

Digital heritage for the future*

YOLA DE LUSENET

A B S T R A C T

This paper provides an overview of various issues related to digital preservation: the nature of digital information, the roles and responsibilities of institutions, selection of materials and choices in preservation, with many references to current initiatives. It argues that traditional concepts like publication and record have to be redefined in the digital world, in order to develop effective preservation policies. Preservation has become a matter of keeping access rather than keeping objects, for which the first requirement is defining the digital object and its essential characteristics. Documentation has moved to the centre of preservation activities, which now have to take place quite early in the lifecycle of materials.

PALAVRAS-CHAVE

PRESERVAÇÃO DIGITAL

POLÍTICAS DE PRESERVAÇÃO

PUBLICAÇÃO NA INTERNET

ARQUIVOS ELECTRÓNICOS

ACESSO A DOCUMENTOS DIGITAIS

R E S U M O

Visão geral sobre vários aspectos relacionados com a preservação digital: a natureza da informação digital, os papéis e responsabilidades das instituições, a selecção de materiais e as opções na preservação, com várias referências a iniciativas actuais. Defende que alguns conceitos tradicionais, como publicação ou documento, têm de ser redefinidos no mundo digital, de forma a desenvolver políticas eficazes de preservação. A preservação tornou-se na questão de manter o acesso, mais do que manter os objectos, para o que o primeiro requisito é a definição do objecto digital e das suas características essenciais. A documentação deslocou-se para o centro das actividades de preservação, que têm agora de ser desenvolvidas desde uma fase precoce do ciclo de vida dos materiais.

>

* This is a revised and shortened version of a discussion paper commissioned by UNESCO in the framework of activities for preservation of the digital heritage; it was subsequently adapted and incorporated in documents presented to the Executive Board in May 2002 (document 20, see <http://unesdoc.unesco.org/ulis/>). The original paper is available at the website of the European Commission on Preservation and Access (<http://www.knaw.nl/ecpa/PUBL/unesco.html>)

INTRODUCTION

Rapidly, the amount of information in the world is becoming so vast that the human mind cannot take it in: it has been estimated that the total **annual** production of information is now approximately the equivalent of 1 or 2 billion Gigabytes.¹ It has also been estimated that 90% of this exists in digital form. And again a large percentage of this exists in digital form **only**, ranging from medical records to movie DVD, from satellite surveillance data to websites presenting multimedia art, from data on consumer behavior collected by supermarkets to a scientific database documenting the human genome.

For those entrusted with collecting and preserving cultural heritage, the question has become extremely pressing as to what of this enormous amount of materials should be kept for future generations, and how to go about selecting and preserving it. With the advent of digital media a new and complex environment has come into being. Not only the media are new, the contents and the means of distribution have also changed dramatically, and new players – among users as well as creators of information – have entered the stage. With so many new types of content presented in new ways, the first question is: what is worth keeping and who should take care of this?

NEW CONCEPTS AND RESPONSIBILITIES

Traditionally, preservation of cultural heritage has been supported by legal frameworks and procedures, which are largely based on formal criteria. Libraries take care of print publications; there is legal deposit of the national production. We know what printed publications are and where they are produced. There is extensive archival legislation defining when and how records must be transferred to archives for selection and preservation. Sound archives collect sound, film archives film, etc.

In the digital world, new types of materials have come into being that is hard to classify by conventional criteria. Multimedia materials combine different types of content with different functionalities. A database is not a fixed object that can be stored in a definitive form; nor can one separate the data from the relationships between them. Websites may combine files with various types of content – data, texts, images, sound – or they may be distributed sites including materials stored on different servers at different locations in the world.

Such mixed or dynamic materials do not fit into traditional categories, and it is not immediately obvious who should be responsible for collecting and preserving them.

Especially on the web, the usual filtering mechanisms of publishers and other agencies that review and select materials worth publishing often do not apply. Although we speak of publishing *on the internet*, it is not at all clear what constitutes an Internet publication. The more than one billion pages of the internet are only partly devoted to materials that resemble what heritage institutions have traditionally been collecting: electronic journals and articles, newspapers, photographs, catalogues and finding aids. On the other end of the scale there are innumerable websites created by individuals and informal groups, which deal with basically every topic under the sun, from digital art to recipes. If the Internet as a whole is seen as a reflection of our society, as a huge open space where a variety of cultural activities are pursued, then preservation will somehow have to deal with these new manifestations of cultural content on the web.

