

# erpastudies

**marine accident  
investigation  
branch**

---

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

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

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

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

## **Table of Contents**

<b>Executive Summary.....</b>	<b>2</b>
<b>Chapter 1: The ERPANET Project .....</b>	<b>3</b>
<b>Chapter 2: Scope of the Case Studies.....</b>	<b>4</b>
<b>Chapter 3: Method of Working .....</b>	<b>6</b>
<b>Chapter 4: The UK Marine Accident Investigation Branch.....</b>	<b>7</b>
<b>Chapter 5: Details and circumstances of the interviews.....</b>	<b>8</b>
<b>Chapter 6: Analysis .....</b>	<b>9</b>
Perception and Awareness of Digital Preservation .....	9
Preservation Activity .....	11
Compliance Monitoring.....	14
Digital Preservation Costs .....	14
Future Outlook.....	14
<b>Chapter 7: Conclusions .....</b>	<b>16</b>
<b>Appendix 1: References.....</b>	<b>18</b>

## **Executive Summary**

The Marine Accident Investigation Branch (MAIB) of the UK Department for Transport (DfT) investigates accidents befalling UK ships or concerning UK waters. MAIB are required to preserve many different types of records relating to marine accident investigations for 25 years before possible permanent preservation by the UK National Archives. They are familiar with the basic challenges of digital preservation and maintain a hybrid record-keeping system with physical artefacts, paper, and digital records.

MAIB's digital preservation strategy is driven by a combination of Departmental record-keeping policy, which is partially based on retention guidance issued by the National Archives, and the activities of Branch management and IT staff. Their approach was developed in-house, although external parties are engaged for some activities such as database migration. The approach is based largely on migration to de facto standards but is still in the early stages of development and is hampered by the absence of specialist digital-archiving and -preservation advice. Preservation of evidence relating to marine accidents is of particular concern, due to the broad array of data formats the evidence can be stored in and the absence of a standard for complex data such as the industry specific 'black box'. However, awareness of digital preservation issues is growing throughout the Branch and they are keen to develop their current practices into a more systematic, long term strategy.

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

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

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

---

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

## **Chapter 2: Scope of the Case Studies**

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

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

This cross-sectoral fertilisation is a main focus of ERPANET as laid out in its Digital Preservation Charter.<sup>2</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

---

<sup>2</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.erpanet.org/www/content/documents/Digitalpreservationcharter4\\_1.pdf](http://www.erpanet.org/www/content/documents/Digitalpreservationcharter4_1.pdf).

object created place unique and specific demands on organisations. Each of the case studies is being balanced to ensure a range of institutional types, sizes, and locations.

The main areas of investigation included:

- perception and awareness of risk associated with information loss;
- understanding how digital preservation affects the organisation;
- identifying what actions have been taken to prevent data loss;
- the process of monitoring actions; and,
- mechanisms for determining future requirements.

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

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

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

### **Chapter 3: Method of Working**

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

The ERPANET interview instrument takes account of the strengths and weaknesses from both, developing a more focused questionnaire designed to be targeted at a range of strategic points in the organisations under examination. The instrument<sup>3</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 an advisory committee.<sup>4</sup>

---

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

<sup>4</sup> See [www.erpanet.org](http://www.erpanet.org) for the composition of this committee.

## **Chapter 4: The UK Marine Accident Investigation Branch**

Marine Accident Investigation Branch: <http://www.maib.dft.gov.uk/>

The Marine Accident Investigation Branch (MAIB) was established in 1989 as a unit of the UK Government Department for Transport (DfT)<sup>5</sup>. Located in Southampton, England, MAIB examines and investigates all types of marine accidents regarding UK ships worldwide and other ships in UK waters. The Branch employs 32 staff, including four teams of accident investigators and administrative staff to deal with records, data analysis, publications and other support services.

