benefit, as well as an increased confidence (which comes from better understanding of the problem) in order to secure a more stable future for the valuable digital assets at risk in this sector.

## **Chapter 1: The ERPANET Project**

The European Commission and Swiss Confederation funded ERPANET Project (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/Swiss Federal Archives), 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 J. McKinney serves as a co-ordinator to the project. An Advisory Committee with experts from various organisations, institutions, and companies from all over Europe gives advice and support to ERPANET.

## **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. It is anticipated that upwards of 500 organisations, institutions and public bodies will eventually contribute to this research. 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 sectors included, but also through the depth at which they will explore the issues.

ERPANET is deliberately and systematically approaching disparate sectors 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.<sup>1</sup> It is of primary importance that disparate groups are given a mechanism through which to come together as best practices for digital preservation are established in each sector.

### *Aims*

The principal aims of the study are to:

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

Potential sectors have been selected to represent a wide scope of information production and digital preservation activity. Each sector may present a unique perspective on digital preservation. Organisational and sectoral requirements, awareness of digital preservation, resources available, and the nature of the digital object created place unique and specific demands on organisations. Each of the

---

<sup>1</sup> 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.erpamet.org/www/content/documents/Digitalpreservationcharter4\\_1.pdf](http://www.erpamet.org/www/content/documents/Digitalpreservationcharter4_1.pdf)

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.

Ten organisations in each sector, and 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 focussed questionnaire designed to be targeted at a range of strategic points in the organisations under examination. The instrument<sup>2</sup> 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 the advisory committee.

---

<sup>2</sup> See Appendix. We include 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.

## **Chapter 4: Introduction to the pharmaceutical sector**

The pharmaceutical industry is highly research-intensive. Due to major advances in research and technology, scientists control all aspects of the manufacture of a new medicine, from research and design through to prototypes and final product development. This new medicine is then marketed and sold to the public.

If research yields a compound that has medical potential, a prototype medicine is developed, which involves conducting pre-clinical tests and clinical trials to ascertain the safety, quality, and effectiveness of the medicine. Before a medicine can be sold, the local authority agency has to approve and certify its safety, quality, and efficacy. During this process a vast amount of scientific data and documentation is produced.

### *Research investment and intellectual property protection*

The drug development process is a high-risk, expensive, and lengthy one. On average it takes 10 to 15 years to produce a new medicine – from the laboratory to approval by the USA Food and Drug Administration (FDA). Only 1 out of initially up to 10 000 screened compounds in the extensive stage of research makes its way through the pre-clinical testing, the clinical testing and the final stage of being approved as a new medicine. For these reasons the development of medicines is extremely costly with an average investment of US\$802 million for a new prescription drug.<sup>3</sup>

Biotechnology is increasingly attractive to the pharmaceutical industry as a new technology that could reduce the costs in drug discovery. The United States dominate the biotech world with eight of the top ten global biotech companies by market capitalisation. This is one of the reasons for the dominance of the United States in the pharmaceutical sector.

It is essential to have patent protection on promising medical compounds. Once such a compound is discovered, a company applies for a patent that is valid for 20 years from the date of application. After the patent has terminated virtually anyone can produce generic medicines<sup>4</sup> and in this way save substantial research costs, while profiting from an abbreviated approval procedure at the relevant governmental agencies.

The pharmaceutical industry is the largest research-performing industry world-wide. Expenditures in R+D are estimated to comprise 18'800 million Euro in 2001 in Europe alone (including Turkey).<sup>5</sup> As a direct result of the costly initial stage of drug research, the pharmaceutical sector has recently seen major mergers and the formation of alliances. Glaxo Wellcome and SmithKline Beecham merged in 2000 to form GlaxoSmithKline, in a deal worth US\$76 billion. In July 2002, Pfizer Inc. acquired Pharmacia Corporation, creating the largest pharmaceutical company in the world with anticipated annual revenues of US\$48 billion. As a result of the rapid and continued growth of the sector, and the costly nature of the industry itself, it is expected that the pharmaceutical industry will continue to form huge multinational companies.

---

<sup>3</sup> PhRMA Industry Profile 2002; Chapter 1, page 6 (refer to the 'Resources' chapter)

<sup>4</sup> A 'generic medicine' is a medicine that is identical, or bioequivalent, to an already approved medicine

<sup>5</sup> EFPIA Industry in Figures 2002; "The pharmaceutical industry: a key asset to the European economy", page 9 (refer to the 'Resources' chapter)

### *Agency regulations*

The pharmaceutical industry is highly regulated by multinational bodies, such as the World Health Organisation (WHO), as well by national governmental agencies, such as the Medicines Control Agency (MCA) in the UK and the Food and Drug Administration (FDA) in the USA. These agencies were established to protect public health and to ascertain that human and veterinary medicines are safe and effective. For this reason these agencies screen every aspect of drug development and production before they can provide market approval.

The procedures for complying with agency regulations demand an enormous amount of effort, information, and time. New medicine submissions can be 50,000 to 250,000 pages in length and it normally takes the agencies more than a year to process and approve of a new medicine<sup>6</sup>.

While attempting to meet the same objectives, regulations between the various national agencies differ. German regulations, for example, are stricter on the retention of documentation on medicines than US regulations, stipulating that it must still be accessible 30 years after the product has been taken off the market. The FDA<sup>7</sup> demands documentation be accessible 5 years after the product has been taken off the market.

