

Dr. Stefan Rohde-Enslin
Institut für Museumskunde
nestor - Kompetenznetzwerk Langzeitarchivierung
(D) - 030 - 8301494
s.rohde-enslin@smb.spk-berlin.de

Dust without Wind? - Long-term Preservation of Digital Data: Media, Formats, anything else?

Summary

Every museum creating or using digital data will sooner or later have to realize that it is not easy to keep such data for the future. The long-term preservation of museum data is a multifaceted problem. A description of the complexity of the problem will be given. Only with an understanding of this it might be possible to find suitable solutions and to make the necessary decisions. Long-term preservation of museum data is more than just storing data, it has to be considered already when the data are created - it is a permanent task. It is not easy to keep something for a long time that cannot be seen, touched or heard. Data are even less than dust in the wind - but that is only one part of the problem.

[A word ahead]

It might hold true for human beings that "All go unto one place; all are of the dust, and all turn to dust again." (Ecclesiastes 3, 20). The "breath of live" makes the difference. Without this special kind of "wind" there might be nothing but dust.

When it comes to digital objects a similar conception might be applied: As long as some care, some changing, some transferring – in a word: some "wind" – is there, they are existing. We shall look for it that there is a strong enough wind blowing in the right direction – towards the future.

[Intro]

This paper deals with long-term preservation of museum data. It is first of all an appeal to give that topic the consideration it demands and to actively and collectively combat the decay. It is not intended here to present the latest results of ongoing research into long-term preservation of data. Instead of that some few spotlights shall be directed on the topic to illustrate its complexity.

The guiding question in the first part of the paper is that about responsibilities at different stages of data-production and keeping. In the second part some initiatives relevant for museums are named to show possible partners from whose experiences museums may benefit and with whom museums may

cooperate. The paper finally deals with considerations about first collective steps to tackle this issue.

[Part I]

This first part shows the complexity of the issue of long-term preservation of digital data in a museum setting. A fully comprehensive overview is not necessary to demonstrate this. Some spotlights thrown on different stages in the life-cycle of museums data might suffice.

[Spot 1]

Even before someone in a museum begins producing data far-reaching decisions took place. Hard- and Software had to be selected as also operators with some technical skills. The requirements the produced data have to fulfil should have been defined. Things like file-format, colour-depth have been chosen, metadata standards might have been considered.

These technical aspects are easily to be seen and they are often discussed. But when one turns to decisions, to the question of who decides on what, one sees more in it.

For example: The decision which scanner to buy necessarily has to involve those responsible for the museums finances and the decision which operators to employ has to involve those responsible for personnel.

The importance of such decisions cannot be overestimated and it is really at this point, where the consideration of long-term preservation has to start because some formats are better suited than others and a too low scan-resolution might produce a file whose worth for being kept shall be doubted.

[Spot 2]

Sometimes some decisions are taken by people outside the museum as for example with digital art. A museum acquiring such pieces of art gets a fixed set of data (and possibly some other things). Often those responsible for the acquisition don't consider the issue of long-term preservation of that data. Those responsible for data-preservation should be involved in the decision to acquire – not all data can be kept.

Because of the need to transfer data to different storage media and to bring them into up-to-date file formats from time to time difficulties may occur if the artist or any other creator of a data-set who holds respective rights denies any changes of the original data. At the time of the acquisition the appropriate agreements have to be made – the creators have to be involved. It might be good to have a common understanding about minimum requirements agreed upon widely by museums so that the artists (creators) will offer there data already in formats which make preservation an easier task.

[Spot 3]

It is much easier now than it was before to duplicate and transfer information. The electronic form tends to increase the amount of information available.

Databases, websites, emails, electronic guides, museums administration data ... different information, different data ... increasingly interwoven among another. It seems impossible to preserve all these data and at the same time all the inherent links. Selections have to be made.

The necessary decisions have consequences. The complexity of the data chosen to preserve decides about the skills and the amount of time necessary to keep them. The amount of chosen data correlates with the necessity to finance more or less sophisticated, extensive technical equipment.

In the formulation of rules for selection many have to be involved. The question tackles not only technical or financial questions but also those of content. The respective curators have to be involved.

[Spot 4]

The last of this few spotlights is directed on data handpicked for preservation by a museum. A place for the intended long-term storage has to be found. Shall it be a storage system inside the museum run by its own employees or shall an institution from the outside be contracted? Shall it be a concerted effort of partners or a clearly commercial relationship? What shall be the criteria for such decisions?