MAIB investigate accidents to determine circumstances and cause, with the aim of improving safety at sea and the avoidance of accidents in the future. It is not their position to apportion liability or blame outside of the fundamental purpose of the investigation, and they do not enforce laws or carry out prosecutions. The framework for reporting and investigating accidents, and the powers of the MAIB inspectors, are laid out in the UK Merchant Shipping Act of 1995 and effected through the UK Merchant Shipping (Accident Reporting and Investigation) Regulations of 1999. These regulations are the foundation of MAIB's work, defining the requirements for reporting accidents and making broad provision for the ordering, notification, and conduct of investigations.

MAIB must preserve records relating to their core activities – the investigation of accidents at sea – for a minimum of 25 years. These consist primarily of database records of marine accident incidents reported to MAIB; evidence collected during marine accident investigations carried out as a result of a reported incident (such as vital 'black box' data, video footage, or documentary evidence); and investigation reports containing conclusions and recommendations resulting from an investigation. With an array of evidence in potentially any format, an absence of black box data standards, and little sectoral guidance or practical experience with digital preservation, preservation of these materials when in digital form becomes a significant challenge. In an effort to address part of this problem, MAIB is working together with other nations to influence the International Maritime Organisation (IMO) of the UN<sup>6</sup> to develop a black box data format standard and policy. They are also members of the Marine Accident Investigators' International Forum (MAIIF)<sup>7</sup>, a non-profit organisation concerned with improving safety and decreasing pollution at sea.<sup>8</sup>

---

<sup>5</sup> UK Government Department for Transport (DfT) website: <http://www.dft.gov.uk/>.

<sup>6</sup> International Maritime Organisation (IMO) website: <http://www.imo.org/index.htm>.

<sup>7</sup> Marine Accident Investigators International Forum (MAIIF) website: <http://www.maiif.net/>.

<sup>8</sup> Whilst MAIIF maintains a manual and guidelines on data to be collected during investigations, their charter does not extend to the post-investigation maintenance or preservation of these objects when in digital form.

## **Chapter 5: Details and circumstances of the interviews**

ERPANET initially approached Ms Cathy Pennock, MAIB's IT manager, in October 2003. MAIB agreed to participate and supplied the time and services of Ms Pennock, Ms Julie Stubbington (Business Records and Investigation Support officer at MAIB), and Mr Simon Harwood (Deputy Chief Inspector of Marine Accidents at MAIB).

The questionnaire was completed by all three parties and returned to ERPANET via email at the end of November 2003. This was supplemented by a follow-up telephone conversation with Ms Pennock on January 12th 2004. In addition to this, Mr David Double (Review Manager of the Records Policy and Review section at the Department for Transport) provided further records policy information via e-mail.

## **Chapter 6: Analysis**

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

- Perception and Awareness of Digital Preservation
- Preservation Activity
- Compliance Monitoring
- Digital Preservation Costs
- Future Outlook

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

MAIB is a small and relatively young branch of government that recognises the value of using and maintaining digital information in their work. Although there is no digital preservation expertise on-site, the staff of MAIB, and in particular the IT section, are keen to improve their digital preservation knowledge and skills. This case study focuses on MAIB's efforts to protect its business-driven information; that is, their database, evidence, and investigation reports.

#### *The main problems*

The main problem MAIB perceives in preserving its digital information lies in the absence of sectoral and departmental advice on how to do so. Respondents reported difficulty in finding guidelines or methodologies to assist them in common preservation and maintenance activities such as database migration procedures and evaluations. Database migrations have been carried out nonetheless, but have required an element of a trial and error approach that is ultimately more costly and error-prone than first anticipated.

There is also a problem caused by the lack of standard data formats for the particular data-types MAIB encounters when collecting evidence. This is particularly so with the industry specific 'black box' and 'white box' voyage data recorder (VDR) data.<sup>9</sup> In 2002, 13 different manufacturers had developed, or were in the process of developing, VDR's. Many use unique and proprietary formats, with their own methods of data compression and retrieval.<sup>10</sup> MAIB may encounter and need to preserve any and all of these during its work, but the more formats it accumulates, the more effort must be expended for effective preservation.<sup>11</sup>

---

<sup>9</sup> Black boxes can contain about 12 hours of data and are triggered by an event. White boxes feed black boxes and can hold much more information.