In 1997 the FDA issued regulations intended to permit the use of electronic technology at pharmaceutical companies. '21 CFR Part 11' – where CFR stands for Code of Federal Regulations – is the specific regulation providing criteria under which the use of electronic records and electronic signatures is consistent with the responsibility of the FDA to promote and protect public health. Submissions to the FDA in electronic form are currently on a voluntary basis.

### *Harmonising the international regulatory landscape*

In order to recoup the enormous investments in the research and development of medicines, an obvious step for pharmaceutical companies is to sell them on multiple national markets. Consequently, pharmaceutical companies have to comply with all the different regulations in their respective countries. The differences in the regulations of the various authorities and bodies worldwide, however subtle they may be, result in an enormous increase in effort, often duplicated, for the pharmaceutical companies. Occasionally these differences will demand a repetition of trials.

The International Conference on Harmonisation (ICH) was formed to tackle this issue by harmonising the technical requirements for registration of pharmaceuticals for human use. Initially created and led by authorities and organisations in Europe, the USA, and Japan, it now has the support of organisations such as the World Health Organisation. Three expert working groups, Efficacy, Quality, and Safety, consisting of a mix of regulators and industry representatives, work together towards creating a series of globally acceptable guidelines.

One outcome of this combined effort is the Common Technical Document (CTD), an interface for industry to agency transfer of regulatory information. The CTD describes the organisation of modules, sections, and documents necessary for the registration of new medicines. It covers all parts of the registration application common to the

---

<sup>6</sup> In 2001 it took the FDA an average of 16.4 months

<sup>7</sup> Due to the influential position of the FDA among the variety of governmental agencies, this report will predominantly make reference to the specific regulations of the FDA

various participating countries. The CTD considerably relieves expenditure when submitting newly developed products. Implementation of the CTD is already in place on an optional basis and will be mandatory from July 2003 for agencies in the EU, Japan, Canada, and Switzerland, and at that date implementation is also highly recommended for the FDA in the USA.

The specification of the electronic Common Technical Document (eCTD) lists the criteria that will allow the electronic transfer of the registration application from industry to regulatory authority. Besides initial registration applications, other submissions of information throughout the lifecycle of the product, such as variations and amendments, are specified by the eCTD and it extends the CTD in these areas. The eCTD promotes and utilises open standards and provides mechanisms for updating or adjusting the CTD to new technologies and criteria. Although paper submissions continue to be accepted, the authorities of the participating countries have committed to allow electronic submissions in the eCTD format from 2005.

## **Chapter 5: Details of interviews**

Various organisations in the pharmaceutical sector were approached for this case study. A short profile of those participating organisations is outlined below. From the majority of companies several employees were interviewed in order to obtain as comprehensive an impression of the current state of digital preservation activities as possible. Unfortunately, this was not possible for all companies, and at two it was only possibly to interview former employees and consultants.

Apart from those listed below, a number of other pharmaceutical companies were approached. Many stated that they were unable to participate because of a lack of experience with the issues. This was the case for many smaller companies focussing on a specific pharmaceutical field (such as mental health or cardiovascular disease), and who sold their products only in a limited number of countries.

### **Astra Zeneca PLC**

<http://www.astrazeneca.com/>

In 1999 Astra AB in Sweden and Zeneca Group PLC merged to become AstraZeneca PLC, in a deal worth US\$35 billion. In 2001 the organisation's sales totalled US\$16.5 billion, with an operating profit of US\$4.2 billion. AstraZeneca employ more than 54,000 workers world-wide. With sales in over a hundred countries, manufacturing sites in twenty, and major research centres in five, the company has a substantial international presence. The corporate headquarters of AstraZeneca are located in London UK, and the research and development base is in Södertälje, Sweden.

AstraZeneca spent more than US\$2.7 million in research and development in 2001. Besides its own resources, collaborations with universities and strategic alliances with research and biotechnology companies give AstraZeneca broad access to technologies and biomedical research.

### **Aventis SA**

[www.aventis.com](http://www.aventis.com)

Aventis was formed in 1999 as the result of a merger between Hoechst AG and Rhône-Poulenc S.A. Globally, Aventis has approximately 71'000 employees. The corporate headquarters are in Strasbourg, France. Its core business sales amount to about €17.59 billion in 2002, which include prescription drugs, human vaccines, as well as an animal health joint venture with the company Merck & Co.

In 2002, Aventis invested €3.14 billion in research and development world-wide. Additionally, Aventis is actively pursuing alliances and partnerships to further enhance internal capabilities.

### **Bayer AG**

[www.bayer-pharmaceuticals.com](http://www.bayer-pharmaceuticals.com)

Although founded more than one hundred years ago, Bayer has maintained a dynamic company profile with numerous acquisitions and reorganisations. Bayer's

business rests on four pillars: HealthCare, CropScience, Polymers, and Chemicals. The Bayer Group is represented on five continents employing 117,000, and with sales of more than €30 billion in 2001. The pharmaceuticals section as part of the HealthCare division has about 18,000 employees and recorded sales of around €4.8 million in 2001.

Expenses of €1.15 million and 3,880 employees are assigned to research and development in the pharmaceutical section, while the research platform of the Bayer Group employs in total around 18,000 employees.

### **Boehringer Ingelheim**

[www.boehringer-ingelheim.com](http://www.boehringer-ingelheim.com)

Founded in 1885 by Albert Boehringer, the business scope of Boehringer Ingelheim now embraces the Americas, Europe, and Asia-Pacific. Ninety-five per cent of the Boehringer Ingelheim group of companies concentrate on human pharmaceuticals with the remaining five per cent in animal health. In 2001 net sales of the company were roughly €6.7 billion with around 28,000 employees.

