

# CONSERVATION OF THE ELECTRONIC NATIONAL HERITAGE FOR THE FUTURE - LEGAL AND POLITICAL CONSIDERATIONS

Erich Schweighofer<sup>1</sup>, Andreas Rauber<sup>2</sup>

***Abstract.** Preservation of electronic media requires a deep intervention in the rights of authors and publishers: Besides the deposit of a copy of the work rights for performing preservation actions on the work are required. Preservation institutions must be allowed to ingest works into their archiving systems. Sufficient information – metadata - has to be provided or extracted from the digital objects. Furthermore, preservation actions may incur significant changes to the objects themselves. The article deals with present developments of deposit of electronic materials, challenges stemming from preservation actions to guarantee long-term access to the deposited materials and resulting open legal questions.*

## 1 INTRODUCTION

The legal deposit of the documentary heritage, or as the UNESCO [25, 26] names it, the *Memory of the World*, requires a decisive change in the methodology of preservation. Digital materials may become unreadable as changes in coding, formats, software, operating systems and hardware occur. Gone are the days when the preservation of the artefacts was sufficient. Digital objects have a relatively short life expectancy as the digital publishing media on which they were originally made available will deteriorate. Even when the carrier material, such as magnetic tapes, may remain intact over long periods of time, it is usually the reading devices that become unavailable, such as legacy tape or floppy disk drives. Existing solutions like refreshment by copying data from time to time to new media or the migration to new hardware and software configurations are costly solutions and, more importantly, have severe limits and effects on the material originally deposited. Furthermore, even when the actual bit-stream on a data carrier is kept available and readable, the objects themselves may turn out to be lost in a sense that they are no longer accessible and usable. Contrary to “conventional” objects, digital ones need a certain environment to be opened and interpreted, e.g. a specific software program for viewing and interacting with them. This software in turn usually requires a certain operating system in a specific version, which again requires a specific set of hardware elements, such as

<sup>1</sup> University of Vienna, Center for Computers and Law, Department of European, International and Comparative Law, A-1090 Vienna, Universitaetsstrasse 2, Austria, Erich.Schweighofer@univie.ac.at. At present at leave and working for the European Commission in Brussels. The expressed views are those of the author.

<sup>2</sup> Vienna University of Technology, Department of Software Technology, A-1040 Vienna, Favoritenstrasse 9-11, Austria, rauber@ifs.tuwien.ac.at.

processors, input/output devices, etc. If any of these become unavailable in a certain point in time, so does the object itself.

The consequence of this fact is that the intertwined linking between artefact and content must be given up for the sake of preservation. By a new approach of preservation, content, meta data and look-and-feel information have to be preserved to a sufficient degree. Interaction with and modification of a digital object may be necessary at several levels and during several stages of a digital preservation process. It is quite obvious that a co-operative rights management has to replace existing one-off solutions of delivery of works to preservation institution.

Digital preservation actions entail a complex set of approaches, requiring different strategies and tools for each individual situation. In a nutshell, the primary goal of digital preservation is to ensure the long-term accessibility of a digital object. Accessibility in this context means to allow a user to interact with the digital object in such a way as to satisfy the needs for which the object was preserved. Depending on the type of object, this may result in different characteristics of an object being preserved. Consider, for example, a simple document written in a current version of Microsoft WORD. This document has many characteristics that may be relevant to be preserved, such as, primarily, the textual content, including any kind of images, equations and tables that may be embedded (in, potentially, a wide range of other file formats, such as different types of image files, tables in Microsoft EXCEL format, etc.). It may also require the preservation of formatting information, such as headers, the usage of specific fonts, page-breaks and numberings. It may further entail macros, active elements that compute certain values in the document (such as inserting a current date). The fact that the document is editable may be an important characteristics, as may be a change history, information on who edited the document last and at which time, etc. all form part of the complex and proprietary Microsoft WORD file format.

To access such a document requires a specific program (e.g. Microsoft WORD, Wordpad, OpenOffice Word), which in turn require a specific version of a specific Operating System, which in turn usually runs only on a specific hardware platform. If any of these become unavailable at a certain point in time, the document in its original form can no longer be accessed, turning into a useless accumulation of 0's and 1's.