<sup>10</sup> See MAIB 2002 Annual report:  
[http://www.dft.gov.uk/stellent/groups/dft\\_maritimesafety/documents/page/dft\\_masafety\\_022759.hcsp](http://www.dft.gov.uk/stellent/groups/dft_maritimesafety/documents/page/dft_masafety_022759.hcsp).

<sup>11</sup> As mentioned in chapter 4 and MAIB's 2002 Annual report (op cit), MAIB has been collaborating with the IMO to develop a simplified VDR that can be fitted to all existing cargo ships in the near future. If adopted, it has the potential to standardise and thus simplify the preservation of this particular record-type across the maritime industry.

### *Asset value and risk exposure*

MAIB has three principal core-business record-types in its collection:

1. Database records of reported marine accidents. The database is an internal record of reported marine accidents that have occurred since 1991. MAIB receive approximately 2,000 accident reports each year; of these, approximately 50 will be investigated.
2. Evidence collected in the course of an investigation into a reported marine accident. For every investigation that is carried out, evidence must be collected; this can take numerous forms, such as log books, charts, interviews, photographs, physical evidence, vessel computer records, audio and video files, and VDR data.<sup>12</sup> No two investigations are the same, and different data, in different formats will be collected as evidence in each one. All evidence collected during the course of an investigation must be preserved but it is neither published nor made available to the public by MAIB.
3. Marine Accident Investigation Reports, produced from analysis of the evidence collected during an investigation. All investigations result in a final report, published as a PDF document and a HTML page on the Internet, and also in hard copy.

MAIB records, reports, and evidence are preserved as direct products of their core business activity. The main reasons for preserving this information are legal and historical: government organisations must comply with legal regulations for preservation of material deemed to have historical value in the future.

MAIB are particularly aware of the historical value of their records. Their holdings serve as a unique historical record of accidents and reported incidents at sea from 1989 to the current day, and hold great value for scholars of the future. They are also keen to see their reports quickly disseminated to as wide an audience as possible in order that lessons can be learned and similar marine accidents avoided in the future. In light of this, they began making reports available over the Internet in 2002 and embarked on a digitisation project for reports held in hard copy produced before then.

None of the data has been assessed in the course of a branch-based risk analysis or business needs analysis.

### *Regulatory Environment*

The main legislation covering MAIB record keeping is the UK Public Records Act (1958).<sup>13</sup> This states that all public records must be made known to the UK National Archives (NA) and gives the NA the powers to issue guidance so they can identify and select which collections or groups of records an organisation ought to retain for possible permanent preservation in the National Archives at a later date. Records that are included in the guidance should be kept by the originating organisation for 25 years, after which they undergo further NA review to determine which are worthy of transfer to the NA for permanent preservation. The NA guidance is directed towards the responsible Department, in this case the DfT, who are responsible for delineating this information through to MAIB.

---

<sup>12</sup> For a more exhaustive list of potential evidence, see the MAIB guidelines on the proceedings of an investigation: <http://www.maiif.net/guidelines.htm>.

<sup>13</sup> See the UK Public Records Act, available from the UK National Archives website at: <http://www.pro.gov.uk/about/act/default.htm>.

MAIB are subject to the UK Data Protection Act (1998), which defines responsibilities for organisations holding records in relevant filing systems containing personal data, that is, information which relates to a living individual who can be identified from that data or other associated data held by a 'data controller'.<sup>14</sup> The DfT has issued guidance on implementation of the Data Protection Act to MAIB and the other units.

Responsibility for implementation of the Acts lies jointly with the Branch's Business Records Officer and the Department. MAIB are also subject to the Freedom of Information Act (Fol, 2000)<sup>15</sup>, which is handled in the same manner as the Data Protection Act.

### **Preservation Activity**

Guided primarily by Departmental policy towards the regulatory environment described above, MAIB preserve both digital and paper assets. They have recently begun to place more emphasis on public access and are in the process of making more of their reports publicly available via the Internet. Preservation became a matter of explicit discussion when MAIB were contacted by ERPANET in 2003.

