Time To Change
TIB

Digital Curation
Training
SUB GOETTINGEN / FH KOELN

Preservation
Policy
ZBW

Tried and Trusted
GESIS

IN THIS ISSUE
4 Papers: Special Issue

ON PAGE 5
Editor Karsten Boye Rasmussen

ON PAGE 31
Membership Information
IASSIST Quarterly

The IASSIST Quarterly represents an international cooperative effort on the part of individuals managing, operating, or using machine-readable data archives, data libraries, and data services. The QUARTERLY reports on activities related to the production, acquisition, preservation, processing, distribution, and use of machine-readable data carried out by its members and others in the international social science community. Your contributions and suggestions for topics of interest are welcomed. The views set forth by authors of articles contained in this publication are not necessarily those of IASSIST.

Information for Authors

The Quarterly is normally published four times per year. Authors are encouraged to submit papers as word processing files (for further information see: http://www.iassistdata.org/iq/instructions-authors). Manuscripts should be sent to Editor: Karsten Boye Rasmussen. Email: kbr@sam.sdu.dk

Announcements of conferences, training sessions, or the like are welcomed and should include a mailing address and a telephone number for the director of the event or for the organization sponsoring the event.

Editor

Karsten Boye Rasmussen
Department of Marketing & Management
University of Southern Denmark, SDU
Campusvej 55, DK-5230
Odense M, Denmark
Phone: +45 6550 2115
Email: kbr@sam.sdu.dk

Deputy editor

Walter Piovesan,
Simon Fraser University

Website editor

Stuart MacDonald,
University of Edinburg

In this issue

5 Editor’s notes
Karsten Boye Rasmussen

6 Guest Editors’ Notes
Astrid Recker, Natascha Schumann

8 Time to Change: Effects and Implications of Digital Preservation in an Organizational Context
by Michelle Lindlar

14 Digital Curation Training: The nestor Activities
by Stefan Strathmann and Achim Oßwald

18 How to Develop a Preservation Policy: Guidelines From the nestor Working Group
by Yvonne Friese

24 Tried and Trusted: Experiences with Certification Processes at the GESIS Data Archive
by Natascha Schumann
Editor’s notes

Special issue: The organizational dimension of digital preservation

Welcome to the special double issue 3 & 4 of the IASSIST Quarterly (IQ) volume 36 (2012). This special issue addresses the organizational dimension of digital preservation as it was presented and discussed at the IASSIST conference in May 2013 in Cologne, Germany.

The two guest editors Astrid Recker and Natascha Schumann from the GESIS Leibniz Institute for the Social Sciences in Cologne have earned special thanks. If you find their names familiar it is because they co-wrote a paper in the IQ 36-2. They are concerned with data preservation and curation at the Data Archive for the Social Sciences, and as a member of the Archive and Data Management Training Center, Astrid also trains others in these areas. Furthermore, they co-chaired the panel on ‘Beyond Bits and Bytes: the Organizational Dimension of Digital Preservation’ at IASSIST 2013, both also participating as panelists in the session. They have now persuaded the other panelists to contribute to this combined special issue. Thanks also therefore to Michelle Lindlar, Stefan Strathmann and Achim Oßwald, and Yvonne Friese.

Articles for the IASSIST Quarterly are always very welcome. They can be papers from IASSIST conferences or other conferences and workshops, from local presentations or papers especially written for the IQ. Authors are permitted “deep links” where you link directly to your paper published in the IQ. Chairing a conference session with the purpose of aggregating and integrating papers for a special issue IQ is also much appreciated as the information reaches many more people than the session participants, and will be readily available on the IASSIST website at http://www.iassistdata.org.

Authors are very welcome to take a look at the instructions and layout:
http://iassistdata.org/iq/instructions-authors.

Authors can also contact me via e-mail: kbr@sam.sdu.dk. Should you be interested in compiling a special issue for the IQ as guest editor(s) I will also be delighted to hear from you.

Karsten Boye Rasmussen
January 2014
Editor
Guest Editors’ Notes

Beyond Bits and Bytes: The organizational dimension of digital preservation

The problem with digital preservation is often reduced to its technical aspects: we seek solutions to storage and bitstream preservation, format migration or building emulation environments. In turn, organizational aspects of preservation are often dealt with almost as an afterthought (see Kenney and McGovern, 2003). However, for digital preservation to be sustainable, technical solutions have to be embedded in an organizational framework complementing and supporting the technological side. Linked closely with the notion of “institutional readiness”, the organizational dimension of digital preservation comprises “the policies, procedures, practices, people” to support digital preservation (Cornell University Library and ICPSR, n.d.). These were the topics addressed in the session “Beyond bits and bytes: the organizational dimension of digital preservation” at the IASSIST 2013 conference in Cologne, on which this issue of the IASSIST Quarterly is based.