Preservation actions try to mitigate this risk by applying some strategies, the most prominent ones being migration and emulation. Yet, none of these can guarantee the complete preservation of all characteristics of a digital object. The main challenges and their legal consequences in terms of permissible actions will be analyzed in detail below.

The remainder of this article is structured as follows. Section 2 briefly reviews the major challenges as well as proposed solutions for the long-term preservation of digital objects, exemplifying the changes that may be unavoidable during preservation actions. Section 3 addresses the legal challenge regarding the initial acquisition of objects and their ingest into a digital preservation system. Section 4 outlines the copyright law, followed by an analysis of access to documents in Section 5. The "editing for preservation" right is discussed in Section 6, that of reverse engineering in Section 7. A short analysis on data protection is given in Section 8, followed by conclusions and an outlook on potential modifications in Section 9.

## **2 PRESERVATION OF THE DIGITAL HERITAGE – TECHNICAL SOLUTIONS AND LEGAL PROBLEMS**

During the last couple of years, a lot of research has been done to define, improve and evaluate single preservation strategies [29]. A good overview over the state of art was prepared by the National Library of Australia [16] and published by the UNESCO as a handbook [25] accompanying the UNESCO charter on the preservation of the digital heritage. It not only describes specific preservation strategies, but also management and legal issues.

The research on technical preservation issues is focused on two dominant strategies, namely *Migration* and *Emulation*. Scientific results on *Migration*, which is at the current time the most common preservation strategy, were published for example by the Council of Library and Information Resources (CLIR) [13], where different kinds of risks for a migration project are presented. Migration requires the repeated conversion of a digital object into more stable or current file formats, such as e.g. converting a Microsoft WORD97 document into the current Office 2000 format (within format-family migration) or converting it, e.g. to Adobe PDF, a simple ASCII/UNICODE text file, a screenshot image, or others. Each of this incurs certain risks and preserves only a certain fraction of the characteristics of any digital document. Conversion to PDF, for example, changes the look-and-feel of the document as it can no longer be edited, may look differently as some fonts may not be available on a given future computer system, edit history and other metadata are lost. Screenshots may preserve the look of a document, loosing the machine-readable content, i.e. the text, whereas a conversion to a text file preserves the content but loses image information, macro interactivity and others. Even migrations within the same format family may incur unwanted and unspecified changes, as is even warned in a special notice when trying to store a document in a different than the original file format.

Work on the second important preservation strategy *Emulation* was advocated by Jeff Rothenberg together with CLIR [20], envisioning a framework of an ideal preservation surrounding. In order to make *Emulation* useable in practice, several projects developed it further. One of them is the CAMILEON project [10], trying to implement first solutions and to compare *Emulation* to *Migration*. Other important projects in the preservation field are the CEDARS project [5], the PADI project [16] or the "digitale duurzaamheid" project [6]. More recently, the *Universal Virtual computer (UVC)* has been proposed as a promising solution [11, 24]. Emulation aims at providing programs that mimic a certain environment, e.g. emulation a certain processor or emulating the features of a certain operating system, allowing users, for example, to run Microsoft WORD on a Linux operating system using the WINE windows emulator. Specifically for more complex digital objects, such as database applications, games and interactive animations, this may well constitute the only possibility to provide a reasonable preservation of the core characteristics of a digital object. Yet, in spite of it providing in its optimal form an original environment, changes may happen due to general technological advancement. Processing speed may significantly change the speed at which old computer games can now be presented, special input devices may no longer be available and cannot be easily mimicked, the look-and-feel of a piece of art designed for an old low-resolution monochrome monitor may not be feasible, yet crucial to the intentions of the artist. More severely, emulation still faces significant technological challenges, as well as requiring a detailed reconstruction of protected information on the internal structure of both hardware as well as software to be emulated.

In the following we will briefly review some of the stages at which legal limitations may be encountered when applying preservation actions. The type of preservation actions that is required and feasible depends on several issues, specifically:

- The way the digital object is being presented for ingest into a preservation system (special

storage device such as CDs, on-line transfer, etc.) and the means of protection applied (copy protection, access control, cryptographic protection of all or parts of the content/functionality e.g. printing-disallowed, watermarks).