If we think of this in terms of *preservation of digital heritage*, we have to be careful that we know what we mean. Conventional approaches to preservation of heritage do not fit comfortably in the digital environment. *Heritage* deals with things that were left to us from the past and *preservation* deals with these things often a long time after they were first created. But the digital world is moving so quickly that we cannot wait for 10 or 50 years to see what will prove to be *heritage*. Because generations of soft and hardware succeed one another so rapidly, anything left to its own devices for an extended period of time will have become inaccessible. Think of WordStar files of 5 1/4 inch disks that are 15 years old and now very difficult to read. In the digital world, the time scale for preservation has shrunk. Steps to ensure that digital materials remain accessible have to be taken very early in their lifecycle. The Public Record Office writes in its guidelines for electronic record management: «Departmental preservation strategies must provide for long-term preservation; that is, for periods of five years or longer».² The message here is that anything that needs to be accessible for more than five years in the digital world requires steps to ensure its preservation, whereas in the world of paper selection for preservation may take place decades later. And now that *long-term* is not really *long term* anymore, it has become necessary to define what we mean by it: 10 years, 50 years, 200 years?

Apart from the need to keep up with technological change, there is the additional problem that websites are changed constantly, and materials vanish without leaving a trace. Estimates for the average life expectancy of a web page vary from 44 days to two years.³ When organizations go out of business or lose interest, whole websites disappear from sight. And in case you may think this mainly happens to informal and small websites: when George Bush took over the presidency, the White House site was wiped clean. The collection of speeches and official communications of the Clinton administration disappeared overnight, breaking a massive amount of links to these materials on other sites.⁴ The material could only be saved because the National Archives and Records Administration (NARA) had archived several versions of the site over the years. This example only goes to show that in the digital world important materials may be lost very easily if no one takes care of saving them early on.

Just as the concept of a *publication* requires a new definition, so in the archival world, electronic records cannot be understood the same way as conventional paper records. There is no longer a fixed object that can be preserved as is, and it becomes necessary to decide which elements actually make up an authentic electronic record. With records being used for years or even decades, they will inevitably have to be moved from outdated environments to new ones, with the risk of changes or loss of content, functionality or original appearance. It is highly unlikely that the original record can be kept as it was once created. Therefore, it has to be determined which are the significant properties of records that must be preserved.⁵ It also has to be established which of the new types of materials appearing on Internet sites should be regarded as records covered by archival procedures and legislation.⁶

CURRENT APPROACHES

In the library world, legal deposit offers one possible approach. Deposit of offline digital products, such as CD-ROM, is in several countries already a legal requirement.⁷ To ensure continued access to online electronic journals, including live links, data and multimedia presentations, in a number of countries libraries are now trying to come to arrangements with publishers about the creation of electronic repositories, as yet often on a voluntary basis.

Several libraries have developed strategies for actively selecting and preserving

websites on the basis of a concept of *publication*,⁸ of which the Pandora project of the National Library of Australia is perhaps the best-known example. Sites selected for preservation should be clearly related to Australia, and priority is given to sites offering content of lasting research value.⁹ *Publication* is defined in broad terms: anything publicly available on the internet is regarded as a publication, and the Library also says: «Distinctions between traditional categories such as books, serials, manuscripts, working drafts and organizational records are blurred in the electronic environment. It is not the intention of the Library to preserve organizational records and similar materials, which are the domain of archives and record management».¹⁰

Some national archives have issued guidelines for the preservation of websites (intranet or internet) of government agencies becomes some of them can be regarded as records. The Public Record Office warns that materials on websites are not always recognized as records: they are often «very different in nature from the traditional image of a *record*. So much so that it can tend to give the impression that no records are present. This can be highly misleading».¹¹

These are selective approaches for preserving web content, in which heritage institutions actively search for relevant sites and select what they believe should be kept. Selection is a difficult and time-consuming process that requires the input of people to make decisions. That is why others believe a comprehensive approach of web harvesting by robots is more feasible. In this way enormous numbers of web pages without any selection for content. It requires an enormous amount of storage space, but that may be cheaper and easier to come by than the labor involved in selecting materials. An example is the Internet Archive, started in 1996 as a private, nonprofit enterprise. It collects freely available web pages worldwide and now comprises over 10 billion web pages or 100 terabytes of data (5 times the size of all the materials held by the Library of Congress). The Internet Archive is freely accessible through the Wayback Machine.¹²

In Sweden and Finland, for a number of years websites with Swedish or Finnish domain names or providing content about the country have been harvested.¹³ This activity of freely available materials is regarded as complementary to the legal deposit of paid materials by established publishers.

