How to Develop a Preservation Policy

Guidelines From the nestor Working Group by Yvonne Friese[1]

ZBW

Abstract

This paper gives insights into the findings of the nestor working group "preservation policy" which was founded in the beginning of 2012. It is led by two of the nestor partners: the German National Library and one of the Goportis libraries, the ZBW – Leibniz Information Centre for Economics.

The working group attempts to help institutions involved in digital preservation to develop their own preservation policies. To support this task, the group has created guidelines for the development of an institutional preservation policy which will be published during the first quarter of 2014. Shedding light on the policy development process and providing guidance concerning the content and structure of a preservation policy, the guidelines describe what a policy is needed for, which content it could have, which staff members should be involved in the development and how its quality can be ensured.

Keywords: preservation policy, digital preservation, guidelines, nestor

Introduction

According to the ISO standard "Audit and certification of trustworthy digital repositories", a preservation policy is a "[w]ritten statement, authorized by the repository management that describes the approach to be taken by the repository for the preservation of objects accessioned into the repository" (ISO 16363, 2012). The standard explains that the policy has to be consistent with the preservation strategic plan. In contrast to the policy, the preservation strategy addresses how the preservation is carried out and therefore focuses on workflows and technical strategies. In practice, the policy and strategy are often (but not necessarily) addressed in the same document which complicates delimiting between the two.

The ISO standard requires that the preservation strategy match the preservation policy and vice versa. For example, an institution cannot state in the policy that 100% of digital content is preserved if the strategy makes it possible to consider only parts of the entire collection for preservation due to technical obstacles.

Preservation policies are an essential tool in digital preservation, serving both the purpose of creating trust and offering a formally binding frame of reference for the preservation activities of a given institution. However, although many institutions in Germany and all over Europe have already begun to engage in digital preservation, only a few have published a preservation policy of some kind (Angevaare, 2011, p. 5). Thus, as part of the 2011 DigCurV survey of training needs, 454 institutions were asked if they engaged in storing digital material. The institutions surveyed were cultural heritage institutions such as libraries, archives, or museums from 44 countries. Respondents mostly came from European countries (81.3%), but also from the United States (12.3%), Canada (1.5%) and a small percentage (4.7%) from other countries (Engelhardt, Strathmann and McCadden, 2011, p. 10).

More than 75% (n = 437) replied that they were involved in digital curation. An additional 18% (n = 437) stated that they were planning to store digital materials in the future (ibid., p. 15-17).Thus, according to the survey, 331 institutions were already engaged in digital curation in 2011 and it is likely that this number has grown over the last two years.

But what is the state of preservation policies? Two resources serve to support Angevaare's claim:

1. With the aim of surveying "the current state of digital preservation policy planning within cultural heritage organizations" Sheldon (2013) collected and compared publicly available policies worldwide and counted 33 documents: 15 from libraries, 16 from archives, and two from museums. It should be noted, however, that Sheldon limited her analysis to published policies which are written in English (see 2013, p. 4). Therefore her study excludes, for example, the BSB (Bavarian State Library, 2012) and the DNB (German National Library, 2013) preservation policies, as neither has an English translation yet.

2. The SCAPE wiki on published preservation policies (last updated in November 2013) lists 40 institutions with a published policy. The list is not limited to English-language material and includes Dutch, German, and Danish policies. As the wiki is built collaboratively and receives updates from many authors from different countries, it seems safe to assume that it is fairly comprehensive even though it surely is not complete.

There is an overlap of 26 published preservation policies found by Sheldon and listed by the SCAPE authors. Sheldon includes four policies not listed in the wiki, and there are 14 policies listed in the SCAPE wiki not taken into account by Sheldon. Hence, 44 institutions with published digital preservation policies are known. Although the different scope and design of the surveys used here is not entirely identical, the numbers support Angevaare's perception: The number of institutions actively archiving digital material (at least 331) greatly exceeds the number of institutions with published preservation policies (at least 44).

The nestor working group

Although a preservation policy is such an important part of an organization's commitment to digital preservation, a certain reluctance to develop and adopt one is understandable. Firstly, a transparent policy which can be accessed by users, partners and investors is a big commitment. Secondly, it can be quite difficult to determine the level of detail and decide on length and scope of a preservation policy. To support the widespread development and adoption of digital preservation policies, in 2012 a nestor working group was formed to establish guidelines for the creation of a preservation policy for memory institutions such as archives, museums or libraries. Its 12 members come from Germany and Switzerland.

