

Does “long-term preservation” equate to “accessibility forever”?

Miquel Termens, termens@ub.edu

Mireia Ribera, ribera@ub.edu

Alice Keefer, alice.keeper@ub.edu

University of Barcelona, Library & Information Science Department
Barcelona, Spain

Abstract

This paper proposes a reflection on the points of convergence between the fields of digital preservation and digital accessibility, in terms of both research and development. The two areas have little exchange between them. But, if we look more closely, we find numerous elements –such as objectives, procedures and unresolved problems– that coincide.

Starting with objectives, each area strives to serve users who, at first glance, are quite different: digital preservation is aimed at future users that will use digital platforms that are still unknown, whereas accessibility focuses on current users with disabilities or within disabling contexts. But on closer look, there are parallels between the two groups of users. In both cases there is a considerable lack of understanding about the true needs of users and many unknowns about their technical usage requirements.

As to procedures, the standards for preservation (ISO 14721:2002 – OAIS) and accessibility (CWA 15778:2008 – Document Processing for Accessibility) share obvious similarities. Both propose a model in which there are entry formats, internal formats, and output or dissemination formats. The criteria for format selection in both fields are frequently quite similar.

Finally, there are common, unresolved problems. In the field of preservation a debate has long existed about which “significant properties” need to be preserved, whereas with accessibility, in the absence to date of serious consideration about elements such as emotional aspects, this debate is just beginning.

In conclusion, there is an evident need and rationale for establishing bridges between the two fields in order for them to learn from one another. If they join forces, it is quite possible that common solutions can be found.

Keywords: Digital preservation; Digital accessibility; Long term access.

1. Introduction

The goal of digital preservation is to allow documents produced in the past to be accessed in the future. For this reason “access” –or “accessibility”– of preserved documents is one of the recurring themes in this area. At the same time the term “digital accessibility” has another meaning: it is the combination of techniques that make it possible for digital documents to be used by anyone, regardless of possible disabilities: vision impairment, motor difficulty, deafness, etc. The aim of this paper is to relate the “accessibility” concept of digital preservation with this second meaning and comment on similarities and differences between the two areas.

Digital preservation and accessibility are two distinct areas, but we believe that it is worth noting the important similarities in the problems that each seeks to resolve. As such, we believe that the way in which solutions are sought in one area can, at the very least, shed light on issues under consideration by the other. For example, both areas have addressed how to select and maintain important document features for subsequent access: to future generations, in one case, and to users



with sensory disabilities, in the other. Another relation between the two areas is the uncertainty surrounding the needs of real users, either because they are future users and we do not know what technology they will be using; or because the technology for assisting them currently advances at such a fast pace that their needs adjust continually to the ever-increasing capacity of new systems.

In spite of these points in common and the fact that both areas work with the same elements –digital objects— they do so in separate ways, in terms of the persons, institutions and standards that are devoted to them.

One example that reveals how preservation and accessibility are not marching in unison is that of open repositories of scholarly material, promoted by universities and other research institutions: given the importance of the stored content, preservation aspects are being given attention but, paradoxically, little attention is being paid to the current accessibility of these same documents. [1]

2. Beneficiaries

On a conceptual level both preservation and accessibility share the broad goals of working to serve all types of users, but the reality does not reflect this ideal. For example, the directives for the accessibility of web content recognize explicitly that they do not include users with cognitive disabilities. [2] Also, even if documents are created following the existing standards, their accessibility is not guaranteed. Some producers in targeting a specific audience may decide that a given property is not essential, even though this will cause the product to be inaccessible for other groups for which the eliminated property may be very important. [3]

Even though generic techniques exist that permit specific documents –or parts of documents—to be accessible to all users, in many cases it is absolutely necessary to know the potential audience in order to apply the most relevant solutions. For example, designing a product for the prelingually deaf may result in a sharp reduction of textual language, even though the resulting product is then ill-suited for persons with visual impairments. Similarly, in the field of commercial publishing it has proven impossible to create via a single production line a digital work that responds adequately to all situations, publishing channels, and needs. Thus, the CWA recommendation gives examples of good practices with diverse scenarios, but it makes it evident that each situation will require solutions adapted to its own users. So in the end technical efficiency and economic viability are the parameters that determine the adoption of particular accessibility solutions.

In preservation the vision is similar, but expressed using a different terminology. As the OAIS standard itself states, the purpose of preserving digital information is to “make it available for a *Designated Community*” [4]: the intended future users of the preserved digital objects. And why is this so? For a very simple reason: the awareness of the difficulty –if not to say the impossibility– of fully preserving all original properties of the digital objects. Again, technical constraints and the need for economic viability lead to solutions in which only some *significant properties*, or *essential elements*, of the objects are preserved. And this poses a question –which elements are essential?– that can only be answered from the perspective of a given designated community. [5][6] Even for the experts this is not an easy matter because the choice of significant properties is subjective, making it difficult to arrive easily at agreements.