Looking back onto the last ten years of “digital preservation developments” in 2007, McGovern sees considerable progress in what she calls the “organizational leg” of digital preservation (no pag.). At the same time she identifies two particularly pressing issues: “the need to integrate the organizational policies for digital preservation into technological implementations and the need to develop and evolve digital preservation skills” (ibid.). It seems safe to assert that awareness of the importance of the organizational dimension has further increased since McGovern’s assessment in 2007, and that further progress has been made in addressing the practical issues arising in this dimension. This is, for example, a result of the increasing commitment of archives and repositories to audit and certification procedures – thus, all the standards in the European Framework for Audit and Certification of Digital Repositories require archives to address the organizational dimension along with the technological.

But despite the progress already made the articles in this issue attest to the fact that a number of challenges still face digital preservation (and individual institutions accordingly) in the organizational dimension. Expanding on the questions of change management, professional development and education, the creation of policies, and certification, the authors draw our attention to issues which we need to continue to address to further strengthen the organizational dimension and its connection to the other legs of the “three-legged stool for digital preservation” (McGovern, 2007, no pag.): technology and (human, financial) resources.

In her paper “Time to change – effects and implications of digital preservation in an organizational context” Michelle Lindlar from the German National Library of Science and Technology (TIB) presents a systematic overview of the prerequisites an institution has to fulfill when introducing digital preservation activities, and of the (organizational) changes this entails. Identifying two “layers of change” (initial, and ongoing), Lindlar carries out analyses of three areas requiring effective change management: preservation strategy, staff, and system choice. She shows which initial changes the introduction of digital preservation will effect in these areas and which ongoing efforts an organization has to make to keep its preservation activities up to par. To support this objective, the author recommends the implementation of an “organizational watch” in addition to a technology and community watch.

The next paper, “Digital curation training – the nestor activities” by Stefan Strathmann from the Göttingen State and University Library and Achim Oßwald from Cologne University of Applied Sciences addresses the issue of professional training and education in digital preservation. The authors describe the efforts made by a working group focusing on qualification issues set up as part of the nestor competence network. To improve education in this field in German-speaking countries, a qualification program has been developed and implemented during the last decade. The article introduces the approach taken to digital preservation training, which is based on e-learning tutorials, schools, seminars, and publications, and sketches future plans of the working group.

In “How to develop a preservation policy? Guidelines from the nestor working group,” Yvonne Friese from the Leibniz Information Centre for Economics introduces another activity from the context of the nestor network. Taking its cue from the observation that a gap exists between the number of institutions involved in digital preservation activities and those who have published a preservation policy, the nestor working group started with the creation of guidelines to help institutions to develop their own preservation policy. By supporting institutions to create a policy tailored to their individual scope and mandate, the guidelines, which will be published in early 2014, can help to close the observed gap and thus support the establishment of more sustainable and reliable preservation services.

In the last paper of this issue, “Tried and trusted: Experiences with certification processes at the GESIS Data Archive”, Natascha Schumann from the GESIS Leibniz Institute for the Social Sciences investigates the organizational impact that a certification process can have on an established
archive. After giving an overview of existing approaches to certification and audit for trusted digital repositories she discusses the GESIS Data Archive's experiences with the Data Seal of Approval and the benefits reaped from preparing the certification process.

Dr. Astrid Recker, Natascha Schumann
January 2014
Guest Editors

References


Abstract
Digital preservation is not a pure technical task – the first paragraph of the OAIS, the probably most cited work in the context of digital preservation, clearly says so: An OAIS is an ‘Archive, consisting of an organization, which may be part of a larger organization, of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community’ (CCSDS, 2012, p. 1-13). But what prerequisites does an institution have to fulfill for ‘digital preservation readiness’? And what effect does digital preservation have on the organization itself? How does the task digital preservation fare in juxtaposition with the necessary resources “organization”, “people” and “systems”?

Digital preservation within an organizational context requires different layers of change: During the implementation process the necessary resources have to be allocated and aligned. Since digital preservation is a cross-sectional task, drawing from the expertise of different stakeholders, this may require organizational change in form of new team structures. Furthermore, digital preservation is a rather fluid task, which requires constant monitoring and adaptability. Both change layers – the initial one as well as the ongoing one – have direct effects and implications on an organization.

Keywords: digital preservation, organization, change processes

Introduction
Digital preservation is often mistaken for a problem of technology. While the rapid technological change we have been facing since the rise of the personal computer and the growth of data that inevitably came with it is certainly the main risk of the preservation process, digital preservation is by no means solely a problem of technology. A frequently used definition of digital preservation is that it is not a promise of safeguarding information for five, ten or fifty years – but rather a promise to develop strategies which meet the constant change imposed on digital objects and their surroundings through the fast-paced development of new technologies (Schwens, 2004).

