Electronic Resources Management and Long Term Preservation (Is the library a growing organism?)

bv

Tommaso Giordano

European University Institute Library Fiesole (Florence)

Preface¹

Digital preservation is a vast and complex issue which involves many aspects and areas of expertise, as is evident from the range of topics discussed during this conference. My paper is rather limited in scope and seeks to develop considerations concerning three questions: the first investigates how the problem of archiving and preservation of electronic journals is perceived by academic libraries, while the second concerns the organizational practices implemented in this area; I conclude with a comparison between the traditional approach and the emerging models of work flow for library activities in this area. The main sources are the professional literature, more specifically surveys and recommendations on electronic collections management and preservation carried out at different times in the academic libraries of the United States and Europe.

Ownership and licensing

In the field of e-journals the license controls access to three levels of data: a) the current year; b) back issues; c) the copy of the current year, provided by the publisher in a determined format, which can be installed on the local site as an archive copy. According to the definition of the Digital Library Federation (DLF), in this paper, 'perpetual access' designates the right to permanently access the licensed materials paid for during the period of the license agreement, while 'archiving right' defines the right to permanently retain an electronic copy of licensed materials for preservation purposes².

The purchase of a print publication has nothing in common with a licensing contract, in the same way that buying an apartment has nothing in common with a rental contract. In the first case we are talking about the acquisition of a physical artifact that automatically provides permanent possession (ownership) of the object involved in the transaction. A license, instead, is a contract that defines the terms of use of a resource for a determined time period; in practice, through this mechanism, libraries rent the use of an information resource without acquiring ownership (and without 'capitalizing') and remain without it when the contract expires. The cancellation, in a financially difficult year, of a subscription to a periodical title in paper format does not jeopardize the possession of the preceding years on the shelves; on the other hand, after the cancellation of a licensing contract, the library is often left with nothing.

The use of a book is governed by copyright law, which gives libraries the right to offer it for free consultation, the right to lend it, the right to reproduce it (regulated by norms which differ from country to country), and the right of perpetual archiving. The electronic license, instead, is governed by a private contract, based on which access is guaranteed for a determined length of time according to methods established by specific clauses negotiated at the time of the contract on a case by case basis. These clauses establish the services allowed (for example ILL / document supply, print and download, etc) and define the users authorized for access; possible variations of the clauses (for example, granting access to the alumni of a university) must be renegotiated with the publisher/provider and have economic consequences. Since they do not have ownership of the resource provided by the license, libraries do not have the right to archive it without the specific authorization of the publisher. In some cases the publishers guarantee perpetual access of previous years, but this provision (about the "perpetuity" of which there are serious doubts) is the object of a specific clause to be negotiated. The archiving clause, too, based on which one acquires the right to preserve a copy of the digital resource obtained from the publisher, is subject to specific negotiation, during which the specifications and the format of the copy that will be delivered to the library are defined. Obtaining the archive copy is just the first step, even if it is an important one, in the process of preservation.

¹ The author wishes to thank Veerle Deckmyn, Aimee Glassel and Fred Birkhimer for their comments on the first draft of this paper.

² Electronic Resource Management: Report of the DLF Electronic Resources Management Initiative. Appendix D: Data Element Dictionary. Washington, D.C. Digital Library Federation, 2004. http://www.diglib.org/pubs/dlf102/

The precondition for passing to the subsequent phases is the availability of adequate infrastructure (a considerable economic commitment) and the necessary expertise to implement long term maintenance procedures and any operation required by technological changes (migration, emulation, etc.). A recent study on 12 important long-term archiving programs, published by CLIR (the Council on Library and Information Resources) ³ describes the complexity of the preservation process, above all when standards and technologies are still as of yet insufficiently developed. In contrast with paper preservation, which could tolerate inertia and discontinuity (it is not unusual for books, or even entire libraries, abandoned for decades to be subsequently recovered without serious damage), digital preservation requires active and constant maintenance. This could mean that, in the digital environment, preservation is a responsibility which not may be possible for every library.