- The type of digital object and its inherent complexity and embedded functionality (e.g. static, single-format document vs. complex document with embedded objects vs. database applications/games).
- The preservation goal, i.e. which type of information from the object needs to be preserved (e.g. content only, look-and-feel, functionality, etc.).

At the moment, the primary focus of activity as well as legal discussion is mostly only on the acquisition of materials. Digital objects, in their completeness or in a selective form, are delivered to e.g. the national library implementing a legal obligation or a particular contractual arrangement. The law (e.g. the obligation to deposit) and the policy (e.g. active collection policy) of the digital preservation institution determine the quantity of the acquisition of materials. The preservation institution obtains and holds the material. Otherwise, an institution may be entitled to actively collect material for preservation, as is the case with some national Web archiving legislation, entitling the digital memory institution to actively harvest pages from the national Web space and store them.

However, long-term preservation requires more rights to handle and/or - specifically – to modify content [25, p. 96]:

At ingest stage, the object will, among other things, be read from the original storage medium (e.g. a CD, DVD, floppy disk, magnetic tape) and transferred to the storage medium of choice in the digital preservation system. This requires the need and permission to access the object and read it in its entirety, migrating any potentially conflicting encoding scheme between different storage devices, bypassing copy-right protection and access control mechanisms.

At the same time, meta-information on the objects will be acquired, be it either manually (i.e. manual editing of metadata information describing the object) or automatically extracted from the object itself (e.g. for digital camera images extracting EXIF header information from JPEG files containing timestamp information when a picture was taken, the camera model it was taken with, shutter speed and aperture, resolution information, etc.) This may require special permission to analyze the internal structure of a digital object, reverse-engineering file format characteristics.

Depending on the chosen preservation strategy, several different processing steps may be applied repeatedly over time, and result in specific legal challenges, such as:

- Extraction: The important content of the digital object with respect to preservation requirements is extracted and preserved separate from the original object (e.g. extracting all text information from a PDF file and storing it as ASCII or UNICODE text file). This loses a wide range of other information characteristics present in the file (e.g. images) that may not be considered relevant for the preservation task at hand, or simply too difficult to preserve, modifying the original object, permanently removing parts of its original content) thus in fact changing the object, splitting part of its content and characteristics into several different objects.

- Conversion: The file may need to be converted to another file format, be it in the same format family (e.g. migrating a WORD 97 document to Word 2000), or into a different format (e.g. converting a WORD 97 document into RTF, PDF, Postscript, OpenOffice, TIFF or JPEG file), again modifying content, functionality, look-and-feel, sometimes at a degree that cannot be determined in advance resulting in a different than the original object.<sup>3</sup>
- Standardisation: Files may be transformed in any way to conform to certain standards imposed by the digital preservation system (such as e.g. down-sampling all image files to a certain resolution, certain colour depth, audio files to a certain bit sampling rate, etc.). If performed prior to ingest into the preservation system, the resulting modifications are performed by the content owner, thus reducing legal challenges. Still, in less controllable environments, this standardization may have to be performed within the preservation system, resulting in the same conflicts as mentioned above. Furthermore, even objects in standardized file formats will eventually be subjected to some kind of further preservation action as a standard drops out of use.
- Encapsulation: Most objects require a specific infrastructure to be accessed, e.g. a website containing a flash animation may require a special version of a flash player, running within a specific Web browser, running on a specific operating system on a given hardware platform. These may be stored together with the object in a kind of virtual storage capsule, distributed across potentially several computing systems. However, specific exemptions from licensing laws may be necessary to permit the multiple storage, installation and invocation of software.
- Emulation: In order to be able to emulate a certain system, either an emulator needs to be provided by the producer of the original software or hardware, which currently is not the case for most standard applications and systems (partially also due to the high complexity of such endeavours). Re-creating such an emulator externally will usually require the de-assembling of the original software, or the microscopic analysis of a hardware chip, extracting the internal structure, to allow a complete reverse-engineering of the required functionality.