Memory institutions do have long-standing experience in the conservation of analogue information artifacts. But how are they faring in the preservation of digital objects? The question of organizational requirements and organizational change in the context of digital preservation is certainly not a new one. In 1996 one of the first high level research activities – the CPA/RLG (Commission on Preservation and Access / Research Library Group) “Task Force on the Archiving of Digital Information” – described the connection between risk mitigation and organizational changes: “Compounding the technical challenges of migrating digital information is the problem of managing the process in a legal and organizational environment that is in flux as it moves to accommodate rapidly changing digital technologies” (RLG/CPA, 1996, p. 6).

Digital preservation within an organization may be seen as a pyramid (see figure 1), where the organization type (i.e., national library, archive, research institution, etc.) and the mandate form the basis for all activities. Preservation activities are realized using a combination of three resources: people, hardware, and software. The resources are tied together in processes and workflows. And lastly, policies form a guideline framework for all activities.

But what are the prerequisites in the single building blocks when it comes to digital preservation? What changes will they face? This paper will briefly highlight needed prerequisites for organizational digital preservation readiness as well as needed ongoing change processes for three areas: preservation strategy, staff, and system choice. While the prerequisites will pick up on common misperceptions and highlight initial change processes, the ongoing changes will show what is needed in the ongoing maintenance of sustainable digital preservation processes.
Preservation Strategy

A preservation strategy formulates the "why and how" of digital preservation and should be the starting point of preservation activity within an institution. Comments one may come across when first talking about digital preservation in an institution may include: "Digital preservation will be solved by technology in a few years" – "We already have all our objects on spinning disks, so they are preserved" – "We've chosen PDF and therefore will never have to revisit our choice of file format" or even "Let's print it all out! Paper will last longer than anything digital".

A preservation strategy shall describe an institution's vision of maintaining accessibility and understandability of digital content and outline ways to achieve this vision. It will lead to one or several refined preservation policies. When formulating a strategy three main factor groups need to be considered: technological factors, organizational factors, and usability factors (see figure 2). Furthermore, the strategy needs to address the multifacetedness of a digital object and take into account the bit layer, i.e. the "ones and zeros"; the logical layer, i.e. the file format, and the semantic layer, i.e. the contextual understandability, alike.

In current discourse, preservation strategy is often used in the context of a chosen preservation action, i.e., migration, emulation and normalization (see Strodl et al., 2007; Van der Hoeven, 2005). In choosing a preservation action strategy, the implications of the aforementioned technological, organizational and usability factors certainly also need careful consideration. Furthermore, the type of content may limit the choice of preservation action – for complex archiving objects such as computer software, migration or normalization are certainly not suitable actions. However, preservation strategy which is only centered around a preservation action is limited to the question of logical preservation. Bit preservation and semantic preservation should be, as mentioned before, addressed in an overarching organizational preservation strategy.

Prerequisite

A preservation strategy can only be formulated based on a thorough understanding of the organization itself. Going back to the preservation pyramid (figure 1), the mandate plays a major role in the preservation intent. In regards to an archival mandate, different regulations may exist on a (sub-)collection basis, meaning that there might be a clear preservation responsibility only for a small part of the holdings and no preservation right for the rest. The mandate may also be tied to funding available for the task of digital preservation. Legal restrictions fall into the category of mandate – or lack thereof.

The role of the organization is closely tied to the consumer or to the designated community which the organization serves. University libraries, for instance, will target a different community than a highly specialized research institute or a company library. Expectations of the organization's stakeholders – data producers as well as data consumers – need to be taken into consideration. Lastly, the holdings need to be analyzed in regards to content type and technological factors such as data carriers, data formats and complexity.
Based on a thorough organizational analysis a preservation vision will be formulated, which will be put into the context of organizational resources and capabilities in order to formulate a strategy. Establishing a preservation strategy and first workflows based on the strategy is therefore the initial change process. While digital preservation efforts have so far been mainly driven by large cultural heritage institutions such as libraries and archives at a national level, more and more smaller organizations are now realizing the importance of digital preservation (Strodl et al., 2011). Their initial change process often starts with a single sub-collection for which an urgent preservation need exists, e.g., research institutions needing to guarantee the long-term availability of a research project’s data.

**Ongoing change**

The three main influence factors into a preservation strategy as shown in figure 2 are technology, community/users and the organization itself. As ongoing change in technology is the root risk to be mitigated in digital preservation practice, the task of “technology watch” is an established one in the preservation context. The same holds true for “community watch” which is the monitoring process targeting data producers and consumers. The intention of the producer and the intended (re-)use of data by the consumer need to be understood to fulfill the task of maintaining the information for the designated community, as defined in the OAIS reference model (CCSDS, 2012).