#### *Policies and Strategies*

As a unit of the DfT, MAIB must adhere to the Department's policies. The DfT has a Departmental Records Policy, accessible to all staff through the DfT Intranet. This policy covers the entire record lifecycle of digital and paper records and states that in the absence of an 'accredited shared [electronic] drive' or Electronic Document Management System (EDMS), units must keep their records in hard copy. MAIB has no such accredited shared drive or EDMS and relies on its network drive for record storage, therefore falling under the hard copy requirements. The policy goes on to say that when organisations relying on hard copy encounter a situation where it is impractical to print an electronic record, for example, a database, that it should be recorded as a related 'dataset' file on the cover of the registered hard copy file.<sup>16</sup> The policy refers to this as the 'hybrid approach'.

Some of MAIB's data is not suitable for printing to paper and cannot be translated into a hard copy environment without a significant decrease in its functionality and evidential meaning. To cater for these types of records, the Branch has adopted the hybrid approach described above. MAIB also wish to maintain access to some of their other records in digital format despite their suitability for paper-based preservation, because it is an effective and efficient way to distribute them. MAIB management and staff have thus worked together to devise strategies for maintaining access to more of the records in digital form, in addition to the majority hard copy requirements. The strategies for preserving MAIB's complex range of data are still in the early stages of

---

<sup>14</sup> See the UK Data Protection Act, available from the Information Commissioner's Office at: <http://www.hms.gov.uk/acts/acts1998/19980029.htm#aofs>.

<sup>15</sup> See the UK Freedom of Information Act, available online from Her Majesty's Stationary Office (HMSO) website: <http://www.hms.gov.uk/acts/acts2000/20000036.htm>.

<sup>16</sup> i.e. the name and location of the dataset should be written on the front cover of the physical dossier/file folder. Alternatively, organisations permitted to store their records electronically can link this data into their electronic filing system and record it in the relevant folder on the EDMS/accredited shared drive. Unfortunately, the policy makes no reference to how to maintain or preserve the database/set in either of the circumstances.

development, but migration to commonly available and accessible formats forms a key feature of their approach.

### *Selection*

Selection is dictated by the inclusion of MAIB records in the guidance issued by the NA. Each unit in the DfT has a file plan addressing administrative and investigative records that specifies which records are included in the NA guidance and how long they must be retained for. MAIB's file plan states that all investigation files (i.e. database records, evidence, and final reports) must be retained for 25 years and the policy described above applies this to both paper and digital records. DfT units are often also issued with Disposal Agreements, which are required before any records can be permanently deleted or destroyed. MAIB do not have a Disposal Agreement, so are not permitted to delete any of the records in their filing systems.

MAIB's file plan is implemented in a Branch-wide registry system, catering for both administrative and investigative files. The Branch Business Records Manager is charged with responsibility for maintenance and implementation of the registry. The registry is intended to cover the entire organisational structure and processes, but there are no measures in place at the Branch level to ensure that this is so.

### *Preservation*

Preservation activities are carried out mainly in-house, complemented by external consultants and support from the DfT when necessary. Activities vary depending on the record type and its requirements, so there are different technical approaches for the database, the evidence, and the reports.

The MAIB database contains records of every accident reported to MAIB since 1991. It is an active static database in which data is never over-written, simply added. It has been migrated on two occasions with some support from private consultants and project management from the DfT. The database was originally a DOS database in Data Ease 4.53.<sup>17</sup> In 1996/7, the DfT attempted to standardise database formats across the different DfT units and selected Access 2.0 as the Departmental database application.<sup>18</sup> MAIB required the same functionality from Access as Data Ease; unfortunately, Access was unable to provide it within the new structure designed by the external consultancy team and the migration subsequently failed. The DfT later relaxed their restrictions and allowed the Branch to select an application that could be more easily manipulated to meet their needs. They chose Oracle 8i<sup>19</sup> and used both private consultants and project management support from the DfT to re-migrate the data from Data Ease into Oracle, this time with greater success.