There are two main streams of thought regarding accessibility [7]: the user-centred design, more inclined to create specific solutions for different communities (the elderly, those with motor disabilities, etc.); and the universal design that promotes the idea of a single design to serve all

publics. Nonetheless, both visions share the belief that documents produced with accessibility in mind will end up being better for everyone. On the other hand, with preservation there is the growing tendency to design systems adapted to a specific community of users, in which the preservation of given significant properties are prioritized over others. As a result in the future we may find documents that are valid for one community, but perhaps totally unintelligible or unusable for others.

These choices –be they related to accessibility’s audience or to preservation’s designated user community— lead to a renunciation of the digital object’s universal applicability and can prove difficult for different sectors’ experts to accept. For example, questions arise such as: Why renounce the subtitles of certain videos? Why not preserve the original typography and colour of a catalogue of artworks?, etc.

3. Problems

Making digital documents and computer applications accessible, as well as preserving all types of digital objects, are stimulating missions, but at the same time difficult to accomplish fully. The basic principles can clash with formidable technical, economic, and management difficulties.

The first difficulty is the broad reach of the missions: at present it is impossible to make all content accessible and to preserve all that needs to be preserved. Priorities must be established and the criteria can vary: the easiest, the most economical, the most scalable, the most heavily used, etc. Therefore, prioritization requires policies to be applied. And the other side of the prioritization coin is the renunciation: of what (for the moment, perhaps) will not be accessible or will not be preserved. The policies of prioritization are painful because implicitly they go against the global aims: some disabled persons will not have access to content that they perhaps will need; others, in the future, will not have access to specific data or testimony from our time.

Traditional accessibility solutions, such as screen magnifiers or screen readers, are built upon the applications and thus are not well integrated into operating systems and other programs. In the long run, the solution will lie in incorporating accessibility into all phases of the development of hardware, software and content, as well as having it present in the workflow of document management. In digital preservation the use of proprietary file formats multiplies the challenges of managing their preservation, as does the plethora of existing formats. A similar reflection can be made regarding the limited support that standardised metadata schema receive from many software applications, not to mention file formats that do not admit metadata.

Legal barriers are also common to both accessibility and preservation. In the analogical world, in many countries the law protects the rights of disabled persons by setting limits to the intellectual property rights in order to allow for the publication of books in Braille. In the digital world, there tend to be fewer exceptions, a situation that leads to increased expense for the rights to publish accessible works. This has opened new fronts for the struggle to broaden rights concerning digital documents. [8] In other cases the problem does not stem from the document itself, but rather from the existence of proprietary reader software that impedes the introduction of elements that that contribute to accessibility.

The problem is similar with preservation. Laws protect the rights of copyright holders by prohibiting the reengineering or decompiling of software, or the modification of content formats, to cite three of the major techniques applied in many preservation scenarios. Certainly, in recent years

there have been numerous initiatives to permit such activities in the context of preservation. But at present, many preservation-related actions currently take place in an environment of questionable legality, at the least. Similarly licenses and usage restrictions –sometimes in the form of Digital Rights Management (DRM)– are also barriers for producing documents that are accessible to all. [9] They also act as barriers for full preservation.

The timing of implementation is also important. The law in many countries protects the publication of accessible versions of textbooks for disabled students, especially the visually impaired. Nevertheless, the procedure for exercising this right can be slow, and may not conclude until after the publication of the commercial work. Therefore, users dependent on the accessible version receive the work much later than others. [10] Similarly, preservation is still seen in many scenarios as an activity taking place at the end of the “normal” life of a digital object and hence is not considered until it is time to “store” the object. At that stage, immediate treatment –for example, migration or other procedures– may be necessary, which might have been spared had the digital objects been created in accordance with preservation requirements.

4. Technological foundation

A comparison of the principal technologies of accessibility and preservation leads us to conclude that there are important similarities both in the procedures and recommendations promoted in each area. The major ones are: the transformation of file formats as a basic technique for facilitating present or future access; the standardization and use of structured and open formats; and the requirement to make full use of metadata.

For accessibility, the use of standards in software and open formats facilitates interoperability and, therefore, the integration of technical aids for reading the documents. The use of structured formats from the XML family facilitates the transformation of documents and, therefore, it too aids in the generation of versions adapted to the needs of different user communities. [11] DAISY is perhaps the format that is currently experiencing the greatest development along these lines, within the publishing sector. [12]

In the field of preservation, many experts have prioritized the reduction of formats used and the appropriate choice of formats for subsequent preservation. The choice of formats is frequently made during the creation, or even during use, of the document and thus remains beyond the scope of preservation actions. However, some programs encourage the use of open, interoperable and standard formats. [13] An appropriate characterization of files and the proper structuring of contents in them can also contribute towards subsequent preservation tasks, such as migration.