The third category also needs to be monitored – the organization itself. Changes in mandate, organizational structure or in resources such as funding, or the skill set of staff will significantly influence the preservation strategy. A continuous organizational watch as well as continuous internal and external lobbying for the preservation cause need to be included.

As the CPA/RLG Task Force pointed out, not only technology (and the community) is constantly changing, but also the organizational environment needs to change with it to fulfill the task of preserving the information throughout those external change processes (RLG/CPA, 1996).

**Staff – “the who”**

As digital preservation is not a mere question of technology, staff is an important factor in any preservation activity. When introducing digital preservation in an organization, various ideas about “the who” in digital preservation exist, such as: “Digital preservation is an IT task” – “The content specialists are responsible for the preservation” – “We have a classical conservation team – it should be their job”. Each of those statements has some truth to it, as digital preservation is a cross-sectional task, including technical aspects, content aspects and legal aspects alike.

However, the nature of being a cross-sectional task may complicate the question of where within an organizational structure digital preservation should be positioned. Furthermore digital preservation activities call for new skill sets. While there has been an increase of job advertisements for digital preservation staff over the past 3 years (Kim et al., 2012), dedicated digital preservation staff is still not the norm for all institutions with a long-term stewardship for digital data (DPOE, 2010).

Job positions related to digital preservation have various titles – data curation librarian, digital preservation analyst, web archiving engagement and liaison officer – to just name a few. But what requirements do organizations actively seek in preservation staff according to job descriptions? A recent study of 110 digital curation job descriptions showed that besides degree and job experience requirements, “working in an information technology intensive environment” (58%) and “familiarity with standards and specifications” (55%) were the most sought after qualifications. The fact that 45% of the analyzed job descriptions called for “project management skills” may be taken as an affirmative indicator for the project-status of many digital preservation programs as well as for the high interdisciplinary and cross-sectional work. Only 23% of the job descriptions listed “working knowledge for curation” as a requirement (Kim et al., 2012).

**Prerequisite**

As mentioned above, the position of digital preservation within an institution needs to be carefully considered and depends on different organizational factors. During the initial phase it is beneficial to position the digital preservation staff as close to the overall management level as possible. This is usually achieved through a first “project” phase of digital preservation and has the advantage of potentially shorter decision making processes.

As a permanent task some organizations position digital preservation within the information technology department – a decision which should be considered carefully, taking into consideration the needed input from other departments for tasks on the logical and semantic preservation layer. It can be assumed that due to the cross-sectional nature of digital preservation, it is a task easier to position within a multi-linear or matrix organizational structure as opposed to a top-down hierarchy organizational structure. One of the first change processes is therefore the formation of a new team or department. Roles and responsibilities for the workflows defined on the basis of the preservation strategy should be defined as early as possible.

Dedicated digital preservation staff should have a sound knowledge of both, information technology and traditional library processes. As an initial change process, staff suited for the intended preservation strategy needs to be allocated (Bähr et al., 2011).

Digital preservation also changes internal and external communication flows. The cross-sectional input needed for digital preservation tasks – e.g., in evaluating format choice options suitable from an organizational, technological and stakeholder point of view – create the need for new internal communication flows. On an external level, communication with data producers and data consumers need to be established to communicate recommendations and to query needs.

Kim et al. point out that dedicated curation staff might also provide reference services or specialized research consultancy to other staff/users (2012).

**Ongoing change**

As digital holdings to be preserved grow, input of “new” departments and stakeholders may be needed. As part of the organizational watch outlined in the previous chapter on preservation strategy, staff should closely watch changes in the digital holdings of the organization and take action where needed. There may also be a temporal need for special staff resources not unforeseen before, e.g., software development resources to develop tools which analyze data or automatically ingest data into a preservation system.
The most important ongoing change when it comes to staff is based on the fluid nature of digital preservation: due to the fast technological and organizational change, digital preservation is a commitment to lifelong learning. While recent projects like the DigCurv project have outlined initial and ongoing curricula for digital curators (Molloy, 2013), the digital preservation community itself forms a valuable information source through various blog, platforms, journals, workshops, and webinars (Bähr et al., 2011). Digital preservation staff must be willing to regularly update their knowledge on state-of-the art processes and technology and maintain close communications with the digital preservation community and the stakeholder community alike.

**System choice**

The last factor in this short analysis of change processes connected to preservation building blocks is the choice of preservation system – the “with what” of the preservation process. Again, first off, common comments heard in institutions when starting with digital preservation activities: “This software will solve all your digital preservation problems” – “How can software help me – software is the root of all problems” – “Software for digital preservation should be selected based on technical criteria only”.