At the moment, the main aim of these initiatives is to save web materials that would otherwise in any case have been lost forever. They show us the evolution

of the Internet, as well as snapshots of our society around the turn of the century. However, that the material has been saved does not mean we can use the sites as they were meant to be used. Links to external sites will in many cases be broken and interactive navigation cannot always be retained. More and more web pages are dynamic, generated *on the fly* by databases hidden behind the static front end of the site. It is estimated that the databases behind websites, together called the *deep web*, contain many more times the amount of information accessible on the surface. The information in those databases cannot be captured by copying the website, as it is not available in ready-made pages at the surface.¹⁴ Moreover, capturing web content is only the first step in a preservation process. After only six years of archiving, there is no saying yet how it can be ensured that these materials will still be available after 25 or 50 years.

WHAT IS PRESERVATION OF DIGITAL MATERIALS ANYWAY?

Once it has been decided what we want to preserve, it has to be figured out what preservation of electronic materials actually means. In the world of print, preservation can be achieved by preserving the paper object or, if that is not feasible, creating a durable surrogate for instance on microform. The equivalent in the digital world would be, for example, to preserve a CD-ROM, or transfer its contents to another type of carrier when the CD itself can no longer be used. However, that does not achieve much, for having all we have preserved then are the bits on a carrier. That does not mean that the information can actually be read and interpreted in the long run.

To keep information accessible, either the programmes that can read the files have to be preserved as well and somehow kept running on new platforms, or the files have to be converted to another format that can be interpreted by new programmes. As the digital world moves on all the time, this is a continuous process. In many cases, sooner or later, information, functionality and/or appearance will be lost, especially with complex, multimedia materials that combine a variety of file formats and applications.

This poses risks for integrity of digital materials: how can it be ensured that the digital object, moving from one environment to the next, remains complete and undamaged? A different but related issue is authenticity, which relates to the trustworthiness of materials, in particular, of electronic records. As records

are used for accountability and as evidence of transactions, it is crucial that the record indeed is what it purports to be. Integrity and authenticity do not only depend on protecting files against intentional changes by unauthorized persons, but also on controlling inadvertent changes resulting from mis-interpretation or mis-representation by computer systems.¹⁵

Because carriers are temporary and environments change, preservation of digital materials cannot be understood in terms of fixed objects that should be kept in their present form. It is first of all a matter of defining the content and properties that need to be represented in future systems. If we think of data files for instance, do we want to keep data as they are, *frozen* as it were, or should future users have the possibility to search, select and sort the data – in which case not only the data but also the software that makes it possible to work with them has to be retained. It may even be necessary, for optimal functionality and access is the primary goal, to upgrade to future systems. Otherwise, future users will be stuck with a level of access and functionality that to them will be very old-fashioned, just as if we were now using punch cards with queries in FORTRAN or COBOL.¹⁶

In other cases the opposite may apply: some materials may have to be represented in a historical context, to be really understood and appreciated, so that future users can experience them exactly as we do now. Think for instance of digital art: for some artists the way the work is displayed (e. g. on a specific type of screen or using a specific browser) is an integral part of the work. Museums now often collect information on artists' intentions to make sure they understand what the work really is and how it is meant to be shown, in order to be able to preserve it as the artist would prefer.¹⁷

So what must be preserved may be different in different cases, but adequate representation at a later stage depends on the identification of the type of content and file formats as well as the software that makes access possible. Only if one knows what one is dealing with can suitable preservation measures be taken. Documentation starts at the lowest level, by describing the characteristics of the bit stream as well as the hardware/software environment capable of rendering the object in its present form.¹⁸

Additional documentation is needed to understand and evaluate what is presented: information as such, without context and background information, will be hard

to *place*. Especially for electronic records, the context of materials is of crucial importance. It makes all the difference for understanding a map with red dots on it whether it was used for geological exploration or military actions, and this cannot always easily be seen from the map itself if it is presented in isolation. It therefore needs to be specified how and when the material came into being, who has held it, and how it relates to other information. Documentation also needs to include data on changes made over time, transfer from one format to another and on authenticity (e. g. by using checksum or digital signatures).¹⁹

Metadata, as all this description and documentation is called in the digital world, to some extent deals with familiar descriptive elements we know from all kinds of catalogues and finding aids, like *author* or *title* or *year* or *collection*.