Investigation reports are preserved primarily in hard copy, previously published by HMSO but published in-house since 2000<sup>20</sup>. In 2002, the Branch began creating PDF<sup>21</sup>

---

<sup>17</sup> DOS stands for Disk Operating System. DOS systems were accessed through a command line interface, as opposed to today's graphical user interfaces. Data Ease is a proprietary database system currently owned by Sapphire International, see: <http://www.dataease.co.uk/>.

<sup>18</sup> Microsoft Access website: <http://www.microsoft.com/office/access/>.

<sup>19</sup> Oracle Corporation website: <http://www.oracle.com/>.

<sup>20</sup> HMSO (Her Majesty's Stationary Office) website: <http://www.hmso.gov.uk/>.

<sup>21</sup> PDF stands for Portable Document Format. It is owned by the Adobe corporation: <http://www.adobe.com/>.

versions of the reports for immediate publication over the Internet from the original Microsoft Word and Quark XPress typesetting files.<sup>22</sup> These are saved on the network drive. MAIB is also creating PDF versions of the reports created prior to 2002, involving digitisation of the hard copy reports that MAIB has in storage.<sup>23</sup>

Evidence causes the real trouble for MAIB, requiring a more pragmatic and complicated approach than the static database and text documents. Marine accident evidence can come in a variety of non-standard formats, including data from proprietary, rare, or customised data systems. It often includes VDR data, key navigational information that can be used to build up an accurate picture of events leading up to an accident. MAIB collect all evidence and convert to widely accessible formats wherever possible. Office and audio files are migrated to the de facto proprietary standards of Microsoft's Office suite and \*.wav format. Evidence consisting of original data from commonly obsolete applications requires MAIB to seek conversion software from external sources. Where this is not possible, the data must remain in its original format, and MAIB must obtain and retain copies of the software (such as old operating systems and programs) if possible. Once registered in the file registry system, evidence (and other software, if appropriate) is written to CD-ROM and stored with the appropriate paper report.<sup>24</sup>

There are no internal guidelines in place with written details for carrying out these technical conversions and MAIB relies on the knowledge and experience of its IT staff to ensure that they remain up to date. The procedures have been developed internally without record-keeping guidance, and the digital records have almost no metadata assigned to them. The only metadata is the reference linking them to the paper records. No media refreshment or further file format migration has yet been scheduled, although the database and network drive is routinely backed-up as part of the regular IT schedule.

### Access

Access is a key issue behind MAIB's digital preservation activities, for both internal and external parties. As mentioned, improving access was the impetus behind MAIB's digitisation project on reports issued prior to 2002. Improving access was also behind the decision to convert internal evidence files to Microsoft's de facto standard formats, in the belief that formats used by the majority of today's users are more likely to be supported by a wider range of products in the future.<sup>25</sup>

Public access to MAIB's records is limited to the final reports, which are distributed in hard copy and in PDF and HTML over the Internet. These formats were chosen to comply with UK e-government accessibility aims, as both are accessible to blind or

---

<sup>22</sup> Microsoft Word website: <http://www.microsoft.com/office/word/default.asp>; Quark XPress website: <http://www.quark.com/products/xpress/>.

<sup>23</sup> Reports consist of a combination of images and text, which is why the files are put into Quark for typesetting before being converted to PDF. This has been the case since 2002. Before then, typesetting was the responsibility of the publisher and MAIB are not in possession of the typeset files. It is simpler to scan the hard copies they have preserved than to re type-set the old files.

<sup>24</sup> Data has been written to CD-ROM since approximately the beginning of 2003. Prior to this, data that could not be printed to paper was stored on a floppy disc.

<sup>25</sup> Whilst there may be a wider range of products claiming to support these formats in the future, the quality of support, and thus also the record as rendered on-screen, will vary from product to product due to the proprietary nature of the file format.

partially sighted users through the use of extra (or 'plug-in') software. Copies of the reports stored on the network drive are uploaded to the web-server for access over the Internet.