Boehringer Ingelheim invested more than €1 billion in research and development in 2001 and sets out to integrate key enabling technologies including genomics, proteomics, and combinatorial chemistry.

### **GlaxoSmithKline plc (GSK)**

[www.gsk.com](http://www.gsk.com)

GlaxoSmithKline was formed in 2000 in a US\$76 billion deal, after a series of mergers and acquisitions. With its headquarters in the UK the company is one of the industry leaders, with an estimated seven per cent of the world's pharmaceutical market. GlaxoSmithKline has over 100,000 employees world-wide. In 2001, sales were US\$29.5 billion, including pharmaceutical sales of US\$24.8 billion.

The organisation has over 16,000 research and development employees based in seven different countries. The company spends about US\$3.8 billion on research and development, and has a leading position in genomics/genetics and new drug discovery technologies.

### **Organon NV**

[www.organon.nl/](http://www.organon.nl/)

Organon is a subsidiary of Akzo Nobel. Besides Organon, Akzo Nobel's Pharma Group comprises the veterinary healthcare company Intervet and pharmaceutical ingredient manufacturer Diosynth. Organon's sales, however, make up to 61 per cent of the total pharmaceutical business unit sales of the parent company Akzo Nobel. In 2001 Organon had more than €2.4 million Euro in sales. Based in the Netherlands, Organon employs around 13,000 people in more than fifty countries.

With 2300 research employees at five centres world-wide, Organon invests 17.5 percent of its annual sales income in research and development. Organon is open to strategic alliances and partnerships, and devotes 30 percent of its research budget on external collaborations.

**Pfizer**

[www.pfizer.com](http://www.pfizer.com)

Pfizer recently announced its merge with Pharmacia. This combination will create the worlds leading pharmaceutical company. Pfizer already held the leading position on the US market, and together with Pharmacia it will also hold this position in Europe, Japan and Latin America. Pfizer alone had US\$25.5 billion in pharmaceutical sales in 2001 and employed 90,000 people world-wide. Its products are available in more than 150 countries.

In 2001 Pfizer had a budget of US\$ 4.8 billion for research and development. As part of this commitment, Pfizer sustains alliances with more than 250 partners in academia and industry.

## **Chapter 6: Circumstances**

The pharmaceutical industry is a hugely complex sector, difficult to penetrate for research purposes. This has had an impact on the information made available for this report.

The complexity of the sector is reflected in the highly specialised information gathered from the variety of agencies, organisations, conferences, and consultancies. The structure of the pharmaceutical companies themselves is often disparate, due to their international nature and magnitude. Even though pharmaceutical companies are well organised and managed, it is difficult to have an overview over all the branches and activities within a company, even as an employee.

Interviewees for this survey approached the topic from very different perspectives and on digital preservation. Not all were informed about digital preservation activities within their own company and as a result provided somewhat inconsistent information. This made it difficult to gain a clear picture of how digital preservation is approached in the company as a whole, but did provide an insight into organisational application and dissemination.

Pharmaceutical companies are often the targets of case studies and have procedures in place to deal with such requests. Not all companies approached, however, consented to participate. Many organisations, especially smaller ones, refused to co-operate because of the topic, necessary resource commitments, or simply because of a lack of interest.

Unfortunately, the FDA's Part 11 Industry Guidance on Maintenance of Electronic Records was only published after most of the interviews were conducted, thus the interviewees could not comment on this influential document. This draft guidance reflects FDA's current thinking on appropriate principles and procedures for maintaining electronic records, and forms part of a series of guidelines that focus on the regulations of "21 CFR Part 11"<sup>9</sup>. It will be especially important to those pharmaceutical companies that want to sell their products on the US market. It offers an interpretation of the respective regulatory framework and provides hands-on approaches to address the preservation of electronic records in the context of the pharmaceutical industry.

---

<sup>8</sup> cf. Chapter 1 – Introduction into Sector

<sup>9</sup> FDA, Regulatory Submissions in Electronic Format – Guidances (refer to the 'Resources' chapter)

## **Chapter 7: 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 pharmaceutical industry is generally aware of the problems surrounding digital preservation. Employees working in the research and development departments are familiar with the mass of data that is produced during the life cycle of a scientific project and the need for this to be preserved.

#### *Main Problems*

External and internal demands combine to produce a variety of problems for pharmaceutical organisations attempting to preserve their digital assets. Among the various problems mentioned in the questionnaires, the most important ones are listed below:

- Comprehensive preservation requirements imposed by national governmental agencies;
- Allocation of responsibility for the development of solutions;
- Management of vast quantities of a variety of data and information types;
- Lack of off-the-shelf solutions.

#### *Asset Value and Risk Exposure*

The principle motivation for digital preservation activity in pharmaceutical organisations is the legal requirement to preserve specific information for a prescribed period of time. If companies do not comply with agency regulations, drug sale might be prohibited, at least for the national market where regulators confirm a violation of the national law. This of course may endanger redeeming the substantial investment<sup>10</sup> in the research and development of the very medicine. In addition to this, digital information is preserved to utilise its reuse value.