Perceptions and practices in the digital context

The terms of the problem were already largely identified in the first phase of the diffusion of e-journals⁴⁵. In this period the main objective of library managers was obtaining the widest possible availability of on-line resources and reinforcing the technological infrastructures of their sites in order to efficiently enhance user access. In any case, at that time the archiving function was assured by the paper copy that the libraries continued to acquire, also because, generally speaking, the licensing models imposed the combined acquisition of both versions (electronic and paper). Already in this period the licensing models suggested by library organizations contained "perpetual access" clauses (i.e. NESLI, ICOLC, EBLIDA, etc.), but these provisions were considered a minor aspect of the contract, while the "core" part of the negotiations with the publishers concerned issues of more immediate impact such as price, IP access, definition of the "authorized users," back issues, clauses for downloading and ILL / document supply. Some studies carried out in other areas and time periods demonstrate that, in practice, the attitude of libraries concerning digital preservation has not made much progress since then. In 2001 a survey of the Boston Library Consortium⁶ showed scarce consideration for archival issues at the level of license negotiations. The archiving clauses were ranked at 5th-6th place in order of importance in the negotiation of licenses. As Jennifer Watson suggested, "the low importance placed on archiving may be partially due to the fact that only 10% of surveyed libraries cancelled print subscriptions." As a result, the libraries, at this time, were not motivated to push for obtaining archival rights. This justification, however, does not seem applicable to the 19 higher and further education institutions in the UK surveyed by JISC in 20038, where it emerges that many of these institutions acquire periodicals in e-only format because they are more convenient for the users and more economical in terms of management cost and storage space. Not much different data emerges from a 2003 survey of ARL, in which only 15% of the libraries request a clause for perpetual access. 9 This attitude is further confirmed by a research carried out in 2004 which shows that 60% of licenses did not include archival rights and 55% did not include perpetual access.¹⁰

In more recent years, with the rush towards the e-only license model, and the resulting concern for the loss of the paper back-up, a greater awareness of the problem has emerged. The impulse, however, has been slow to result in concrete actions for various reasons of a practical nature, among which the scarce availability of resources to invest; at the base of all these reasons can be found the conviction that there are other priorities more urgent than preservation. This attitude is common in both Europe and the United States, as several

³ E-Journal Archiving Metes and Bounds: A Survey of the Landscape by Anne R. Kenney, Richard Entlich, Peter B. Hirtle, Nancy Y. McGovern, and Ellie L. Buckley. Washington D.C., CLIR, September 2006.

⁴ Ann Okerson, Who Owns Digital Works?, "Scientific American", July 1996, p.80-84

⁵ Sandy Norman, Copyright and fair use in the electronic information age, "IFLA Journal", 23 (1997), p.295-298

⁶ Jonathan Nabe, *E-journal Bundling and its Impact on Academic Libraries: Some Early Results.* "Issues in Science and Technology Librarianship", Spring 2001 http://www.istl.org/01-spring/article3.html

⁷ Jennifer Watson, *You Get what you pay for? Archival Access to Electronic Journals*, "Serials Review", 31(2005), 3. p.200-2005

⁸ Maggie Jones, *Archival E-Journals Consultancy: Final Report*. Report Commissioned by the Joint information Systems Committee (JISC), October 2003. http://www.jisc.ac.uk/uploaded_documents/ejournalsfinal.pdf

⁹ Mary M. Case, *A Snapshot in Time: ARL Libraries and Electronic Journal Resources*. "ARL Bimonthly Report", 235 (August 2004). http://www.arl.org/newsltr/235/snapshot.html

Sharon Farb, *Libraries, Licensing and the Challenge of Stewardship*, "First Monday", 11 (2006), 7. www.firstmonday.org/issues/issue11 7/farb/index.html

studies confirm¹¹. This implies the opinion, at times explicitly stated, that preservation is an issue that does not concern every library; it should instead be delegated to libraries and agencies with this specific responsibility and to the publishers.

A survey presented at the Fall 2006 ICOLC meeting offers an insight into the position of library consortia on the preservation issue. ¹²As is well known, the consortia are, together with the libraries and the publishers, major players in the field of the distribution of scientific information. Their role is becoming increasingly influential both because they represent more or less vast aggregates of libraries and because they are the main partners in the negotiation of licenses with publishers. ¹³ In the survey 35 consortia participated (18 from North America, 13 from Europe, 1 from India, 1 from Japan, and 2 international, for a total of 1241 licensing contracts). All the consortia declared that they considered long-term preservation a matter of primary importance. However, in the responses about the priority given to the archiving clause in the licensing negotiations, this promise appeared largely unfulfilled. In fact, in the question regarding which clauses are specifically requested in the contracts, only 14 consortia reported requesting a copy for the permanent archive; only a few were able to provide details on the format of data provided, which could mean that the data acquired for the archive are not used. If nothing else, this demonstrates the discrepancy between the intentions declared in theory and the practices implemented in reality concerning digital preservation. This same attitude emerges from another survey carried out in 2005 by the University of Innsbruck Library¹⁴ on 335 libraries of different types (national, public, and university) in the 25 countries of the European Union. The overwhelming majority of those interviewed claimed to consider long-term preservation a "very important" (75.6%) or "rather important" (22.9%) task; in spite of this, only 37.7% reported having begun programs for digital archiving, some of which (about half) had not yet even chosen the software to adopt. In conclusion, from the studies here briefly summarized there emerges a cultural awareness of the problem of digital preservation which, however, in the majority of cases is not met by consequential actions on the practical or policy level. In other words, the priority of libraries is today that of maximizing access; while recognizing the strategic value of preservation, they still believe that it is a task to demand of others (national libraries, trusted third part repositories, consortia, publishers). A short-term vision emerges, where the statistics of use and a discount, however modest, from the publisher carry more weight than any political or cultural consideration. It should be investigated how much of this tendency derives from the crashing waves of the digital revolution and how much, instead, from older and more deeply-rooted convictions.