Resulting copyright problems are:

- Digital files or specific objects require usually protected software for presentation and use. Thus, it may be necessary to acquire rights and store not only the object as such, but also the complete environment including all layers of software and hardware or hardware emulators that are required. A complex file may incorporate several independent file formats, each of which may contain specific limitations regarding actions permitted.
- Technological measures for copyright protection and access control constitute a specific problem. Such impediments may need to be removed or bypassed for preservation activities. However, they may need to be re-imposed or at least re-imposable upon access – be it to ensure confidentiality of the documents, or to simply demonstrate the look-and-feel and functionality of a digital object (who had the right and the power to access which part of

<sup>3</sup> In some cases, even seemingly trivial changes may constitute significant obstacles, specifically when it comes to e.g. pieces of art: artist may, for example, write poems using a specifically designed font set, and using a very specific spacing and arrangement of characters on the page layout, all of which may undergo subtle and unavoidable, sometimes even undetectable or undocumented changes in migration steps (given the resources available for performing large-scale preservation activities).

information in a specific object).

- Preservation actions may entail necessary changes, that sometimes are subtle and hard to detect. Limitation of liabilities for damages incurred by these unnoticed changes are a crucial issue, as is a need and method to specify which characteristics of a digital object are crucial and may thus not be changed under all circumstances.

In addition to that, providing access to a digital archive needs regulations concerning the following questions:

- Time and extend of providing access requires balancing the needs of both the digital preservation institution, the content owner, the range of rights owner in the value chain, as well as the individual with respect to privacy issues.
- Collection of user information (e.g. a Web archive allowing to trace all individual statements and changes made by a person across the world throughout his or her entire life) may require specific legal access limitations as the collected data as such constitutes a completely new set of information, potentially severely inflicting privacy regulations or requirements.

The question of a “publication” will need ongoing adaptation to the various new forms of electronic communication. At the moment, it is discussed under which conditions a website and various other new forms such as mailing lists, discussion forums, log files and others, are a publication (being available = published, minimum number of users? etc.) and when they are actually also be perceived as such by the persons creating them.

The legal situation can be described as ambiguous. Many jurisdictions are still in the process of clarifying legal frameworks of rights and how they should be managed. Fragmented or collaborative creation of digital materials makes the identification of rightholders very difficult or impossible. Good rights management practices are not sufficiently in place. Voluntary contractual arrangements offer only a solution for a small period of time in order to master the various preservation technologies. Therefore, the lawmakers should establish a sufficient legal environment for preservation of the digital heritage.

### **3 DEVELOPMENT OF NEW RULES FOR THE DEPOSIT OF ELECTRONIC PUBLICATIONS**

As shown above, the principle of “simple” deposit of publications is not sufficient for the digital world. Even the basic principle is jeopardized as it is no longer secured that preservation institutions may own a copy with an exhausted distribution right. Such rights are typical for analogue, tangible and static objects like books and journals but not for digital publications. In many cases electronic publishers rent the tangible product or provide on-line access that prevents exhaustion of distribution and forbids lending. Therefore, a different approach is needed leading to a change in the conditions for use and copying of the deposited electronic material [21, 22].

Countries with provisions on legal deposit of electronic publications (see for an overview [4, 8, 17, 27]) have mostly extended only the obligation to deposit to some or all such digital

materials (e.g. France [2] or Austria<sup>4</sup>). More advanced are provisions allowing preservation measures (e.g. USA<sup>5</sup> or Canada<sup>6</sup>). Explicit provisions on ingest to a digital preservation system, acquisition of meta-information, extraction of digital objects, conversion, standardization or encapsulation are still not existing in legal deposit laws.

The Google's Library Project [9] may represent the best example of such a preservation system. Although the project aims only at "[a]n enhanced card catalog of the world's books", the archive of scanned books with a full text book search shows many similarities. Rightholders have strongly criticized the approach of Google of developing such an "extended card catalogue" without a sufficient contractual basis with the authors [1].

As preservation institutions do not have the necessary rights so far for a long-term preservation, contractual arrangements may provide a solution for the time being. However, clearing rights is difficult and resource-intensive. In this paper, it can only be referred to some existing models like the *NESLi2 Licence for Journals* [18] or the *OCLC/RLG Working Group on Preservation Metadata* [19]. Although these contractual arrangements seem to be successful it is evident that they have to be incorporated in legal deposit acts in the near future in order to guarantee the completeness of the collection of the national heritage.