Of course software is neither at the root of, nor is software itself the answer to all digital preservation problems. Software is, however, an integral part of digital preservation. It is needed for file analysis such as file format identification, technical metadata and validation; for preservation actions like migration or emulation; for accessibility with rendering software and for data management in repositories or digital preservation systems. In all of those tasks software is only as good as the processes and workflows that implement and use it. And – especially in the case of data management in repositories and digital preservation systems – software is only as good as the metadata stored with the object.

In this chapter we are not focusing on singular file analysis or rendering tools, but on larger preservation “systems” where the choice needs to be made between a system off-the-shelf (commercial or open source), a custom built system or an available service.

The system selection should be based on both organizational and technological decision criteria.

Custom built systems certainly have the highest degree of transparency. Quality and modularity of the system can be determined by the organization itself and licensing costs can be kept to a minimum. When extending existing open source repository systems, some form of community support is usually available. This comes at the cost of internal resources for initial integration and development as well as ongoing internal IT staff requirements for support and development, as the system needs to be adapted to meet new needs. Depending on organizational resources, the time from project start to roll-out of the solution can be longer than in the case of off-the-shelf systems or available preservation services.

Off-the-shelf systems usually require little to no internal IT resources for development. The resources needed for integration differ from solution to solution. Usually the time from project start to roll-out of the system is comparatively short. The continuous development of the solution is usually taken care of by the software provider who may also offer support or service models.

This, of course, may go hand in hand with licensing costs and is tied to a dependency on the software vendor. Also, as off-the-shelf systems try to target a wide customer base, there are usually drawbacks in fulfillment of institutional needs and integration of other existing systems may be tricky.

Available preservation services certainly have the lowest staff cost and no direct hardware and software cost. As there is little to no local technological implementation necessary on the organization’s side, the time from project start to roll-out of the solution is relatively short. The organization itself usually has no direct control over the data and all preservation actions are based on the decisions of the service provider. Access is usually only possible in pre-defined scenarios, so called “trigger events”.

**Prerequisite**

In order to make a system choice, the overall preservation strategy needs to be in place. As part of the preservation strategy, an analysis of organizational parameters like budget, staff and legal restrictions should have been evaluated. Depending on the strategy a number of prerequisites for the system choice may exist; the mandate, for example, may bind the organization to storage within its own premises, which rules out preservation services. Furthermore, technological parameters like the existing software infrastructure play a role in system architect decisions but also in regards to the availability of interfaces, if existing systems such as catalogue systems or retrieval platforms shall be integrated into the chosen system. If the prerequisites are fulfilled, the initial change process is a system choice and introduction based on the aforementioned criteria.

**Ongoing change**

It has been mentioned in the previous chapters that processes need to be adapted as new materials need to be preserved or new technology becomes available. This of course means a continuous change of the system, e.g., in form of including new workflows or in form of extending the system to suit new technologies. Monitoring the suitability for the digital preservation processes of all parts of the system itself should be an integral part of the aforementioned “technology watch” process. Scaling the system as the holdings grow is another change process which is to be expected.

An at first sight surprising, but very important preparation for a change process is the planning of the exit scenario from the get-go. As part of trustworthy digital preservation an organization needs to know how to completely extract data out of a system and move it to a new one. This scenario needs to be planned for from the start, as it may become necessary to change the system due to different factors, such as vendor problems or new state-of-the-art processes.

**Conclusion**

This paper highlighted organizational prerequisites and ongoing changes needed in a digital preservation process. Three areas were analyzed: preservation strategy, staff and system choice. It was demonstrated that all three sections require initial change processes in an organization when first introducing the task of digital preservation, as well as ongoing change processes in a continuous process.

As it has been pointed out by the RLG/CPA task force, digital preservation is a fluid process and the organization involved in it...
needs to adapt its processes alongside the task, to be able to meet the challenge (RLG/CPA, 1996). These changes hold true for all parts of the organization involved in the digital preservation process. A main outcome of this analysis is the importance of “organizational watch” – the re-evaluation of organizational factors in the preservation strategy. While technology watch and community watch are often included in best practice recommendations for digital preservation processes or as requirements in certification procedures, a constant self-analysis in form of an “organizational watch” is usually not mentioned yet.