It also handles relationships between separate files by specifying for instance the organization of different files into a larger whole, like the chapters of a book. As from a digital image of for instance an illustration it cannot be seen where it is meant to go in a text file, these structural relationships have to be made explicit. Metadata also includes a lot of technical documentation that has become necessary because we need to know how the computer can access the information **before** we can see it or read it and try to understand what it means. In digital preservation there is a shift from preserving the carrier, the object itself and the information it contains, to preserving ways of accessing the information. We are moving from preservation **for** access to preservation **of** access. You cannot preserve what you see on a screen, that is only the temporary manifestation of a digital file, and we need to make sure we know how to recreate this temporary manifestation in new environments.

This is a technological issue, but the choice of technological solutions depends on the requirements for future representation: «any technological choice we make has inescapable implications for what will (and will not) be preserved. In the digital case, we must choose what to lose».²⁰

If I have not emphasized technological aspects so far, it is because I believe that we should not start from the technical end. A technological solution is no solution if we do not know what we want to achieve, and once we understand where we need to go, the technology will somehow be developed. Some of the larger national institutions are working on technical solutions, often in cooperation with industry. In the end, the technology is not something

heritage institutions will have to develop themselves. If they can only define the requirements for digital preservation, then others, companies that have all the technical know-how can do the job.

Of course an understanding of the principles behind the technology is definitely essential. You have to be aware that media are transitory carriers that serve their function only for a limited period of time and that transfer to new media is an absolute necessity. You have to be aware that soft – and hardware become obsolete in years rather than decades and that although successive versions of programmes may be compatible, software producers do not usually support compatibility over a long period. You also have to be aware of the value of using standards, for documentation as well as for platforms and file formats, and that proprietary software is problematic not only because it is protected and the source code is not available, but also because it is often inadequately documented, so conversion of files remains very much a black box. Having said all that, I still believe the efforts of the heritage sector would be misdirected if we focused on digital preservation as a technical issue. Heritage institutions have specific expertise in selecting materials worth preserving and understanding how they are used, and that expertise should be used to define where we want to go.

COOPERATION AND NEW FRAMEWORKS

The other area, which requires a lot of joint effort, is the division of tasks and the creation of frameworks that support digital preservation. Because preservation requirements have to be taken into account very early on, even at the point when material is created, «the first line of defense against loss of valuable information rests with the creators, providers and owners of digital information».²¹ The traditional roles of creators and keepers of information, as distinct responsibilities, have to be abandoned in the digital world.

Creators should be made aware that choices made at the time of creation affect the possibilities for later preservation. That is why archiving institutions like the Public Record Office have issued guidelines for record-creating agencies, to make them aware of their responsibility in creating records that can indeed be maintained over time.

Cooperation with creators and owners of information is also crucial because of copyright. Copyright legislation places such strict limitations on copying

that even transferring files to the library's system may constitute an infringement. In some countries copyright law has been extended to allow copying by heritage institutions for preservation purposes. There are also examples of voluntary arrangements between libraries and publishers that cover deposit and copying for purposes while restricting access.²² However, rights management is developing into an extremely complex area and not all aspects can be covered by agreements between publishers and libraries. When a digital product relies on proprietary software owned by third parties, the creator of the content does not usually hold the rights to the software. Software vendors have so far hardly been involved in preservation efforts and software is not usually covered by deposit legislation.²³ A dazzling array of rights may be associated with websites combining mixed materials from various sources. A general agreement on the principle of the right to copy for preservation will therefore have to be sought to make copyright aspects of preservation more manageable.

Ideally, responsibility for preservation is shared by creators and keepers, each maintaining materials during a certain stage of their life cycle. However, as there is a risk that creators are not sufficiently aware of the need for continuing maintenance, deposit regulations should help to ensure that materials are indeed transferred to an archiving institution. For this, a distributed system of trusted, digital repositories needs to be created, capable of keeping materials alive for the long term and making them available to users as agreed with the depositor. National libraries and archives are at present taking on this role. There are, also a number of specialized research institutes and data archives that clearly see a role for themselves. Apart from those already actively involved, there is a range of other institutions that may have a task in preserving certain types of materials (digital photographs, sound, art, broadcasting materials) or preserving materials for a specific community (institutions with a local or regional task, research institutes in a specific discipline). A distributed system of digital archives makes it possible for institutions to specialize, by focusing on specific types of materials or on serving a specific community.²⁴