Reuse is primarily an issue for data created and gathered during research and development. It is generally agreed that the final reports of this work must be retained. This is in the financial interest of each pharmaceutical company, since

---

<sup>10</sup> see the introduction on the sector, Chapter 4

these intellectual assets add to the company's actual value. Reuse of scientific data has generated two directly opposing opinions. One view is that the preservation and subsequently reuse of research data is of little value. Due to the high specificity of research and the fast pace of technology, experiments should be repeated to obtain more accurate results. In addition, data generated by scientific machinery is very fragile as each supplier and the actual scientific machine may have its own specific data type. These types are particularly volatile, with software updates being released up to two times a year. This makes it more difficult to preserve and reuse this type of scientific data. The other view takes the position that this data constitutes a substantial asset of the pharmaceutical industry. Repeating experiments would significantly impede progress in research. Experiments and data can provide answers to questions that were not posed in the initial scope of the project. Experiments can raise new questions triggered by an unexpected discovery. In addition, there may be successive generations of scientific equipment for research and development with an equal quality, or paradigm shifts may cause equipment in a transitional stage to be no better than the former generation. Therefore, repeating experiments with other, newer equipment does not necessarily yield better results.

Both of these views have valid arguments to substantiate their claims, and it is unlikely that the debate will be resolved in the near future. Only individual company culture and policy framework will dictate whether or not its scientific and research data should be preserved and reused.

Preservation solutions and strategies are still in the early stages of development for many of the organisations, and their success is not guaranteed. The risk involved in a commitment to a specific course of action proved to be of some concern translating into uncertainty and unwillingness to commit resources.

Companies are increasingly turning to the use of digital information, as it simplifies and accelerates all company processes. It can be expected that governmental agencies will be able to accelerate the approval of medication if the registration data is available digitally. Smaller companies, however, may not have the resources to install the necessary systems and structures to be compliant with the regulations, and consequently they will continue to submit paper-based registration for medication, creating serious competitive disadvantages for smaller companies.

Some organisations identified historical value in their assets and have preserved some paper-based records to reflect this. This activity, however, is not a high priority, and whether this will be extended to include digital resources, however, remains to be seen.

### *The Regulatory Framework*

The requirements imposed by national governmental agencies such as the FDA form the main challenge for the preservation activities of pharmaceutical organisations. FDA Regulations demand that information about medicines is preserved for the length of the expiration (batch expiry date) of the last product on the market, plus an additional five years. In practice this might well be a period of fifty years or even longer.

The data that is legally required to be preserved can be divided into two sections. Firstly, all information necessary for the registration of new medicines, including a comprehensive documentation of the research and development, must be retained. Secondly, documentary data must be recorded and preserved during the

manufacturing process. In practice, this can generate hundreds of thousands of paper pages.

The regulations concerned with laboratory studies are formulated in the Good Laboratory Practice (GLP). A GLP was first introduced by the FDA. Subsequently a compilation of GLPs was published by the OECD<sup>11</sup> and other countries have adopted local GLPs with minor deviations. Similarly, a standard for the design, conduct, performance, monitoring, auditing, recording, analysis, and reporting of clinical trials has been developed as the Good Clinical Practice (GCP).

Once a medicine is approved for the market, agencies have to ensure that the medicines are properly manufactured. This is enforced by inspections at the manufacturing sites, which are conducted regularly (annually by the MCA<sup>12</sup>, for example). Current Good Manufacturing Practices (cGMP) by national agencies regulate that products are consistently produced and controlled to the appropriate quality standards. In order to comply with the relevant regulations and as a compilation of the company's specific practices, companies establish Standard Operating Procedures (SOP), which are protocols and procedures for a laboratory's program.

Due to governmental agencies' broad requirements of preserving data, the digital objects that need to be retained include research documentation, which concerns not only the final but also intermediate reports, as well as the original data created by scientific equipment as part of experiments. The scope of digital objects is very broad, from simple Office documents to SAS<sup>13</sup> applications used for clinical data analysis and raw scientific data preserved various formats.

### **Preservation Activity**

Organisations direct their preservation activities in response to the attributed value of the assets they hold and the risks that these and the organisation in are faced with. A drive towards action is indicated through the development and recommendation of policies and strategies and the practical activities of selection, preservation, storage, and provision of access to the organisation's valuable assets.

#### *Policies and Strategies*

Although at senior management level problems concerning digital preservation are a concern, there does not appear to be a cohesive company strategy as evidenced by the existence of only a few digital preservation policies in place. This may in part be due to a cautious management attitude in response to concerns regarding the costs and risks involved in digital preservation. Some companies plan to develop preservation policies as part of projects to articulate system specifications, but as such no policies have been developed. This lack of formal organisational position articulating the assets at risk and the strategies and solutions developed and recommended to mitigate these risks increases the difficulty of understanding an organisation's preservation activities.

---

<sup>11</sup> OECD, Organisation for Economic Co-operation and Development. <http://www.oecd.org/>

<sup>12</sup> MCA, Medicines Control Agency in the United Kingdom. <http://www.mca.gov.uk/>

<sup>13</sup> The company SAS (<http://www.sas.com/>) provides information technology solutions, and is essentially the biggest provider of scientific analysis applications in the pharmaceutical sector.

### *Selection*

The criteria on what data to select and retain is largely based on agency regulations. Such guidelines are created by the departments responsible for organisational records management and disposition, and are subsequently implemented corporate-wide. There are clear selection policies and operating procedures for the paper-based records of the organisation. These selection criteria can largely be transferred to the digital world.

There are retention schedules in place for paper-based records at some of the pharmaceutical companies. However, some interviewees stated that this should be addressed in more depth, as management plans for the retention of digital assets has not yet been formulated. Document management systems in place currently lack the capability to assign digital objects to a specific schedule and to dispose of them accordingly at a certain point in time. A possible reason for this could be that costs for their physical storage have dropped and will likely continue to do so. As a result some pharmaceutical companies do not consider it worth the effort to filter and select specific digital documents as it is cheaper to retain all of the data instead.