The working group is part of nestor (network of Expertise in long-term storage and availability of digital resources), the German-language competence network for digital preservation founded in 2003. Initially funded by the BMBF (Federal Ministry of Education and Research) in two phases (2003-2006 and 2006-2009), since 2009 nestor is acting as an independent, self-financing network. As of today, it has 16 members, mostly German memory institutions such as libraries, archives and museums (see figure 1). There are more institutions interested in becoming a nestor partner, so the number of partners is likely to grow further.

Currently, nestor consists of eight working groups for different important digital preservation tasks and topics, e.g. AV-Media, Cost, Rights and Emulation. The network is also engaged in standardization work and

has developed three national standards over the last three years, among others the catalogue of criteria for trustworthy digital archives (DIN 31644). Since 2013, German digital archives have the possibility to receive the nestor seal for trustworthy digital archives which is based on these criteria (nestor, 2013a). In addition, creating guidelines and making international standards accessible to the German-speaking community belongs to the tasks of nestor and its working groups. E.g. a translation of the OAIS model into German was published in 2012 (nestor, 2013b). nestor also monitors the state of digital preservation and curation in Germany and the German-speaking countries, and just recently a Baseline Study of the Digital Curation of Research Data in Germany (also available in English, Neuroth et al., 2013) was carried out.

Starting with a review of already existing preservation policies (The National Archives, 2009; NLNZ, 2011), the working group noted that policies vary considerably in length, depth, and detail. For example, the preservation policy of the National Library of New Zealand (2009) and Archives New Zealand (ANZ) includes parts that - due to fast technical changes – would have to be updated quite often and in our opinion should rather be included in the preservation strategy.

As no German guidelines on this topic exist, the working group reviewed existing English guidelines on policy development (The National Archives, 2011). These were used to decide which parts are important for the German community as well. The resulting German-language guidelines, which will be published in the first quarter of 2014, consist of five main chapters:

1. Goals of our guidelines
2. Use of a policy
3. Development of a policy (motive, responsibility, publication, relation to other related documents and strategic papers)
4. Possible content of a policy
5. Updates of a policy (policy watch)

In the following, an overview of the most important findings of the preservation policy working group is given. It is these findings which form the basis for the content of the guidelines.
.

Goals of our guidelines

As already emphasized, a preservation policy is an important element in securing long-term-access to digital objects. Digital preservation depends on technical as well as organizational issues, and a policy serves to address these. It demonstrates the

Figure 1 nestor partners

commitment and responsibility of the archiving institution and adds to its trustworthiness.

It is to be expected that it will be common practice in a few years for all Digital Archives to have a published preservation policy, and accordingly the pressure for every institution to get involved in this topic grows. Against this background, the work of the nestor group aims to simplify the task of writing a preservation policy and to raise the awareness of the need of a publicly visible policy, especially for a German audience. Our guidelines provide a tool box: the institutions using it decide themselves which parts will be relevant for their institutional policy and on this basis create a policy suitable for their needs. Thus the guidelines aim to assist in the development of a policy, but they will not dictate any mandatory rules as the needs of the different institutions and digital archives are very heterogeneous. Accordingly, the guidelines inform users about the impact, use, typical questions and difficulties of (creating) a policy. They help them to unmask their blind spots and increase awareness of dependencies and consequences. In addition, a generic policy example – abstracted from already existing policies – gives an idea to the users of the guidelines of what a policy might look like.

Purpose of a policy

In general, the purpose of a policy is to show why – and, possibly, how - an institution is involved in digital preservation and to define its benefits (The National Archives, 2011). It demonstrates that an institution is part of the preservation community and is aware of important standards. More specifically, the purpose of a policy derives from its audience, or, simply: its users. The latter can be internal or external users.

For internal users such as staff members, a policy forms an important basis for decisions. It can also serve to mitigate possible financial cuts: affected staff can point to the policy and insist at least on the budget needed for the minimum standards the institution has publicly committed to maintain.

For external users, for example the "consumers" of the digital assets, a policy supports the building of trust as it creates security that the assets will remain accessible, citable, and usable for the long term. The same is true for the data producers, who have an interest that their findings serve future users as well. Having a published preservation policy means that stakeholders will have transparent information about what the archive does to secure long-term access, as will potential clients who are considering outsourcing the digital preservation of their assets. Thus, the institutional preservation policy is likely to be the basis for service level agreements between the archiving institution and any third party (Beagrie et al., 2008).