#### **4 COPYRIGHT AND PRESERVATION OF ELECTRONIC PUBLICATIONS**

The WIPO Copyright Treaty [28]<sup>7</sup> does not harmonize the regime of limitations and exceptions to the rights granted to authors of literary and artistic works (Article 10 WIPO Copyright Treaty). In an Agreed Statement concerning Article 1(4) of the WIPO Copyright Treaty, the contracting parties recognize that the exceptions (Article 9 of the Berne Convention) „fully apply in the digital environment“. The Agreed Statement concerning Article 10 permits contracting parties to carry forward and appropriately extend limitations and exceptions into the digital environment. The limit will be always the so-called tree steps test. Such limitations or exceptions are allowed only for special cases. They should not conflict with a normal exploitation of the work and should not unreasonably prejudice the legitimate interests of the author (Article 10 (2) of the WIPO Copyright Treaty).

Legal deposit or permanent storage of electronic publications is without question a special case. This exception is well established in existing copyright laws of contracting parties of the Berne Convention. Normally, access to deposited material does not conflict with a normal exploitation of the work because access can be given only to a limited number of users at the same time on-site. This may be seen differently when the commercial exploitation period of the product is run out as no legitimate interests of the authors or rightholders are prejudiced. On the contrary, the preservation institution acts in the interest of the authors. Preservation of the

<sup>4</sup> Anbietungs- und Ablieferungspflicht bei sonstigen Medienwerken nach dem Mediengesetz, Official Gazette (Austria) II No 65/2001.

<sup>5</sup> Digital Millennium Copyright Act, Title IV (DMCA) 1998 ([www.loc.gov/copyright/legislation/dmca.pdf](http://www.loc.gov/copyright/legislation/dmca.pdf)). The DMCA expressly allows three digital preservation copies of an eligible copyrighted work.

<sup>6</sup> Library and Archives of Canada Act - <http://laws.justice.gc.ca/en/L-7.7/index.html> - and National Library Book Deposit Regulations, 1995 - <http://laws.justice.gc.ca/en/N-12/SOR-95-199/index.html>.

<sup>7</sup> Cf. <http://www.wipo.int/treaties/en/index.html>.

electronic publications belongs to the legitimate interests of the authors. The preservation institution can provide access to the materials resulting in broader distribution and make the publication available to the public when the product is sold out or not distributed any more by the rightholder. This *mind-share* for the authors should not be underestimated.

The main aim of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society<sup>8</sup> (Copyright Directive) ([12], [14] p. 495, [30] p. 1009) is the creation of a general and flexible legal framework „in order to foster the development of the Information Society in Europe“ (recital 2). The protection of the substantial investment in creativity, innovation and network infrastructure should „foster substantial investment in creativity and innovation, including network infrastructure, and lead in turn to growth and increased competitiveness of European industry, both in the area of content provision and information technology and more generally across a wide range of industrial and cultural sectors“ (recital 4).

Implementing the international obligations of the WIPO Copyright Treaty, the Copyright Directive grants the authors a reproduction right (Article 2), a right of communication to the public, including the right of making available works or other subject matter (Article 3) and a distribution right (Article 4). The exceptions (Article 5) are exhaustive.

Article 5(2)(c) of the Copyright Directive - the relevant provision for preservation - allows exceptions *“in respect of specific acts of reproduction made by publicly accessible libraries, educational establishments or museums, or by archives, which are not for direct or indirect economic or commercial advantage”*.

Therefore, it is up to the national law to find appropriate solutions for preservation of digital materials. The obligation to legal deposit is easily covered; migration to new media can still be subsumed as an act of reproduction. However, “editing for preservation” does not comply with this provision. Such acts are the prerequisite of the author and are not covered by any of the exemptions of copyright.

The Copyright Directive protects technological measures and rights-management information as any circumvention of any effective technological measures is prohibited (Article 6 of the Copyright Directive). Preservation requires sometimes a circumvention of effective technological measures. The Copyright Directive does not contain an explicit exception but Member States should take appropriate measures to ensure that the exceptions or limitations of Article 5(2)(a), (2)(c), (2)(d), (2)(e), (3)(a), (3)(b) or (3)(e) are available to the beneficiaries. It is unclear if a Member State can allow a preservation institution to circumvent effective technological measures.