References


NOTES

1. Michelle Lindlar works as a technical analyst and as a preservation researcher at the German National Library of Science and Technology (TIB). As part of her work, she is currently involved in the EU FP-7 DURAAK project, where she leads the digital preservation work package. After having worked as a system administrator and IT project manager, she joined the field of digital preservation in 2009 – and hasn’t regretted it for a minute since. She can be reached by email at: michelle.lindlar@tib.uni-hannover.de

2. Publicly available examples for two rather detailed preservation strategies are the joint digital preservation strategy of Archives New Zealand and the National Library new Zealand (Joint Operations Group, 2011) and the British Library’s preservation strategy for the years 2013-2016 (British Library, 2013).
Abstract
Digital curation is currently not very well covered by university curricula in the German speaking countries. Nevertheless there is a strong demand for well-educated staff in this field. As part of the project “nestor”, a transnational partnership of academic institutions in Germany, Switzerland, and Austria, a comprehensive qualification program based on e-learning tutorials, schools, seminars, and publications has been established to meet this demand.

Keywords: training, education, digital preservation, digital curation, nestor.

The nestor network
nestor, the Network of Expertise in long-term STOrage of digital Resources in Germany is a cooperative initiative of libraries, archives, museums and other parties interested or involved in digital preservation. From 2003 until 2009 nestor was funded by the German Ministry of Education and Research (BMBF). After 2009 it was transformed from a project into a sustainable membership organization, funded by the partners involved. At the moment (autumn 2013), 14 partner institutions are formal members of the nestor network.

However, the formal members are only the core of the network. Most of the nestor activities are organized and conducted in the open nestor working groups (WG). Currently about 60 institutions are involved in several working groups focusing on topics such as media, cooperative long-term preservation, standardization, policy, cost, etc. These groups are constituted and terminated based on need. Only the WG Qualification is based on a more formal cooperation and therefore organized in a different way than the other WGs (see nestor MoU Group Homepage).

The nestor WG Qualification
The nestor WG Qualification is acting since 2005 and organized around a Memorandum of Understanding (MoU) first signed in 2007 and renewed in 2009. The MoU provides an effective level of commitment but is at the same time less formal than a regular contract. This form was chosen because a formal contract bringing together so many different institutions might have involved too much bureaucracy. The MoU has been signed by 12 partner institutions from the educational sector and one coordinating nestor partner (see table 1). The group is involved in education for libraries, archives, museums, and IT in the German speaking countries (Germany, Austria, Switzerland).
(see table 1 page 15)

The WG’s objective is to stimulate and promote qualification in the field of digital preservation. The involved institutions cooperate in the development of course-materials, mutually accept credit points of courses regarding digital preservation and seek to establish a cooperative and distributed master’s degree program in digital curation as a long-term objective. Early on, in 2006, the WG’s agenda was shaped by the results of a survey on the coverage of digital preservation topics in Bachelor- and Master’s programs provided and planned by LIS departments in Germany, Austria, and Switzerland (see Oßwald and Scheffel, 2008). The lecturers from each institution stated that none was able to cover the topic in its entirety or to keep pace with the variety of the current developments in the whole field. Moreover, subject-related needs (e.g. of museums) resulted in a focus on special topics of digital curation in each institution. This initial situation was a good starting point for a cooperation between the different institutions and lecturers.

Project activities
The nestor WG Qualification – hereafter called MoU Group – offers five major lines of activities in digital curation-related qualification to meet existing needs:

- nestor seminars
- nestor schools
- nestor publications
- development of e-learning tutorials
- development of a cooperative curriculum
nestor seminars and schools

The nestor seminars on special topics (or for special audiences) take place occasionally. However, they constitute a strongly limited channel for the distribution of knowledge because of the demands they pose on the lecturers’ time and the topics covered, which often address a smaller audience with a higher degree of specialization. The MoU Group tried to expand the coverage and to relieve the lecturers by producing a video DVD with two recorded introductory seminars for self-study. As this approach was not very successful, the group decided that producing some essential publications was a more effective way of distributing knowledge (see below).

The series of nestor schools offered by the MoU Group is a success story. Since 2007 seven nestor schools were conducted. For the duration of three to six days attendees and renowned lecturers come together for an intense exchange of ideas in the privacy of a remote location. The courses are a combination of lectures and practical exercises, augmented with several social activities. The participants, a heterogeneous audience comprised of professionals from libraries, archives, museums, companies, and administration as well as students, are presented a perfect opportunity to build and enlarge their professional networks, to make contacts with colleagues, and to discuss ideas and conceptions. Upon completing the course, attendees receive a certificate equivalent to 2 ECTS credits (European Credit Transfer and Accumulation System). The school