Internal access to the digital records on MAIB's network drive is controlled by login-based access and security privileges associated directly with the operating system. Access to the records stored on CD-ROM or floppy disc can be gained with the help of the IT department, which maintains several PC's running different versions of Windows and with different media drives. MAIB provides some advice on accessing black box data in their Investigation Guidelines: in circumstances where information from a VDR is available but the investigating State does not have the appropriate reading facilities, it should seek and use the facilities of another State, bearing in mind the timeliness of the availability of the facility.<sup>26</sup> The discs are stored with the paper records in a humidity controlled environment and are the responsibility of the Business Records Officer.

### **Compliance Monitoring**

There is no compliance monitoring in place regarding the preservation of digital information at MAIB. This is largely due to the fact that they are required to keep their records in hard copy. MAIB's difficulty in complying with that policy indicates that compliance monitoring of their hybrid system would be beneficial.

### **Digital Preservation Costs**

MAIB have not yet made a case for digital preservation funding and, as yet, none has been allocated. Digital preservation is not seen as a separate activity, although the Branch believes that funding would be available were they to make such a case. Respondents were unable to provide further answers on aspects of MAIB's or on the Department's finances.

### **Future Outlook**

MAIB are very keen to improve their understanding of digital preservation and how to achieve it. They estimate that their current strategy, using a hybrid approach of paper and digital records systems and media, will only last them for the next 2 – 3 years. As a result of this and their experiences with ERPANET, they hope to develop in the next few years a sustainable migration policy and strategy that will detail file formats and media suitable for use with all of their different types of records over the long term.

The lack of sectoral guidance on preserving digital marine accident records is also an issue the Branch would like to see addressed. More information about the policies followed by other national marine accident investigation units is needed, as is evaluation of these policies over the short and medium term, to better develop new policies and to minimise the effort and cost each one must expend to achieve the same thing. They are aware of the cost investment this will require but expect to see more money allocated for digital preservation in the future, as the urgency grows and the subject attracts more attention.

The DfT are currently implementing a project for full electronic document and records management across the whole Department. Thus, at some point in the near future,

---

<sup>26</sup> There are several other 'States' which operate a similar organisation for marine accidents, including Australia, Hong Kong, and Japan. Other countries incorporate this function into a Transport Accident or Transport Safety organisation that considers all types of transport, not just maritime.

MAIB will also be undergoing a complete migration to electronic documents and records management. This should significantly change and strengthen their electronic record-keeping infrastructure.

## **Chapter 7: Conclusions**

MAIB have taken the first steps to ensuring that the data they hold in digital format is successfully preserved and remains accessible over time. They appreciate the value of their data and the staff are committed to finding an effective solution to digital preservation. Some issues still require further attention, but MAIB have begun addressing them at an early stage and thus their preservation problems are not insurmountable.

The wide array of evidential data to be kept, in varying formats and of different record types, is problematic. Some of this data can be converted to standards, but the standards must be well chosen and the conversion must be checked to ensure that a faithful representation of the record is captured. Using proprietary standards without any evaluation procedures increases the risks of obsolescence and damaged integrity. Other data cannot be converted to standards, and in such instances the organisation must preserve not only the file but also the file application, or else they must rely on external parties to provide rendering facilities. This takes some organisation and can be very costly if the number of unique formats continues to rise. In addition to this, the media on which these files are stored (i.e. floppy disc or CD ROM) must also be monitored for signs of degradation and the hardware (i.e. drives) to read them must also be maintained.

The challenge is compounded by the lack of digital archiving and digital preservation skills within the Branch, and the lack of guidance issued on such matters by the Department. This is evident from MAIB's digital preservation activity, which is characterised by a lack of technical preservation aspects and digital record keeping considerations. Conspicuously absent from MAIB's approach are technical and record keeping metadata, media refreshment, secure offsite storage, use of open standards (see above), authenticity, and integrity checks. These aspects of MAIB's preservation activity are inextricably linked to the fact that the Branch is operating a hybrid record keeping system within the context of a record keeping policy designed primarily for active records that does not sufficiently account for long term preservation of digital records.