## **5 ACCESS TO DOCUMENTS AND COPYRIGHT**

The question of access comprises a right of communication to the public of the deposited material and the right to lend tangible materials. Users should be allowed to make copies of

<sup>8</sup> OJ L 167, 22/06/2001, p. 10. Proposal: COM (97) 628 final.



parts of the electronic product within the limits of the national regime for copying. In practice, the distinction between on-site and on-line access is decisive.

Article 5(3)(n) of the Copyright Directive is quite clear in this question in restricting access to on-site users in allowing to “[u]se by communication or making available, for the purpose of research or private study, to individual members of the public by dedicated terminals on the premises of establishments referred to in paragraph 2(c) of works and other subject-matter not subject to purchase or licensing terms which are contained in their collections”. Preservation institutions can display bought or rented electronic products on-site and can allow users to view, listen and browse the material for the mentioned purposes.

Personal reproduction is dealt with in Article 5(2)(b) of the Copyright Directive: It is allowed “in respect of reproductions on any medium made by a natural person for private use and for ends that are neither directly nor indirectly commercial, on condition that the rightholders receive fair compensation which takes account of the application or non-application of technological measures referred to in Article 6 to the work or subject-matter concerned”. Preservation institutions can allow – taking into account the national law - on-site users to make a digital or paper copy for private or educational purposes provided that the authors and publishers get fair compensation.

On-line use – e.g. access and dissemination of the deposited or permanently stored electronic publications by the preservation institution – may spoil the economic interests of authors and rightholders. The three steps test of Article 10(2) of the WIPO Copyright Treaty does not lead to a clear result and it seems to be left to the national legislator to balance between the rights of authors and publishers and the necessities of the preservation institution. However, the Copyright Directive seems to leave so far no loophole for on-line access to deposited material. Therefore, rightholders may forbid or limit access to the deposited or permanently stored electronic product or the copying of documents on-line. The preservation institution has again to refer to contractual arrangements.

It seems that the Copyright Directive is too rigid in this direction as the three steps test of Article 10(2) of the WIPO Copyright Treaty does not prevent a more liberal approach to on-line access. In the quite many cases where the period of commercial exploitation of the work has expired but the quite lengthy period of copyright still lasts it should be allowed for preservation institutions to give on-line access.

## **6 “EDITING FOR PRESERVATION” AND COPYRIGHT**

Long-term preservation requires besides extensive copying and migration in an archiving system also some “editing” of materials, in particular necessary modifications due to the preservation methods or changes in hardware and software. Copyright law does not allow any changes of the work without the permission of the author. At the moment, only contractual arrangements between authors, publishers and preservation institutions can give them the rights for such “editing for preservation”.

In the medium-term, it is obvious that such rules have to be included in the deposit laws. However, as such editing seems to infringe existing copyright treaties, an amendment of them would be necessary. It is evident that very precise rules for such editing have to be developed. Contractual arrangements may be very helpful for establishing a proper legal framework.

The ongoing discussion between open source and copyright may also provide some guidance as – in the long-term – every piece of cultural heritage will be public domain and accessible to everyone. Preservation will secure that the work as such still exists and remains accessible.

## **7 REVERSE ENGINEERING**

Authenticity of long-term preservation requires maintaining to the utmost possible degree the same environment of presentation, e.g. the look-and-feel, the functions of the various programmes and data formats etc. As mentioned above, storage and migration will be performed by a huge computer system containing the national heritage. In the medium term, existing hardware and software will be useless for access to these documents. Current PCs with current versions of Microsoft WINDOWS and Microsoft OFFICE will disappear and with them the hardware and the software package allowing access in a technical way. Therefore, this environment may have to be emulated by reserve engineering of functionalities of hardware and software.

Decompilation and reverse engineering of such programmes may infringe the copyright of the rightholders. Article 6 of the Council Directive 91/250/EEC of 14 May 1991 on the legal protection of computer programs<sup>9</sup> is quite restrictive ([14], Blocher in [30], 205). On the one hand, the “authorization of the rightholder shall not be required where reproduction of the code and translation of its form within the meaning of Article 4 (a) and (b) are indispensable to obtain the information necessary to achieve the interoperability of an independently created computer program with other programs”. Preservation institutions have licenses of such programmes and re-engineering is necessary to achieve interoperability. On the other hand, however, the aim is not interoperability but substituting the programme as such and thus not covered by the Copyright Directive.