Some of this can be achieved by joint efforts of the heritage community itself; some of it requires cooperation, national and international. Over the past year, both UNESCO and the European Union, under the Spanish presidency, have passed resolutions that urge governments to take steps to preserve the digital memory. The General Conference of UNESCO passed a resolution

in October 2001, and this was the start of a programme on preservation of digital heritage which includes the formulation of a charter for digital preservation – to be presented at the next General Conference in 2003 – the preparation of technical guidelines by the National Library of Australia and consultation meetings in various regions to provide feedback.²⁵

These initiatives will not solve the immediate challenges institutions are facing now but they are extremely important to increase awareness at a political level of what is at stake. Without legal foundation, central regulations and adequate funding, efforts of the heritage sector will prove to be futile. And simultaneously, none of these political initiatives will have any real effect without serious lobbying with governments and EU agencies for adequate implementation. There are formidable obstacles on the way, but it is promising that the issue is beginning to be recognized, and now that the first steps are being taken, it is up to the heritage sector to give direction to these developments.

NOTES

¹ This estimate has been corrected for duplication: it refers to **unique** information. Only 0.003% of the total annual production of information is originally produced in print format. In evaluating these figures one should of course remember that quantitative measurements in terms of storage space do not say anything about the relative *weight* of contents: one photograph easily take as much space to store as several volumes of text in ASCII format. See: *How much information?* <http://www.sims.berkeley.edu/how-much-info>

² *Management, Appraisal and Preservation of Electronic Records*, vol. 2: Procedures, chapter 5: Preservation of electronic records, 5.14. <http://www.pro.gov.uk/recordsmanagement/eros/guidelines/procedures5.htm>

³ KENNEY, Anne R. et al. Preservation risk management for web resources. *D-Lib Magazine*, vol. 8:1, January 2002. <http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/january02/kennedy/01kennedy.html>

⁴ WIGGINS, Richard. Digital preservation: paradox and promise. *Library Journal/netConnect* Spring 2001, <http://www.libraryjournal.com/digital-preservation.asp>

⁵ «With the rapid development of information and communications technology in government a wider range of record types will emerge: website (hypertext) documents, multimedia documents, digital audio and video, and dynamically interlinked documents. Many of the developments in desktop information technology will tend to blur the boundaries between types of records, and increase the problems of capturing and retaining all elements of a record.»

<http://www.pro.gov.uk/recordsmanagement/eros/principles0.htm>. For an extensive discussion of principles, see *InterPARES Preservation Task Force Final Report*, Draft October 2001, http://www.interpares.org/documents/ptf_draft_final_report.pdf and Jeff ROTHENBERG and Tora BIKSON, *Carrying Authentic, Understandable and Usable Digital Records Through Time*. Report to the Dutch National Archives and Ministry of the Interior, 1999.

⁶ See e.g. *Archiving Web Resources: A Policy for Keeping Records of Web-based Activity in the Commonwealth Government*, National Archives of Australia, revised version January 2001, http://www.naa.gov.au/recordkeeping/er/web_records/policy_contents.html, and *Management of Electronic Records on Websites and Intranets*.

An ERM Toolkit, Public Record Office, December 2001, http://www.pro.gov.uk/recordsmanagement/eros/website_toolkit.pdf

⁷ For instance in France, Norway, Canada, Germany and Austria; for discussion and overview see <http://www.nla.gov.au/padi/topics/67.html>

⁸ E.g. the Bibliothèque Nationale de Québec has formulated the principle that «networked publications are as important as traditional publications» and initiated a programme that is consistent with their legal deposit programme which assumes that networked publications, just like traditional publications, will be deposited at the beginning of their 'active life'. Selection includes «independent, coherent publications, monographs, serials» and excludes for instance institutional websites «taken as a whole». <http://www.bnf.fr/pages/infopro/ecdl/quebec/sl027.htm>

⁹ <http://pandora.nla.gov.au/selectionguidelines.html>

¹⁰ *Selection Guidelines*, 3.3, <http://pandora.nla.gov.au/selectionguidelines.html>

¹¹ *Management of Electronic Records on Websites and Intranets: an ERM Toolkit*, Public Record Office, December 2001, p.7. http://www.pro.gov.uk/recordsmanagement/eros/website_toolkit.pdf

¹² <http://webdev.archive.org/index.php>

¹³ In Sweden, different categories are collected: (1) addresses ending in .se; (2) web servers located in Sweden but whose addresses end in .com, .org, and .net; (3) pages by Swedish producers on a foreign server (very popular: the .nu domain of Niue, a small nation in the Pacific; nu is Swedish for 'now'); and (4) *suecana* – pages abroad with content about Sweden, e.g. travel or translations of Swedish literature.