Finally, a policy can be also useful – or even mandatory – for certification and audit processes and certainly will help acquiring third-party funds.

Developing and publishing a policy: The why, how, who and where

There are multiple motives for starting to develop a policy. An obvious reason would be the beginning of digital preservation activities, but as the findings cited above show, this is rarely the case. In fact, one reason not to adopt a preservation policy early on in the process of building a digital preservation system is that this policy is likely to be revised frequently as the system and the experience grow and the workflows are implemented.

For example, the Marriot Library of the University of Utah in Salt Lake City, USA, revised its policy three times during the last three years. In contrast to the many institutions which conduct a digital archive without having developed a preservation policy yet, the Marriot Library published the first version of its preservation policy in 2010 – two years before they purchased the digital preservation system they are using today. They deliberately developed a policy so early because they felt it would help them to decide which preservation software to purchase once there would be something suitable available for them. In the case of the Marriot Library, writing the policy helped to shape the preservation program and to raise awareness about digital preservation plans and actions among the staff members. It is likely that there will be another revision once the digital archive is well established and fully implemented in the library workflows (Keller, 2012).

In contrast, the German National Library (2013) and the Bavarian State Library (2012) had already been engaged in Digital Preservation for a number of years before they published their policy. In these cases, it was preferred to set up the policy after the Digital Archive was established and the full extent of the system was known.

Furthermore, technical or organizational changes within the institution could be the reason to start the development of a policy: an external evaluation of the institution, or – as mentioned before – an audit or a certification of the Digital Archive. Depending on the organizational structure of the institution, a number of different staff members can be responsible for developing the policy content. Possible scenarios are described in the nestor guidelines. In most cases, both members of the management and practitioners are likely to be involved. The development process and later adjustments of the policy will be time-consuming, especially if many staff members need to be involved. If possible, it is therefore recommended to keep the number of involved persons to the necessary minimum.

Where and how the policy is published is partly dependent on its scope. A policy might contain confidential matters and therefore will only be published within the respective institution. This might concern the whole policy or just certain chapters. The language used in the policy strongly depends on the target group. Usually the national language is used and often an additional English translation for an international audience is created. Generally, the language used has to be comprehensible for a wider audience and should avoid technical terms.

Additionally, the policy will most likely refer to other documents or strategic papers. It is recommended, for example, to address technical solutions not in the policy text but in other, related documents, as this content is likely to change very fast. As for the description of ingest workflows and preservation strategies like migration, these are better explained in the preservation strategy plan instead of in the policy (The National Archives, 2011, p. 7). It is also highly important to ensure that the policy does not conflict with laws, rules or tasks of the institutions or already existing policies, for example the preservation policy for printed material.

Due to the relative novelty of the field, digital archives are often still in a development phase or in a very early stage of productive use.

Therefore, the status quo of a given archive is often still not stable enough to frame certain principles. As mentioned above, this could be one of the reasons why many institutions seem hesitant to publish a (final) policy. In these cases it is possible to express the status quo of an archive or to create an "aspirational policy" (The National Archives, 2011, p. 6), but both possibilities bear a risk of having to revise the policy fairly soon.

Policy content: The what

The areas covered in a policy can vary a lot. Analyzing 33 policies in the English language, Sheldon (2013, p.6) observes that some are only one page long, whereas others consist of 30 pages or more. From the point of view of the nestor working group it is therefore an important task of our guidelines to give an overview of possible content of a policy and to emphasize the consequences that adding a particular content item will have for future work and the need to update the policy regularly. Again, the guidelines refrain from prescribing too much because each institution will have very individual needs and there will be no "one size fits all" solution.

The policy content is the main chapter of our guidelines as the possibilities are diverse and multifaceted. Therefore, only a selection of possible aspects can be highlighted in this paper. In creating its guidelines for policy content, the working group took into account Beagrie's model of a preservation policy (2008; see table 1), the findings of Sheldon's analysis (see table 2), the practical experience of the members of our working group, and our own analysis of existing policies we consider to be a good example.