Therefore, preservation institutions have to conclude contractual arrangements allowing reengineering of existing hardware and software solutions for the purpose of access to the national heritage. In the medium term, preservation institutions should have statutory rights for reverse engineering of standard IT environments.

## **8 PROTECTION OF DATA**

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data<sup>10</sup> (Data Protection Directive) [7] provides sufficient guidance as keeping of user data is not allowed without an agreement of the users or if no legal provision or contractual obligation provides a legal basis (Articles 6 and 7 of the Data Protection Directive). In practice, the balancing between the – quite automatic – data acquisition and the – technically unsolved – “forgetting of a data repository” remains a tricky issue.

More problematic is the fact that huge archiving repositories will allow a much more sophisticated access to personal data. Names with some context data are quite good search

<sup>9</sup> OJ L 122, 17/05/1991 p. 42.

<sup>10</sup> OJ L 281, 23/11/1995 p. 31.

terms even in such huge archives. It cannot be presumed that the various data subjects have agreed that published personal information may be stored forever and searched appropriately. Establishing some limits for access or anonymisation can be a quite appropriate solution, like the practise established by courts for the publication of its decisions<sup>11</sup> [3, 23].

## 9 CONCLUSIONS

The preservation of electronic publications turns out to be a much more difficult task than that of other conventional media. Simply acquiring the object via deposit or active collection is not sufficient. On the contrary, significant interventions into a set of rights are required in order to be able to extract the information necessary from a given digital object and its software and hardware environment. Additionally, changes to the object as a result from preservation actions are unavoidable, requiring permission to inflict changes to an object in order to preserve it.

The WIPO Copyright Treaty leaves this important question to the national legislation. However, limitations and exceptions are subject to the three steps test. The Copyright Directive allows reproduction and communication to the public for the particular purposes of preservation institutions. So far, implementing laws are quite conservative and extend only the obligation to deposit. Extensive preservation rights are still subject to contractual arrangements. Yet, a more international solution to these challenges is definitely desirable, requiring international discourse and agreements.

## 10 REFERENCES

- [1] AUTHORS GUILD, Authors Guild Sues Google, Citing “Massive Copyright Infringement”, Website, September 2005 Press Release, [http://www.authorsguild.org/news/sues\\_google\\_citing.htm](http://www.authorsguild.org/news/sues_google_citing.htm) (2006).
- [2] BIBLIOTHEQUE NATIONAL DE FRANCE, Guide du dépôt légal 2004 destiné aux établissements dépositaires du dépôt légal imprimeur, [http://www.bnf.fr/pages/infopro/cooperation/po\\_dli\\_2003.htm](http://www.bnf.fr/pages/infopro/cooperation/po_dli_2003.htm) (2006).
- [3] BING, J., Electronic Copying, Complex 10/90, Tano, Oslo (1990).
- [4] CHARLESWORTH, A., Legal Issues Relating to the Archiving of Internet Resources in the UK, EU, USA and Australia: a Study Undertaken for the JISC and Wellcome Trust, [http://www.jisc.ac.uk/uploaded\\_documents/archiving\\_legal.pdf](http://www.jisc.ac.uk/uploaded_documents/archiving_legal.pdf)(2003).
- [5] CURL EXEMPLARS IN DIGITAL ARCHIVES, University of Leeds, Website, <http://www.leeds.ac.uk/cedars> (2006).
- [6] DIGITALE DUURZAAMHEID, ICTU, Den Haag, Netherlands, Website, <http://www.digitaleduurzaamheid.nl> (2006).
- [7] DROBESCH, H., GROSSINGER, W., Das neue österreichische Datenschutzgesetz, Juridica, Wien (2000).
- [8] GATENBY, P., Legal deposit, electronic publications and digital archiving – The National Library of Australia’s experience, <http://www.ifla.org/IV/ifla68/papers/071-124e.pdf>, Glasgow (2002).
- [9] GOOGLE BOOKS LIBRARY PROJECT, <http://books.google.com/googlebooks/library.html> (2006).
- [10] HEDSTROM, M., LAMPE, C., Emulation vs. Migration. Do Users Care?, RLG DigiNews, Dec. 2001 , Vol. 5, No. 6 (2001).