### *Preservation*

One of the challenges for long-term preservation in the pharmaceutical sector is the vast amount of data that is created and that needs to be managed. Traditional documentation consists of millions of paper pages for a single medication. With the increasing usage of digital information, an enormous amount of digital objects need to be retained.

Unsurprisingly, it is the larger multi-national pharmaceutical companies that are actively developing solutions for digital preservation. Projects have been set up to analyse and specify preservation requirements, and system building is outsourced to external vendors. Recognising the complexity of the problems involved, such plans have been cautiously set up and tend to evolve gradually.

Document management systems have been purchased and installed at all the large companies, while smaller companies are increasingly investing in these information management tools. In pharmaceutical research and development, where information can stem from myriad different sources, these offer the possibility of integrating digital objects for convenient and efficient use, and to structure the enormous amounts of information to be processed. Pharmaceutical organisations do not wish to develop technical solutions to manage long-term preservation themselves, and it is felt that these tools should be developed by external, specialised suppliers. Presently there are no off-the-shelf solutions available that cover the specific digital preservation needs of the pharmaceutical sector.

Other organisations have chosen to wait and see what develops in this area. However, suppliers will only produce solutions in response to specifications compiled by other pharmaceutical companies. Despite the differences between the various pharmaceutical companies, there is a certain commonality in the requirements for long-term digital preservation. Though these organisations are well aware of the necessity and urgency to undertake preservation activities, they consider the problems too big for them to tackle by themselves, and the associated costs and risks too high.

Handling the sheer size of the data is largely a question of whether storage facilities can cope with the volume of records. The companies therefore have to legally supply

the finances to acquire sufficient storage. The different formats that the data can be stored as also has to be taken into consideration, ranging from MS Office formats to more fragile formats specific to scientific equipment.

Two of the interviewed companies have extended their document management systems so they are capable of fulfilling certain archival tasks. In contrast to this, four of the companies set up or plan to set up completely separate facilities, with a document management system on the one side and an archive on the other. Organisation of the archive and document management is designed to remain closely linked. It should, for example, be transparent to the user whether a document is still active. Once a document is finalised and captured into the archive, the document is automatically transferred to the physical facilities of the archival system.

The design of data storage facilities varies greatly amongst pharmaceutical companies. Half of the interviewed companies have established, or plan to establish, a centralised archive. Necessary security precautions are taken for these facilities, which include storing the data on backup media and at mirror sites that are at geographically different locations. Redundant storage is carried out not only for security reasons, but to increase availability of the data and the performance of the system.

The other half of the interviewed companies use a distributed system that may consist of a mixture of different operating systems and hardware technology. These distributed systems address the security of the data, but involve a comparably bigger organisational effort and strategies can vary between the different locations.

In both cases, and for all of the organisations questioned, the data is currently kept online on the hard disks of storage servers. Previously, it was significantly cheaper to store all archived data offline on external storage media such as magnetic tapes. Consequently, there are decommissioned archives at several pharmaceutical companies that used to save all their data on external, optical media. Nowadays, the price for storing the data online is comparable to offline (or nearline) storage systems. Since pharmaceutical organisations are very cautious about the safety of their data, these hardware facilities are preferably located on their own premises. The maintenance and day to day running of this service is mostly outsourced to third parties.

### *Migration*

Some of the more proactive companies have led the way by taking the pragmatic approach of migrating their data to new formats when the necessity arises. PDF is now becoming the standard format of choice. There are still large amounts of data in TIFF format, but companies are in the process of converting these objects to PDF. For example, one company has just finished migrating legacy data from TIFF formats to PDF. This method, however, is followed on an ad-hoc basis – it is not yet expressed in organisational policies or in a long-term digital preservation strategy. To minimise risk, companies retain the digital objects in their original format alongside the converted version. The additional storage capacity this requires is not an issue at this point of time. One company even retains the TIFF version besides the original and the PDF format of the digital object. This is done since TIFF is a relatively simple format, and given the format specification a TIFF reader can be developed easily. Although the FDA does not require that digital objects are retained in their original form, it does demand that the "ability to process an electronic record's information

throughout its records retention period should be preserved"<sup>14</sup>. A major argument for adopting PDF is the fact that agencies accept it for submissions and, generally, it has evolved to be the most prevalent format. To increase the stability of the PDF format, some companies restrict the use of its special features. Specific digital objects, however, can not be represented in PDF format. These include Oracle databases as well as data in the SAS format. A long-term preservation strategy for these types of format remains yet to be addressed.

In the future XML might become an important preservation format and companies are alert to its development. This does not necessarily mean that XML will replace PDF as an industry format, but rather that the two formats would be used in a complementary fashion with XML for structural information and PDF for content and pictures. For the time being, it is generally felt that there are not enough tools available for XML, and so it has yet to be commonly adopted. At best there should be consent on standard formats. Suppliers of scientific equipment should be encouraged to use those formats. However, many indicate a hesitancy towards this as it would demand time, money, and effort to agree on a standard. In their opinion, conversion to other formats can be done automatically within individual organisations to meet their own needs, and would be a minor effort and cost in comparison to that needed for the standardisation of scientific equipment and technical machinery necessary in research and development in the sector.