Perceptions and practices in the pre-digital context

Some indications in this regard can be found in the debate concerning collection management several years ago, when digital technology had arrived but its impact on library activities was still rather limited. Particularly interesting is a large in-depth study on preservation policies and practices in British libraries in the early 1990s conducted by John Feathers, Graham Matthews and Paul Eden¹⁵. The research, based on a survey carried out in 488 libraries of various types (a large representative sample of public, academic and special libraries) shows how, in the majority of libraries considered, "preservation is a very minor concern, or not designated as such and essentially aimed at insuring that current materials are available in sufficient quantities and in usable conditions, while drawing on external resources – both document supply services

¹¹ Jennifer Watson, You Get What You Pay for?, cit.

¹² Tommaso Giordano, *Survey on Archiving and Long Term Preservation - A View from the Consortium Side*. "ICOLC Paper presented at ICOLC Fall 2006 (8th European Meeting), Roma. http://www.aepic.it/conf/viewabstract.php?id=178&cf=7.

¹³ Tommaso Giordano, *Library Consortia in Europe*, "Encyclopedia of Library and Information Science" 1.1 (2005). 10 Jan. 2007,

http://www.dekker.com/sdek/abstract~db=enc~content=a713602119

Monika Krimbacher; Michael Neuhauser, Martina Vogl, Survey on the Long-Term Preservation of Digital Documents in European Libraries. Re-USE. Project; University Innsbruck Library, 2005, http://www.uibk.ac.at/reuse/docs/d_6.7study_european_digital_repositories.pdf

¹⁵ John Feather, Graham Matthews, and Paul Eden, *Preservation Management: Policies and Practices in British Libraries*. Aldershot (UK), Brookfield (USA): Gower, 1996

and network services – for other material and data" ¹⁶. "We do not preserve but rather exploit our stock" ¹⁷, as one university librarian interviewed during the research efficiently synthesizes this philosophy.

On the other hand the contradiction emerges between the growing awareness of the problem on the part of most library professionals and the preservation programs of individual libraries. "There is a perceptible gap between the aspirations and achievements of many librarians." In most libraries the preservation activities are not listed as such or the library policy regarding this specific function is not explicitly declared. In libraries there exists a *de facto* activity of preservation, evident in the procedures of collection management, a kind of 'passive preservation' or, rather, we can talk of "preservation by inertia," using the exact expression of the authors in the study already cited: "there is no time to weed, or no incentive to do so. But some, perhaps more, seems to be an expression of a genuine belief in the long term value of the collection and a consequent desire to preserve them for future generations." This attitude can be found in many other countries. More recently the observations of J.P. McCarthy, from the University College Cork (Ireland), confirm these attitudes: "There was never an underlying commitment to a policy of a comprehensive collection; after all, we are a teaching and research institution and our primary aim has been to serve current need and demand" ¹⁹.

On the conceptual level there has emerged a surprising coincidence between the attitudes of librarians concerning digital preservation as discussed above and the policies which emerged from the study of British libraries carried out in a context which was still substantially paper-based, as was the situation in the early 1990s.

We can say, in general, that the development of print collections in most university libraries has been realized according to the "just in case" model, that is to say, the accumulation of material as a means to enhance availability in case of need. "In this sense," admitted McCarthy, "it is more realistic to say that what we have collected is a reflection of passing need rather than a long term one."

How many librarians can identify with this honest pragmatism! In effect, the practice of "preservation by inertia," fed by collections developed by the pressures of "passing needs," has deeper and more widespread roots than was initially apparent. It is based on the assumption that the redundancy of the collections and the multiplicity of the libraries, including the national libraries, insure the preservation of cultural heritage. This mindset—generally followed in the library world— has worked reasonable well until now, permitting us to find a document wherever it is located, and preserving the cultural heritage as well. Unfortunately, this system is not transferable to the digital context!

Models

A comparison between the traditional approach (pre-digital context) and the emerging organizational model (digital context) can help us better define the relationships between collection management and the activities aimed to insure the long term preservation of library resources.