Whether from the vantage point of accessibility or of preservation, the volume of digital production is so great that it is virtually impossible for all files to be handled appropriately after the fact, e.g., after their creation. This leads to the recommendation that files should be standardised at their point of origin, as a means of reducing variability. It should not surprise us, then, that accessibility experts are promoting the adoption among publishers of standards and common formats, for both webs and textbooks. [14] This would enable the publishing chain to generate specific products with varying presentations and formats geared to the needs of each user. In preservation, there are many more sources of content generation, since digital objects created within the publishing world account for only a small fraction of the total number of items to be preserved. Some current attempts for influencing how digital objects are created are centred within the public administration and some scientific fields and it remains to be seen if and how it will spread to other areas in the future.

5. Conclusion

Accessibility and preservation serve different objectives even though they act on the same types of materials. In this work we have seen some of their similarities: in strategies, in approaches to challenges, and in technological underpinnings. We have also seen that some proposed solutions are stymied by pre-existing legal conditions. Likewise we have shown how practical concerns –such as technical expediency and economic viability—can lead to actions that are frequently more limited than the respective movements’ overriding aims.

These common elements lead us to believe that a greater degree of understanding between the two communities would be beneficial to both sides. Surely each could learn something from the other and, in so doing, shed more light on its own approaches. Also, collaboration would be beneficial in order to reach common objectives, such as the promotion of open and structured formats.

Finally, it is worth remembering that the two communities maintain close relations with certain stakeholders: universities, public administration and libraries. Equally important to both is the expansion of e-government as the main transforming engine for practices related to the creation and management of digital content. This common ground could facilitate points of encounter and contribute to working together towards common solutions.

6. References

1. Kelly, Brian (2006): “Accessibility and Institutional Repositories”. *UK Web Focus*. <http://ukwebfocus.wordpress.com/2006/12/12/accessibility-and-institutional-repositories/>
2. *Web Content Accessibility Guidelines (WCAG) 2.0. W3C Recommendation 11 December 2008* (2008). <http://www.w3.org/TR/WCAG20/>
3. CWA (2008): *CEN/ISSS CWA 15778:2008 – Document processing for accessibility*. <ftp://ftp.cenorm.be/PUBLIC/CWAs/DPA/CWA15778-2008-Feb.pdf>
4. Consultative Committee for Space Data Systems (2002): *Reference model for an open archival information system (OAIS). Blue Book*. <http://public.ccsds.org/publications/archive/650x0b1.pdf>
5. Cedars Project (2001): *The Cedars Project Report. April 1998 – March 2001*. <http://www.leeds.ac.uk/cedars/pubconf/papers/projectReports/CedarsProjectReportToMar01.pdf>
6. Hedstrom, Margaret; Lee, Christopher A. (2002): “Significant properties of digital objects: definitions, applications, implications”. *Proceedings of the DLM-Forum 2002*. Luxembourg, Office for Official Publications of the European Communities. p. 218-223.
7. Seale, Jane K. (2006): *E-Learning and Disability in Higher Education: Accessibility Research and Practice*. Oxford, Routledge.
8. Commission of the European Communities (2008): *Green paper: copyright in the knowledge economy*. Brussels, Commission of the European Communities. http://ec.europa.eu/internal_market/copyright/docs/copyright-infso/greenpaper_en.pdf
9. Kramer, Elsa F. (2007): “Digital Rights Management: Pitfalls and Possibilities for People with Disabilities”. *Journal of Electronic Publishing*, 10 (1). <http://hdl.handle.net/2027/spo.3336451.0010.106>

10. Keil, Sue; Parris, Delith, Cobb, Rory; Edwards, Angela, & McAllister, Richard (2006). *Too little, too late - provision of school textbooks for blind and partially sighted pupils*. London, Royal National Institute of the Blind.
11. Paepen, Bert; Engelen, Jan (2002): “Using XML as a Reading Enabler for Visually Impaired Persons”. *8th International Conference, ICCHP 2002. Lecture Notes in Computer Science*, 2398, p. 382-389.
12. Kahlisch, Thomas (2008): “DAISY: an opportunity to improve access to information for all”. *Information Services and Use*, 28 (2), p. 151-158.
13. Arms, Caroline; Fleischhauer, Carl (2005): “Digital formats: Factors for sustainability, functionality and quality”. *Proceedings Society for Imaging Science and Technology (IS&T) Archiving 2005*. Washington DC. p. 222-227.
14. Lockyer, Suzanne; Creaser, Claire; Davies, J. Eric (2005). “Availability of accessible publications: designing a methodology to provide reliable estimates for the Right to Read Alliance”. *Health Information and Libraries Journal*, 22 (4), p. 243-252.