In addition, MAIB are creating and preserving data for two different reasons, which means they have different requirements in place and, in effect, different aims. The first reason is to meet legal and historical requirements, which they achieve by preserving as much of the data as possible in hard copy and complementing it with a CD-ROM of digital evidence that cannot be printed out onto paper. The second reason is access provision. MAIB are making many of their reports available online to increase access options, and as record-keeping is sufficiently addressed for the paper records, it is not a primary concern for the digital ones. Many of the technical and record keeping aspects to digital preservation have therefore become inconsequential, but they will again be relevant if the Branch attempts to rely on these versions of the records when full electronic documents and records management has been implemented. The Department is currently engaged on a project to do just that.

The strength of the UK's drive towards e-government, providing a purposefully designed electronic infrastructure which accounts for digital record keeping issues, should help MAIB in facing some of these challenges.<sup>27</sup> E-government is about more

---

<sup>27</sup> The UK approach scores reasonably well in the EC Information Society October 2003 benchmark report of e-government in Europe, coming 8th place (out of 18) in levels of online sophistication, and 6th place in the number of services fully available online. See

than just improving access by moving services online. It also concerns (often major) organisational restructuring of the back office, the area where policies are made and records are kept. In the UK, it incorporates significant centralised research by the National Archives into practical long term preservation of digital records. This support infrastructure is yet to be fully developed, but MAIB should find it much easier to co-ordinate its approach once this has been done. The future shift to a fully digital record keeping system will make a big difference, introducing amongst other things a more strictly controlled digital file storage environment.

Despite the many challenges, ERPANET is confident that MAIB will be able to preserve its digital records in the future. Although they may experience some problems accessing the data stored on discs, it is not yet too late to design a suitable strategy for the rest of their records that will build on their existing format and application choices (i.e. PDF and Oracle) and, furthermore, incorporate vital metadata and authenticity considerations. In this respect, MAIB should pay close attention to the activities of the digital preservation department at the National Archives. The NA's PRONOM file format registry system provides technical information on file formats used to store digital records and software that can be used to render them. They also provide guidance on file format selection, storage media, graphic file formats and image compression through their publicly accessible website.<sup>28</sup> MAIB should make use of these resources and co-operate with the NA when developing their strategy for the long term. Once the unique challenges are more keenly understood by all of the government parties involved, then they will be easier to solve.

---

[http://europa.eu.int/information\\_society/eeurope/2005/doc/highlights/whats\\_new/capgemini4.pdf](http://europa.eu.int/information_society/eeurope/2005/doc/highlights/whats_new/capgemini4.pdf) for the full report. For further information about the e-government strategy in the UK, see <http://www.e-envoy.gov.uk/Responsibilities/EGovernmentStrategy/fs/en>.

<sup>28</sup> This information was available on their website on February 3rd 2004. Several more publications are planned for the future, on such topics as digital signatures, encryption and compressed archive file formats. See <http://www.pro.gov.uk/about/preservation/digital/archive/default.htm>.

## **Appendix 1: References**

UK Government Department for Transport (DfT) website: <http://www.dft.gov.uk/>.

International Maritime Organisation (IMO) website: <http://www.imo.org/index.htm>.

Marine Accident Investigation Branch: <http://www.maib.dft.gov.uk/>.

Marine Accident Investigators International Forum (MAIIF) website:  
<http://www.maiif.net/>.

Digital Preservation at the UK National Archives:

<http://www.pro.gov.uk/about/preservation/digital/archive/default.htm> .

E-Government in the EU:

[http://europa.eu.int/information\\_society/eeurope/2005/all\\_about/egovernment/index\\_en.htm](http://europa.eu.int/information_society/eeurope/2005/all_about/egovernment/index_en.htm).

E-Government in the UK – Office of the e-Envoy: <http://www.e-envoy.gov.uk/Home/Homepage/fs/en>.

Information Commissioner's Office (UK): <http://www.informationcommissioner.gov.uk/>.

Her Majesty's Stationary Office (HMSO): <http://www.hmso.gov.uk/>.

All hyperlinks functional on 24th February 2004.

## CONTACT DETAILS

### **ERPANET Coordinator**

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

Tel: +44 141 330 4568  
Fax: +44 141 330 3788

Coordinator@erpanet.org

## ERPANET STAFF

### **directors**

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

### **coordinator**

Peter McKinney

### **editors**

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

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