### *Metadata*

Metadata forms a significant part of any data management and archival system. Departments responsible for preservation have already established metadata frameworks and indexes for paper documentation. Metadata includes information referring to the agency regulation the document is subject to. Most of this metadata is created by the user with additional information added for document management and preservation purposes. The metadata sets that are applied in a paper-based context will be used as a basis for metadata sets in a digital context, but must be enhanced with additional fields necessary for the management of the document in the digital environment. This includes information on the file format, the operating system, and the application with which it was created, as well as a traceable history of the document, and possibly checksums or encrypted signatures to verify the integrity of the digital object. A large amount of metadata is required in a pharmaceutical context. One company stated that they have over 100 metadata fields, and the metadata set is not yet finished. With such a large number of metadata fields to be captured, it will be necessary to generate the major proportion in order to increase the likelihood of coherent and consistent metadata for any given digital object.

### *Access*

Provision of access to relevant information and related contextual information was investigated. Retrieval of digital information was considered an essential part of the preservation process. It proved useful to discover what methods of control were in place to protect assets from unauthorised use, alteration, or destruction.

There is broad consensus among the pharmaceutical companies that global access to stored information should be possible for employees, and that tools for convenient and efficient reuse of data should be installed. Access and accessibility, which includes creating indexes and installing adequate tools for retrieval of the archived

---

<sup>14</sup> FDA, Draft Guidance for Industry: 21 CFR Part 11 – Maintenance of Electronic Records

digital objects, will have to be addressed by the departments responsible for the archives as part of the development of system specifications.

Pharmaceutical companies are understandably cautious about who has access to their repositories, and the valuable information contained in them. They are also legally obliged to guarantee the integrity of the digital objects that contain all relevant information about their medication. Well-designed and tested procedures are required for rights management, so that only authorised personnel restricted to particular domains can retrieve digital objects. If employees are not authorised to view specific objects, they will be able to search only descriptive information, such as the author, the title, or keywords. Once a relevant digital object has been identified, they have to ask permission to access it. Another issue in this respect will be managing the secrecy level of classified digital objects that could change over time. Copyright does not pose a major concern to pharmaceutical organisations as the majority of digital objects are internal corporate ones. Privacy issues do arise in the clinical environment, where personal data about patients is commonly retained. The procedures that are or have been in place formerly are considered sufficient to guarantee privacy also in a digital context.

### **Compliance Monitoring**

There is a need to monitor preservation processes at pharmaceutical companies. Paper-based records must be compliant with the regulations of governmental agencies and Standard Operating Procedures (SOPs) have been formulated for selecting the material, enriching it with metadata, and handing it to the department responsible for the archives. The archives take all necessary steps to retain the paper material and organise it so that it can be retrieved when needed. While these processes may be applicable in a digital context, this transfer remains to be undertaken. As yet, the organisations questioned have no framework for monitoring actions specifically concerning digital preservation. This is due to the lack of formal policies and strategy towards the long-term preservation of digital assets. Pharmaceuticals organisations are, however, working to overcome this. Once policies, strategies, and subsequently SOPs have been formulated, training for employees will be organised and a compliance framework put in place.

Detailed external inspections by governmental agencies may take weeks, and FDA inspectors will visit more than 15,000 facilities a year. The FDA normally encourages correction of any problems that might be found. If this is not done, the FDA can impose legal sanctions. After pharmaceutical companies have overcome the current transitional phase from paper-based to digital processes, the regulatory agencies can be expected to scrutinise carefully the company's procedures to ensure digital preservation. The FDA's CFR 21 Part 11 requires that all necessary steps are taken "to ensure the authenticity, integrity, and confidentiality of electronic records"<sup>15</sup>. As part of guaranteeing this, digital objects must be managed and retained in controlled processes throughout their existence – their history must be documented, and their integrity ensured. For this reason all events that have an impact on an archived object have to be documented accurately in a traceable history, and this audit trail has to be preserved throughout the existence of the digital object. In addition, computer systems housing digital objects have to be validated<sup>16</sup> to comply with

---

<sup>15</sup> FDA, CFR 21 Part 11, § 11.10

<sup>16</sup> FDA, Draft Guidance for Industry: 21 CFR Part 11 – Validation

agency regulations, which refers to the “confirmation by examination and provision of objective evidence that computer system specifications conform to user needs and intended uses, and that all requirements can be consistently fulfilled”<sup>17</sup>.

### **Digital Preservation Costs**

Pharmaceutical companies intend to fund the installation of digital preservation facilities as required by the respective authoritative agencies. However, since the solutions for digital preservation are still in a relatively early stage of development, pharmaceutical organisations are hesitant to assign personnel and financial resources to this area of concern. The costs and risks involved in developing solutions for digital preservation are considered substantial.

Interviewees in this case study did not provide extensive details on how much money is at disposal for digital preservation in their companies. At some companies the research and development divisions finance digital preservation activities, while at other it is considered a corporate investment.

Different basic investment characteristics between organisations can be distinguished. Organisations that set up dedicated projects to address digital preservation activities will have specific resources allocated to them. This scenario is more common with larger organisations that have already invested considerably in digital preservation. These projects are designed to proceed in smaller incremental steps with fixed milestones reflecting the rather cautious attitude towards developing solutions. On the other hand, those companies actively implementing archival functionality, and which employ an 'ad-hoc' migration strategy, are confident that the money will be available if it turns out to be necessary to conduct another migration step. Whether it is necessary will be considered on the basis of a cost-risk analysis at that point in the future, most likely if there is a change in government regulations.

### **Future Outlook**

When asking interviewees about their view on the current state of digital preservation, the efficiency of organisational structures in place, available funding, and the adequacy of approaches and systems, several issues were raised.