<table>
<thead>
<tr>
<th>Table 1: Institutions participating in the nestor MoU group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Partner</strong></td>
</tr>
<tr>
<td>Cologne University of Applied Sciences, Faculty of Information and Communication Sciences, Institute of Information Science</td>
</tr>
<tr>
<td>University of Applied Sciences – Hochschule Darmstadt</td>
</tr>
<tr>
<td>Georg-August-Universität Göttingen, Göttingen State and University Library (coordinating nestor partner)</td>
</tr>
<tr>
<td>Humboldt-Universität zu Berlin - Faculty of Arts I - Institute of Library and Information Sciences</td>
</tr>
<tr>
<td>Archives School Marburg - University of Applied Studies for Archival Science</td>
</tr>
<tr>
<td>University of Applied Sciences Potsdam, Department of Information Science</td>
</tr>
<tr>
<td>Stuttgart State Academy of Art and Design</td>
</tr>
</tbody>
</table>
events have been conducted in cooperation with several digital preservation projects (DELOS, DPE, DigCurV) and with some support from industrial partners (PDF/A Competence Center, EMC, Sun Microsystems). Throughout, the events received very high evaluation marks from the participants.

**nestor publications and tutorials**

Because of a lack of course materials and to reach a broader audience compared to the nestor seminars, the nestor MoU Group initiated the production of the nestor handbook (Neuroth et al., 2010). This German-language “encyclopedia” bundles the recent state of knowledge on digital long-term preservation and its various components. It collects contributions from over 50 authors on more than 630 pages and has been revised and expanded several times. Beside the open access online edition (version 2.3), a printed book (version 2.0) is available. The handbook tries to cover the whole field of digital preservation and contains chapters on topics like: State-of-the art of legal aspects, Preservation Policies, The OAIS (Open Archival Information System) reference model, Trusted digital repositories, Formats, Relevant standards important in digital curation (e.g. PREMIS), Strategies of digital preservation, etc. The handbook is very well accepted as a standard work on digital preservation in the German speaking community and it is used by students as well as by practitioners.

In 2012 a comprehensive baseline study on research data management and curation was published (online and print) in German language (Neuroth et al., 2012). It is the analysis of a structured survey regarding the curation of research data in Germany and holds reports from eleven academic disciplines (medicine, astronomy, humanities, and social sciences among others.). A condensed English-language version, entitled “Digital Curation of Research Data. Experiences of a Baseline Study in Germany” has just been published online and in print (Neuroth et al., 2013).

Since 2007 digital preservation e-learning modules – based on the e-learning platform Moodle – have been developed within student projects. Created by students for students, these tutorials are mutually exchanged between the nestor MoU partners and used in their university courses. Several tutorials, dealing with topics like ‘Introduction to Digital Preservation’, ‘The OAIS model’, ‘Formats’, ‘Digital Preservation of GIS-data’ etc. have been maintained and updated regularly. Adhering to the concept “e-learning tutorials by students for students” adopted by the nestor MoU Group means that the modules can only be developed in combination with university courses. This approach has some implications: the development of e-learning tutorials proceeds very slowly and must be integrated in the regular curriculum. The initial input given by the teachers is extremely high. At the same time the acceptance of using these tutorials is high as well. Despite this acceptance it has become apparent, however, that there are issues to secure the quality of the tutorials regarding the content – which needs frequent updating – as well as the didactics, which should respond to various preconditions of students of different levels. Therefore the tutorials are expected to become obsolete very soon. At the moment, the group discusses other ways of organizing the teaching material to keep the time and effort needed reasonable. The new concept is not only a big challenge but also a great chance to cooperate and collaborate.

**Towards a shared digital preservation curriculum**

As already mentioned none of the partners active in the nestor working group is able to set up a preservation-centered curriculum on its own. Instead, the partners have agreed to realize this on a cooperative basis. Accordingly, as a long-term goal the MoU Group intends to set up a cooperative curriculum. This plan received support by experts and the administration of several universities in the German speaking countries. In the Memorandum of Understanding the thirteen partner institutions declared their intention to cooperate in developing building blocks for a digital preservation curriculum and adjusting modules focusing on selected topics. One first step in the direction of a cooperative curriculum is the agreement that credits earned in digital preservation-related courses can be transferred between the universities involved and are therefore accepted in the local curricula.

The curriculum development activities led to the engagement of the MoU Group in the context of the Digital Curator Vocational Education Europe project (DigCurV; http://www.digcur-education.org/). The MoU Group was a founding partner of the project and contributed in various ways to the development of the DigCurV Curriculum Framework (see DigCurV, 2013).

The cooperation between nestor and the educational institutions in the MoU Group results in several synergies. The nestor MoU Group has stimulated reflections and awareness of digital curation even in those universities where the topic has been part of courses or programs for years. As a consequence results of the nestor network have influenced the content and the quality of lessons and courses provided. Not by intention but as a secondary effect the nestor MoU Group has initiated intra- and cross-sectoral cooperation where competition seemed to dominate. On the national level the nestor network in Germany has gained insights and created links to universities and qualification bodies. This has improved the awareness and understanding of issues related to qualification for preservation and curation activities. Initiatives, promotional programs (e.g. by the German Research Foundation) and activities now address qualification issues regularly.