<http://www.kb.se/kw3/ENG/Default.htm>.

For Finland, see Kirsti LOUNAMAA and Inkeri SALONHARJU, EVA – The acquisition and archiving of electronic network publications in Finland, *Tietolinja News* 1, 1999. <http://www.lib.helsinki.fi/tietolinja/0199/evaart.html>.

¹⁴ One could theoretically copy both the database and all the software necessary to generate the pages and in this way preserve the functionality; however, as the programmes

run on the server, one would need to gain direct access (as reserved for the webmaster) in order to copy them.

¹⁵ «The overall process of preservation must be continuous. If there is ever a point where we cannot reasonably assert that the record continues to carry its original message intact, we can never thereafter assert that it is authentic. It is important to recognize that while the process must be continuous over time, the activities that constitute the process are discrete steps. Each instance where the way the information is represented changes – whether moving between storage and use or between storage media or subsystems – is a potential point of failure, a weak link where the entire chain could be broken.» *InterPARES Preservation Task Force Final Report*, Draft, October 2001, p. 88, http://www.interpares.org/documents/ptf_draft_final_report.pdf.

¹⁶ THIBODEAU, Kenneth. Building the archives of the future. *D-Lib Magazine*, Vol. 7, N. 2, February 2001. <http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>
Cp also: «Preservers should assume that future users would want to use the best available technology for access to the records.

The design of preservation systems should take into consideration the need to be able to interface with evolving technologies for information discovery, retrieval, communication and presentation.» *InterPARES Preservation Task Force Final Report*, Draft, October 2001, p.19, http://www.interpares.org/documents/ptf_draft_final_report.pdf

¹⁷ See Howard BESSER, *Longevity of electronic art*, February 2001, <http://www.gseis.ucla.edu/~howard/Papers/elect-art-longevity.html> and the Guggenheim Variable Media Initiative (<http://www.guggenheim.org/variablemedia>), a proactive program that asks artists to provide guidelines for presenting their works in new environments.

¹⁸ The model that is now widely adopted to gain a better understanding of which elements and processes are needed for preserving any kind of digital information, the Open Archival Information System (OAIS), distinguishes a data object (the bit stream) and representation information (which enables the interpretation of the bit stream into meaningful information). See OCLC/RLG Working Group on Preservation Metadata, *A recommendation for content*

information, report, October 2001, p. 3.
<http://www.oclc.org/research/pmwg/contentinformation.pdf>

¹⁹ Several organizations have now published recommendations for preservation metadata for digital resources; see for instance the CEDARS project
<http://www.leeds.ac.uk/cedars/OutlineSpec.htm>
and the final report of the RLG Working Group on Preservation Issues of Metadata
<http://www.rlg.org/preserv/presmeta.html>

²⁰ ROTHENBERG, Jeff and BIKSON, Tora. *Carrying Authentic, Understandable and Usable Digital Records Through Time*. Report to the Dutch National Archives and Ministry of the Interior, 1999, p. 6.

²¹ WATERS, Donald and GARRETT, John. *Preserving Digital Information. Report of the Task Force on Archiving of Digital Information*, Commission on Preservation and Access/Research Libraries Group, 1996, p. 37.
<http://www.rlg.org/ArchTF>

²² The Koninklijke Bibliotheek, the national library of the Netherlands, and the Dutch Publishers Association have made such an experimental agreement specifying storage, copying, access etc. See <http://www.kb.nl/kb/dnp/overeenkomst-nuv-kb-en.pdf>

²³ *Trusted Digital Repositories: Attributes and Responsibilities*. An RLG-OCLC Report, Final version, May 2002, p. 19-20, <http://www.rlg.org/longterm/attribswg.html>.

²⁴ The concept of a distributed system of digital archives was first discussed in Donald WATERS and John GARRETT, *Preserving Digital Information. Report of the Task Force on Archiving of Digital Information*, Commission on Preservation and Access/Research Libraries Group, 1996, p. 37. <http://www.rlg.org/ArchTF>. The report RLG published last year builds on to the suggestions of WATERS and GARRETT; see *Trusted Digital Repositories: Attributes and Responsibilities*. An RLG-OCLC Report, Final version, May 2002, <http://www.rlg.org/longterm/attribswg.html>.

²⁵ See the pages on "e-heritage" on the UNESCO website (<http://www.unesco.org>).