For some companies digital data is currently stored in various locations, in separate systems, and in different departments. As this requires increased resources, budget, and time, there is a demand for a more global, integrated solution. A more global solution might also lower the organisational overhead that adds up with each local system and may, consequently, lead to lower costs. In additions to this, is felt that scientists in the local departments should not have to be concerned about managing and maintaining information and should be allowed to concentrate their efforts in their area of expertise. The hope is that tools can be put in place to enable this management responsibility to be transferred away from the data creators and users.

---

<sup>17</sup> FDA, Draft Guidance for Industry: 21 CFR Part 11; Electronic Records; Electronic Signatures – Glossary of Terms. <http://www.fda.gov/cber/gdlns/esigglos.htm>

Awareness about digital preservation might be increasing, but the perception of the problem remains only the very step. Processes and practices that support the preservation of the digital material must be promoted in all areas and situations throughout the company. This starts at the creation of the documents, where digital objects should be created in a standard format. It will be important to organise seminars and training at the pharmaceutical companies to teach employees in applying the appropriate practices. While this occurs in a few companies, this can still be improved. Furthermore, management needs to assign the required resources and time to employees in order to support digital preservation.

Certain interviewees indicated that communication with external sectors could be beneficial since new ideas and different perspectives would contribute to a more efficient process for developing solutions. More intense consultation with other sectors could create global systems and infrastructures for digital archiving and preservation built on a common basis.

## **Chapter 8: Conclusions**

There is a significant level of awareness of digital preservation issues in the pharmaceutical sector, manifested predominantly in the concerns exhibited by organisations attempting to comply with the regulations imposed on them by government agencies while responding to market pressures to move to more efficient and effective digital processes. Solutions to preserve relevant documentation for the length of time required by governmental agencies are urgently needed. Efforts to attain solutions focus on the field of pharmaceutical research and development, as this is where regulations are most pressing and where significant challenges exist due to the amount of scientific data produced and the range of formats in which it is produced. An adequate strategy for digital preservation and a long-term commitment to the associated tasks and responsibilities still remain to be established. The main impediment to action is the concern that to develop and implement an adequate system for digital preservation would be extremely costly in terms of time, human resources, and most important, money.

Despite this, the sector is responding to these demands, primarily by putting pressure on software suppliers to offer adequate solutions, and data management and preservation systems are gradually being installed. Solutions need to be supported from the top of the company, and so management needs to commit to sustaining and furthering digital preservation activities. It will therefore be important for pharmaceutical companies to develop clear policies and a comprehensive digital preservation strategy geared for the long term.

While the pharmaceutical sector currently focuses on research and development with respect to the preservation of digital objects, strategies for digital preservation will have to be developed for other organisational functions. For example, manufacturing processes produce huge amounts of data. The adequate retention of this documentation is required by agency regulations as well.

Pharmaceutical companies should not attempt to solve the problems alone and in an isolated manner. Currently, co-operation within the sector is very low. Even though pharmaceutical companies hold very sensitive information, close co-operation could be beneficial. Collaborative action could be a way to attain more robust solutions in less time, and in the long-term lead to a reduction in costs. Similarly, communication and co-operation with organisations in other sectors is very low. Other sectors have had varying degrees of success with digital preservation, and evidence of success and failure as well as other findings may offer valuable transferable information for the pharmaceutical sector. A better understanding of digital preservation in the pharmaceutical sector in relation to other scientific sectors will emerge with the progress of ERPANET case studies.

The practical experiences of pharmaceutical companies will be important to all organisations with a stake in digital preservation. Due to the pressure of the market and agency regulations, the pharmaceutical sector can be expected to develop a pragmatic approach and solutions that other sectors may identify with.

## **Chapter 10: Resources**

### *Governmental agencies*

European Medicines Evaluation Agency (EMA), EU; <http://www.emea.eu.int>

Food and Drug Administration (FDA), USA – Centre for Drug Evaluation and Research; <http://www.fda.gov/cder>

Ministry of Health, Labour and Welfare (MHLW), Japan; <http://www.mhlw.go.jp>

Medicines Control Agency (MCA), United Kingdom; <http://www.mca.gov.uk>

Links to governmental agencies in Europe: <http://heads.medagencies.org>

### *Organisations*

European Directorate for the Quality of Medicines (EDQM), Pharmacopoeia Europe; <http://www.phEur.org/>

European Federation of Pharmaceutical Industries and Associations (EFPIA); <http://www.efpia.org/>

EUDRA, European Network in the field of human and veterinary pharmaceuticals; <http://www.eudra.org>

International Conference on Harmonisation (ICH); <http://www.ich.org>

Japan Pharmaceutical Manufacturers Association (JPMA); <http://www.jpma.or.jp>

Organisation for the Economic Co-operation and Development (OECD); <http://www.oecd.org/>

World Health Organisation (WHO); <http://www.who.int/medicines/>

Pharmaceutical Research and Manufacturers of America (PhRMA); <http://www.phrma.org/>

RegSource, gateway to global regulatory information; <http://www.regsource.com>

### *Regulations*

ICH, the electronic Common Technical Document (eCTD); <http://www.ich.org/ichectd.html>

FDA, Title 21 CFR Part 11 – Electronic Records/Signatures; [http://www.fda.gov/ora/compliance\\_ref/part11/](http://www.fda.gov/ora/compliance_ref/part11/)

FDA, Regulatory Submissions in Electronic Format – Guidances; <http://www.fda.gov/cber/esub/esubguid.htm>

FDA, Guidance For Industry – 21 CFR Part 11; Electronic Records; Electronic Signatures – Maintenance of Electronic Records. Draft, July 2002; <http://www.fda.gov/cber/gdlns/esigmaint.pdf>