Table 1: Policy content suggested by Beagrie (2008)	
1	Principle statement (benefits)
2	Contextual links (relation to other strategies and documents)
3	Preservation objectives
4	Identification of content (scope of digital content)
5	Procedural accountability (responsibilities)
6	Guidance and implementation
7	Glossary
8	Version control (review of the policy)

The working group decided not to include all these criteria in its guidelines because from our point of view a compact policy with a manageable number of topics is easier to develop and to maintain. Thus, some of the topics identified by Sheldon (e.g. Preservation Planning, Storage, Duplication, and Backup) might better be placed in a preservation strategy, which addresses more technical topics like preservation planning, storage, duplication and backup, and which will have to be revised more often.

It is evident that the objective and the scope of the policy should be embedded in the general strategy of the institution and has to be compliant with its focus, priorities and tasks. It is important to define this objective in time and to address it within the policy.

Table 2: Common policy content identified by Sheldon (2013)	
1	Access and Use
2	Accessioning and Ingest
3	Audit
4	Bibliography
5	Collaboration
6	Content Scope
7	Glossary/Terminology
8	Mandates
9	Metadata or Documentation
10	Policy/Strategy Review
11	Rights and Restriction Management
12	Preservation Planning
13	Rights and Restriction Management
14	Roles and Responsibilities
15	Security Management
16	Selection/Appraisal
17	Staff Training/Education
18	Storage, Duplication, and Backup
19	Sustainability Planning

The goals of preservation, e.g. maintaining the usability, authenticity and integrity of the archived digital objects, can be a main part of the policy, as this is the heart of all preservation activities and of particular interest for the target group. A policy can also address how these preservation goals will be reached. Furthermore, it is recommended to name the responsible units within the institution, those responsible for the archiving workflows, and the staff member or members responsible for the content and the updates of the policy itself. As there will most likely be some fluctuation in the staff, it is recommended to only point out staff functions rather than including names.

From the perspective of the working group, other important topics for a policy are:
• The organizational structure of the institution (including secure funding for the future)
• Mandate of the institution
• Legal and technical framework
• Principles of digital curation, e.g. maintaining integrity, authenticity and accessibility
• Protecting sensitive data from unauthorized access (e.g. medical research data).

Some points are not mandatory but could be useful depending on the scope of the policy:
• Purpose and scope of the archived digital material

- Staff and other resources used for digital preservation (as this can also change over time, a rough estimate might be enough)
- Preservation activities (information in detail might lead to regular updates).

The process and criteria for the selection of digital objects for the archive could also be part of the policy. Furthermore, the access to different collections – if the institution is a light archive with user assess – can be an important part of the policy as well. Again, however, as digital collections are growing, the policy would possibly have to be extended quite often. Thus, if an institution does not want to update the policy regularly, it might be a good decision to deal with this issue in another, related document. The scope of the archive collection could be described in such a document and the description of newly acquired material could then be added to this document to avoid that the policy has to be edited too often.

Policy watch: updating and evaluation

Among the members of the working group opinions about whether or not a policy should be changed, and how this should happen, diverge. On the one hand, by revising its policy regularly, an institution can show that it actively watches technology and developments in digital preservation and keeps the preservation policy up to date. On the other hand, a preservation policy should be a commitment for the long term, something the institutional staff, stakeholders and clients can build and rely on. Accordingly, if the decision to change the policy is made, it is a matter of trust to make the reasons for updates and changes transparent and to archive the older versions and keep them accessible, for example on the institution's website. It is possible to indicate the next review date within the policy, as the National Archives have done (The National Archives, 2009, p. 10). Of course, such a review might reveal that there is no need to change the policy.

A policy update becomes necessary, if the policy no longer matches the daily work. For example, if an institution had a dark archive and adds an access component, the policy is likely to lack guiding principles for this. It will be necessary to extend the policy in order to cover access to the archived collections, and this will have to happen in a transparent and comprehensible way. If the policy includes detailed technical aspects, there will likely be a need to adjust it quite often to account for technical developments and changes of workflows. One possibility to deal with this issue is to state it in the policy and thus announce it from the beginning.

An evaluation of the preservation policy can include the question whether the policy has met its goals. For example, it might be necessary to adjust existing workflows to the policy in this context. This is the best case scenario. An evaluation might also reveal that the reality cannot be adjusted to the policy and the policy has to be changed because certain procedures cannot be implemented. Due to the lack of experience in this still relatively new field this possibility cannot be ruled out. Again, it is recommended to create transparency in this case by giving comprehensive explanations about the changes.

As the example of the Marriot Library mentioned above shows, a policy can be revised because the first draft of the policy has been published at a very early stage and therefore has to be updated more often and extensively as the implementation of the actual workflows take place.