In any case this is again not only a technical question, it is also one of long-term financial plans and it might be one of a long-lasting relation between the museum and an outside institution. The management of the museum has to take such far reaching decisions.

[Synopsis of Spotlights]

There are many more points in the life-cycle of digital data in museums worth to direct spotlights on. Here only four were chosen to demonstrate that the question of long-term preservation is not only a question about file-formats, storage-media etc. and that many different people are necessarily among those involved in decision making.

While those responsible for information and/or documentation often know about the issue, others, like curators or those responsible for finances, are often unaware of their responsibility and the museum management might not yet imagine which decisions it will have to take. In small museums, where there is no designated documentation/information officer there remains no one fully aware of the task.

[Part II]

In regard to long-term preservation of digital data the situation in museums is only slightly different from those of others in the cultural heritage sector. This second part of the paper names some initiatives already on their way to address the issue. They offer points for information and cooperation, listed in an arbitrary order. Certainly there are much more national or international initiatives not listed here. The ones named shall only be taken as examples.

PADI

Most known in the digital-preservation community is PADI, the “Preservation Access to Digital Information”. It is basically an information-gateway run by the National Library of Australia (<http://www.nla.gov.au/padi/>). Most of the relevant documents can be found here.

DPC

Closely working together with PADI is the DPC, Digital Preservation Coalition, in the United Kingdom (<http://www.dpconline.org/graphics/index.html>). The coalition which has many member-organisations from different sectors, among them museums, published a very helpful “Handbook” in print and in a web version. (<http://www.dpconline.org/graphics/handbook/>)

nestor

The German project “nestor” is also cooperating closely with PADI. Under the guidance of the main German library, organisations from the cultural heritage sector are collectively working for rising awareness and promoting relevant discussions. (<http://www.langzeitarchivierung.de>). “nestor” published situation reports on different aspects of digital preservation.

Conservation Online

A good address for finding documents is “Conservation Online – Resources for Conservation Professionals” in the United States of America (<http://palimpsest.stanford.edu/>). The site covers more than the preservation of digital data. Many useful documents may be found here.

North Carolina - Echo

On the website of “North Carolina – Exploring Cultural Heritage Online” one finds a very practical guideline for digital preservation which corresponds to the needs of smaller institutions.

(<http://www.ncecho.org/documents/DGCh6Preservation.pdf>)

IASA

The “Technical Committee” of the International Association of Sound and Audiovisual Archives” (IASA) prepared a paper on “The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation, Standards, Recommended Practices and Strategies”. This is a very helpful compilation of minimal technical requirements for audio files, which might be useful not only in the setting of audio archives, but also in museums.

(<http://www.iasa-web.org/iasa0013.htm>)

UNESCO

In March 2003 UNESCO published its “Guidelines for the Preservation of Digital Heritage” (Prepared by the National Library of Australia). Many aspects of the issue of long-term preservation of digital data are tackled here. These comprehensive “Guidelines” try to reach all of those who have a hand in the

future of digital data of cultural organisations
(<http://unesdoc.unesco.org/images/0013/001300/130071e.pdf>).
The “Guidelines” are available in English, French and Spanish.

As mentioned in the introduction to this short list it is made up of more or less arbitrarily chosen organisations. It is presented here to visualize that there is already an ongoing international discussion and a growing awareness in many countries of the world. These are activities of which we as museums may profit from and contribute to.

It often is the library sector where the discussion of long-term preservation of digital data started. The need to look for ways to preserve digital data becomes urgent the more the whole process of digitisation reaches the museum world.

[Part III]

This third part consists of an appeal for concrete action. To recapitulate: The introduction stated: We have to do something! Part one focussed responsibilities and the fact that many more than only documentation or information officers have to be involved in the necessary decision making. The second part illustrated that there are already initiatives in the cultural heritage sector looking into the issue.

We shall take the relevant experiences already collected in the cultural heritage sector and combine them with our own and others. With this we shall prepare a checklist in such a way that it is intelligible to and helpful for everyone who will have to bear a part of the responsibilities in making the right decisions for the preservation of museum data be it in a small, medium or big institution.

An embryonic and preliminary version of such a checklist is currently in preparation. Please feel free to participate in CIDOCs Digital Preservation Group.

It's not enough to know how the wind blows; we have to make wind ourselves.