FDA, Guidance For Industry – 21 CFR Part 11; Electronic Records; Electronic Signatures – Validation. Draft, August 2002;  
<http://www.fda.gov/cber/gdlns/esigvalid.pdf>

FDA, Office of Regulatory Affairs: Enforcement Policy: 21 CFR Part 11; Electronic Records; Electronic Signatures (CPG 7153.17). Compliance Policy Guide, Section 160.850. [http://www.fda.gov/ora/compliance\\_ref/cpg/cpggenl/cpg160-850.htm](http://www.fda.gov/ora/compliance_ref/cpg/cpggenl/cpg160-850.htm)

*General information on the pharmaceutical sector*

EFPIA – European Federation of Pharmaceutical Industries and Associations  
The Pharmaceutical Industry in Figures, 2002  
[http://www.efpia.org/6\\_publ/Infigures2002.pdf](http://www.efpia.org/6_publ/Infigures2002.pdf) (7.32 MB)

PhRMA – Pharmaceutical Research and Manufacturers of America  
Pharmaceutical Industry Profile, 2002  
<http://www.phrma.org/publications/publications/profile02/index.cfm>

US Business Reporter  
Pharmaceutical Industry Profile, January 2002 (revised)  
[http://www.activemedia-guide.com/pharmaceutical\\_industry.htm](http://www.activemedia-guide.com/pharmaceutical_industry.htm)

Margaret Sherp and Pari Patel  
Europe's Pharmaceutical Industry: An Innovation Profile: Summary EIMS Studies (European Innovation Monitoring System), January 2002 (last updated)  
<http://www.cordis.lu/eims/src/eims-r32.htm>

PriceWaterhouseCoopers  
Pharmaceutical Sector Insights – Analysis and Opinions on Merger and Acquisition Activity Annual Report, 2001  
<http://www.pwcglobal.com/pharmainsights>

UK Analytical Partnership  
Pharmaceutical Sector Report In: UKAP Good Regulation and Competitiveness Network Sector Studies, October 2001  
<http://www.chemsoc.org/pdf/ukap/pharma.pdf>

Alfonso Gambardella, Luigi Orsenigo and Fabio Pammolli  
Global Competitiveness in Pharmaceuticals – A European Perspective.  
Report prepared for the Directorate General Enterprise of the European Commission, November 2000  
<http://europa.eu.int/comm/enterprise/library/enterprise-papers/paper1.htm>

U.S. Environmental Protection Agency  
Profile of the Pharmaceutical Sector.  
In: EPA Office of Compliance Sector Notebook Project, September 1997  
<http://www.epa.gov/compliance/resources/publications/assistance/sectors/notebooks/pharmaceutical.html>

*Publications specifically on Digital Archiving and Digital Preservation in the context of the Pharmaceutical Sector:*

Lesley Richmond, Alison Turtin, Julie Stevenson  
The Pharmaceutical Industry: A Guide to Historical Records  
ISBN 0754633527; Ashgate Pub Co.; October 1, 2003

Philip Lord  
Preserving Scientific and Technical Records - Strategies and Solutions  
European Archives News, DLM 02 May  
[http://www.dlmforum2002.org/download/philip\\_lord.PDF](http://www.dlmforum2002.org/download/philip_lord.PDF)

Princeton Softech  
Archiving Complex Enterprise Databases. Improving Application Performance,  
Availability, and Reliability White paper, April 2001  
[http://www.21cfrpart11.com/files/library/data\\_management/archiving\\_%20complex\\_enterprise\\_databases\\_white%20paper.pdf](http://www.21cfrpart11.com/files/library/data_management/archiving_%20complex_enterprise_databases_white%20paper.pdf)

Philip Lord  
Moving forward with industry in electronic archiving: business process benefits,  
European Archives News, DLM 99 October

David Ryan  
5x3 to EAD, Society of Indexers - Annual Conference 1999  
Title: "What companies require is not only a digital nervous system but also a digital long term memory"

Jean Samuel  
The Intellectual Infrastructure Underpinning Records Management within Pfizer  
Central Research  
<http://www.aslib.co.uk/rmj/1998/apr/02.html>

David Bowen  
Practical issues in implementing a central electronic archive, European Archives  
News, DLM 96 December

Alan Murdock  
Roles and responsibilities in managing an electronic archive, European Archives  
News, DLM 96 December

Philip Lord  
Strategies and tactics for managing electronic data records: a view from the  
pharmaceutical industry, European Archives News, DLM 96 December

Ulf Andersson  
Short version of the Sesam report - Philosophy and rules concerning electronic  
archives and authenticity, European Archives News, DLM 96 December

Astra - Finance, Information Systems+Technology, Information Systems; Ulf  
Andersson  
SESAM Philosophy and rules concerning electronic archives and authenticity  
28. Feb 1996  
<http://www.si.umich.edu/e-recs/Sesam0/>

## **Appendix: Interview Instrument**

## ERPANET Case Study

### 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?

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)

16. Do these policies apply across the entire organisation?

17. How are these policies implemented?

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

- Yes
- No

If YES, Please specify.

---



---

---



---

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

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)

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

**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?

- Yes
- No

If YES, 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**

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

- Outsourced
- In-house

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

---



---



---

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

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?

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

- Yes
- No

If YES, Please specify.

---



---

- 
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.

<i>Technique</i>	<b>Specify Type/Technology Used</b>	<b>Information Preserved</b>
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
- Others

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 others)? 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) does 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. Others (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.

## **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  
Samir Musa

[www.erpanet.org](http://www.erpanet.org)