A look into the future

The nestor preservation policy working group aims to publish its guidelines in the first quarter of 2014. Subsequently, there will be a workshop for practitioners and possibly other follow-up activities. The guidelines will be available as an open access resource (in German; an English translation is currently not planned).

The topic of preservation policies will also feature in nestor's upcoming best practice wiki[2]. The wiki will supplement the publication of the guidelines and will provide a competence network for the discussion of practical questions and issues. It will also serve as a platform to address yet unresolved or even unknown aspects of drafting and maintaining preservation policies. For example, the majority of the institutional policies Sheldon (2013) examined, address the issue of collaboration. Currently, this is not part of our guidelines, although some of us participate in a digital preservation consortium. Apparently, there are still blind spots to be detected by us and by others!

References

Angevaare, I., 2011. Policies for Digital Preservation. [pdf] Available at: <http://www.digitalpreservationsummit.de/presentations/angevaare.pdf> [Accessed 10 December 2013]

Bavarian State Library, 2012. Sicherung des in digitaler Form vorliegenden Wissens für die Zukunft – Die Langzeitarchivierungsstrategie der Bayerischen Staatsbibliothek. [pdf] Available at: <http://www.babs-muenchen.de/content/dokumente/2012-11-22_BSB_Preservation_Policy.pdf> [Accessed 10 December 2013]

Beagrie, N. et al., 2008. Digital Preservation Policies Study, Part 1: Final Report October 2008. [pdf] Available at: <http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1final-report.pdf> [Accessed 10 December 2013]

Engelhardt, C., Strathmann, S. and McCadden, K., 2011. Report and analysis of the survey of Training Needs. [pdf] Available at: <www.digcur-education.org/eng/content/download/3322/45927/file/Report%20and%20analysis%20of%20the%20survey%20of%20Training%20Needs.pdf> [Accessed 10 December 2013]

German National Library, 2013. Langzeitarchivierungs-Policy der Deutschen Nationalbibliothek. [pdf] Available at: <http://d-nb.info/103157140X/34> [Accessed 10 December 2013]

Keller, T., 2012. Digital Preservation Policy. [online] Available at: <http://www.lib.utah.edu/collections/digital/digital-preservation.php> [Accessed 10 December 2013]

International Organization for Standardization, 2012. Space data and information transfer systems – Audit and certification of trustworthy digital repositories. ISO 16363:2012. Washington D. C., USA

nestor, 2013a. nestor seal for trustworthy digital archives. [online] Available at: <http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html> [Accessed 10 December 2013]

nestor, 2013b. Referenzmodell für ein Offenes Archiv-Informations-System. nestor-Materialien 16. [pdf] Accessible at: <http://files.d-nb.de/nestor/materialien/nestor_mat_16-2.pdf> [Accessed 10 December 2013]

Neuroth, H., et al., 2013. Digital Curation of Research Data. [pdf] Available at: <http://nestor.sub.uni-goettingen.de/bestandsaufnahme/Digital_Curation.pdf> [Accessed 10 December 2013]

NLNZ, 2011. Digital Preservation Strategy. [pdf] Available at: <http://archives.govt.nz/sites/default/files/Digital_Preservation_Strategy.pdf> [Accessed 10 December 2013]

The National Archives, 2009. Preservation policy. [pdf] Available at: <http://www.nationalarchives.gov.uk/documents/tna-corporate-preservation-policy-2009-website-version.pdf> [Accessed 10 December 2013]

The National Archives, 2011. Digital Preservation Policies: Guidance for Archives. [pdf] Available at: <http://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf> [Accessed10 December 2013]

SCAPE Wiki, 2013. Published Preservation Policies. [online] Available at: <http://wiki.opf-labs.org/display/SP/Published+Preservation+Policies> [Accessed10 December 2013]

Sheldon, M., 2013. Analysis of current digital Preservation Policies. Archives, Libraries and Museums. [pdf] Available at: <http://www.digitalpreservation.gov/documents/Analysis%20of%20Current%20Digital%20Preservation%20Policies.pdf?loclr=blogsig> [Accessed10 December 2013]

NOTES

1. Yvonne Friese is affiliated at the Leibniz Information Centre for Economics in Kiel. The main focus of her work is on digital preservation of research paper and digitized material and organizational issues around digital preservation. Her contact email is y.friese@zbw.eu.

2. The wiki has already been established as a test and will be made available for public access in the future.