The need to manage efficiently electronic resources and to incorporate them into their services pushes libraries to profoundly reconsider the procedures and the technical tools to adapt to the changing situation. One of the principal aspects of this evolution concerns the workflow for electronic resources processing. In this context the term "work flow" indicates the totality of library tasks and operations necessary for the efficient management of the resource during the different phases of its life cycle. The new management systems which support these operations—Electronic Resource Management Systems (ERMS)—refer to procedures agreed upon in the library environment, 2 including the following phases: a) selection (product identification and consideration, and trial); b) acquisition (licensing negotiation, technical evaluation,

¹⁷ Ivi,, p. 78

¹⁶ Ivi, p. 79

¹⁸ Ivi, p. 37

¹⁹ J.P. McCarty, *The Print Block and the Digital Cylinder*, "Library management", 26 (2005), 1-2, p. 89-96 ²⁰ Ihidem

²¹ Timothy D. Jewell, *Selection and Presentation of Commercially Available Electronic Resources*, Washington, D.C.: Digital Library Federation and Council on Library and Information Resources, 2001. Available at http://www.clir.org/pubs/abstract/pub99abst. Ellen Finnie Duranceau, *Beyond Print: Revisioning Serial Acquisitions for the Digital Age*, "The Serial Librarian", 33 (1998), 1-2, p.83-105.

²² Electronic Resource Management: Report of the DLF Electronic Resources Management Initiative, cit.

business negotiation); c) implementation process (including cataloguing and promotion); d) maintenance and review.

In the following table some characteristics of the two models are identified. As can be seen, while the decision-making process begins in selection phase (a), its crucial moment is actually located in acquisition phase (b) and is connected to the process of negotiating the license with the supplier. This activity includes the agreements relevant to the clauses for perpetual access and for permanent archival copy. Also in this phase, the technical conditions are evaluated and the economic and legal sections of the license agreement are concluded.

| | PRINT | DIGITAL |
|---------------------|--|--|
| Mission | Maximise the availability of documents for current needs ("just in case") | Focus on information access for current needs |
| Preservation Policy | De facto, not explicitly declared | Should be explicit and declared (related to the licensing policy) |
| Selection | On the assumption that the item will be kept in the library forever | Should include decision on long term preservation (archival right clause) |
| Acquisition | Based on ownership, use (perpetual use and archival right included) supported by copyright law | Based on the license for use of the resource, codified by a contract |
| Format | Stability and relative capacity of self preservation | E-documents are unstable and not self -perpetuating |
| Functions | Preservation not identified as a function in itself | Preservation as specific function |
| Infrastructures | Normal infrastructure (for access and retention) may be sufficient | Ad hoc infrastructures and investments are necessary |
| Collection care | Discontinuity is tolerated | Permanent care needed; discontinuity is not permitted |
| Decision making | Implicit, "by inertia", decisions on preservation may be postponed | Structured, proactive, <i>ex ante</i> approach, procrastination not admitted |

There are many considerations to be made about this scheme, and the ramifications are wide-ranging and profound. Here I would like to highlight two organizational aspects which appear particularly relevant to the current discussion. The first regards the decision-making process: while in the traditional system the decisions inherent in long-term preservation were implicit, nearly imperceptible, and could be spread out through the years, in the digital context preservation requires clear and specific decisions for every resource. Decisions relevant to the long term preservation must be made beginning in the selection and acquisition phases, and cannot be postponed. If we examine the electronic resource management work flow, it is evident that the trial and the negotiation of access and archiving clauses are crucial moments for long term preservation. This establishes an *ex ante* (prior) decision-making process, which requires a more proactive and structured approach than the traditional system. On the other hand, the decision to preserve (or not) a certain resource presupposes the existence of a specific infrastructure and sustainable long term planning. It is well known that digital preservation is a costly operation which cannot be supported by the current budget provisions: since it requires considerable investment and long-term commitments it should be considered as a high-level strategic issue, one which goes far beyond the administrative autonomy normally granted to libraries.

Conclusions

Surprisingly, the general attitude of librarians on preservation issues has not changed significantly in the last 15 years, despite the great change that has occurred in knowledge management and cultural communication systems. On the other hand, the increased awareness of the problem has not reduced the gap between perception and practice. On the organizational level, however, the difference between the traditional

approach to collection development and the emerging model is quite radical, as is the professional culture that they respectively imply.

However, we are not dealing solely with a "cultural" issue; it is also a structural matter of vast dimensions that undermines the business model, which has until now supported libraries. There is in progress a radical shift from an economic model based on the accumulation (and 'capitalization') of the resources acquired to a model based on renting resources for temporary use with no heritage and no guarantees for the future. It is not a *change* – it is a *genetic mutation* in libraries, which is challenging the foundations of modern librarianship. "The library is a growing organism," Ranganathan's 5th law declares; the sustainability of this principle is now an open question.

14.01.2007