<sup>11</sup> The Austrian courts can be considered as good practice. Cf. <http://www.ris.bka.gv.at>.

- [11] HOEVEN, J.R. VAN DER, DIESSSEN, R.J. VAN EN MEER, K. VAN DER, Development of a Universal Virtual Computer (UVC) for long-term preservation of digital objects, *Journal of Information Science*, Vol. 31 (3) (2005).
- [12] KUCSKO, G., *Geistiges Eigentum*, Manz, Wien (2003).
- [13] LAWRENCE, G. W., KEHOE, W.R., RIEGER, O.Y., Walters, W.H. KENNEY, A.R., *Risk Management of Digital Information: A File Format Investigation*, CLIR, Washington, DC. (2000).
- [14] LLOYD, I. J., *Information Technology Law*, fourth edition, Oxford University Press (2004).
- [15] MUIR, A., *Copyright and licensing for digital preservation*, Update, June 2003 (2003).
- [16] NATIONAL LIBRARY OF AUSTRALIA, PADI Preserving Access to Digital Information, Website, <http://www.nla.gov.au/padi> (2006).
- [17] NATIONAL LIBRARY OF AUSTRALIA, PADI Preserving Access to Digital Information, Website, Legal deposit, <http://www.nla.gov.au/padi/topics/67.html> (2006).
- [18] NESLi2, Model Licence for Journals, <http://www.nesli2.ac.uk/model.htm> (2006).
- [19] OCLC/RLG Working Group on Preservation Metadata, *Preservation Metadata and the OAIS Information Model: a Metadata Framework to Support the Preservation of Digital Objects*, [www.oclc.org/research/pmwg/pm\\_framework.pdf](http://www.oclc.org/research/pmwg/pm_framework.pdf) (2002).
- [20] ROTHENBERG, J., *Avoiding Technological Quicksand: Finding a viable technical foundation for digital preservation*, CLIR, Washington, DC. (1999).
- [21] SAMUELSON, P., *Copyright and Digital Libraries*. In: *Comm. of the ACM*, Vol. 38, No. 3 (1995).
- [22] SAMUELSON, P., *Encoding the Law into Digital Libraries*. In: *Comm. of the ACM*, Vol. 41, No. 4 (1998).
- [23] SCHWEIGHOFER, E., *Vorüberlegungen zu Data Mining und Datenschutz am Beispiel Österreichs*, *Datenschutz und Datensicherheit* 21 (1997).
- [24] ULLRICH, D., *Digitale Langzeitarchivierung*, *Jahresbericht 2004 der Gesellschaft für wissenschaftliche Datenverarbeitung*, <http://www.mpg.de/bilder/BerichteDokumente/dokumentation/jahrbuch/2004/gwdg/forschungsSchwerpunkt1/pdf.pdf> (2004).
- [25] UNESCO, *Guidelines for the Preservation of Digital Heritage*, Prepared by the National Library of Australia, CI-2003/WS/3, March 2003, Paris (2003).
- [26] UNESCO, *Memory of the World*, [http://portal.unesco.org/ci/en/ev.php-URL\\_ID=1538&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/ci/en/ev.php-URL_ID=1538&URL_DO=DO_TOPIC&URL_SECTION=201.html) (2006).
- [27] UNIVERSITY OF AMSTERDAM, INSTITUTE FOR INFORMATION LAW, *Copyright Aspects of the Preservation of Electronic Publications*, *IViR Reports* 7, [www.ivir.nl/Publicaties/koelman/KBeng2.doc](http://www.ivir.nl/Publicaties/koelman/KBeng2.doc) (1998).
- [28] VINJE, TH. C., *The New WIPO Copyright Treaty: A Happy Result in Geneva*, *EIPR*, Vol. 5 (1997).
- [29] WALLE, J. V. D./MACKENZIE OWEN, J. S., ELDEP, *A study on issues faced by national libraries in the field of deposit collections of electronic publications*, *Final Report*, NBBI (1996).
- [30] WALTER, M. (Hrsg.), *Europäisches Urheberrecht, Kommentar*, Springer, Wien New York (2001).

*Cited websites were last visited in January 2006.*