**Summing up**

Since 2007 the MoU Group – comprising of 13 Higher Education Institutions in the German speaking countries Germany, Austria, and Switzerland – has initiated and contributed to the awareness of the importance of qualification issues in the field of digital preservation and curation in the countries involved. Members of the MoU Group and their universities have provided curriculum-based courses, a variety of qualification events like, for example, the nestor School events, and publications such as the nestor handbook, a state of the art publication on issues, methods and best practice solutions in the field of digital preservation and curation. They are still involved in improving the curriculum-based qualification as well as further collaborative education activities in the field. As a side effect of the MoU Group’s activities, qualification in the field of digital preservation and curation have become a regular topic of related programs and activities in and beyond the nestor context.

**References**

NOTES

1. Stefan Strathmann is Head of the Competence Center on Digital Preservation and Research Data at the Research and Development Department of the Göttingen State and University Library. He is one of the representatives of the nestor MoU Group on qualification in digital preservation. He can be reached by email: strathmann@sub.uni-goettingen.de.

2. Achim Oßwald is Professor at the Institute of Information Science, Cologne University of Applied Sciences. His focus of research and teaching is IT applications in the LIS field including digital curation issues. He is one of the representatives of the nestor MoU Group on qualification in digital preservation. He can be reached by email: achim.osswald@fh-koeln.de

3. In the context of this article we will use the terms ‘digital preservation’ and ‘digital curation’ synonymously.
How to Develop a Preservation Policy

Guidelines From the nestor Working Group by Yvonne Friese

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).

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 well as 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.

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

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 wiki3. 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


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.
Abstract
The need to prove and improve trustworthiness is an issue not only for new archives but also for those who have been doing this job successfully for a considerable time. But what happens when an established data archive faces the challenges of certification and audit processes? Founded in 1960, the GESIS Data Archive for the Social Sciences has been "in the business" for more than 50 years and is one of the oldest archives in Germany to preserve electronic resources for the long term. Driven by a growing awareness of the needs of its stakeholders, who have to be sure that the data they produce, use, or fund is treated according to common standards, the Data Archive started a process of audit and certification within the European Framework for Audit and Certification.

After giving an overview of the GESIS Data Archive and the European Framework for Audit and Certification, this article describes how existing workflows were evaluated with regard to the requirements of the chosen level of certification. While the workflows themselves are already in place, in some cases the evaluation process showed a lack of appropriate documentation. Suitable documents have to be created and made available to the public. In some cases, this process has to be accompanied by discussions within the institution about the mission and goals of the archive. In our experience, these can be very productive and lead to a common understanding and an improvement of services.

Keywords: Digital preservation, trusted digital repositories, certification.

Figure 1: Research Data Life Cycle
The GESIS Data Archive and its organizational context

As an infrastructure institution for the social sciences, GESIS not only carries out research but provides services in all phases of the research data life cycle (see figure 1), from the conception to the archiving and re-use of social science research. Among others, it provides consultation in methodology, develops software tools for research, and offers information services.

The GESIS Data Archive for the social sciences is one of five departments of GESIS and has been providing comprehensive data services for national and international comparative surveys for several decades. One of its main tasks is to make research data available for re-use. To support this goal, the data archive has an explicit mission for long-term preservation, which is also laid down in GESIS’s by-laws. Accordingly, among GESIS’s primary objectives is the archiving, documentation, and long-term preservation of social sciences data, including the indexing of data as well as the high-quality enhancement of particularly relevant data to prepare them for re-use (GESIS Constitution § 2).

Workflows within the archive are organized according to an archival life cycle, ranging from pre-ingest (incl. acquisition) to ingest and processing to archival storage up to the dissemination of data. The central functions of the OAIS reference model (CCSDS, 2012) can be mapped to the existing structure of the archive (see Schumann and Recker, 2013).

However, although the GESIS data archive already has working procedures and processes in place to ensure the preservation of its data, there is still a need for further activities. For example, in some cases documentation of workflows and defined interfaces between different steps of the preservation process are lacking, and some definitions of information packages are not up to date. By addressing these issues in a systematic fashion, the archive aims to further increase its trustworthiness.

A need for trust

A definition of a trusted digital repository is given by the RLG/OCLC Working Group: “A trusted digital repository is one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future” (Research Libraries Group, 2002, p. 1). But what does that mean in detail? Audit and certifications standards such as the nestor Catalogue of Criteria for Trusted Digital Repositories on the one hand add aspects from an IT security perspective – for example, “authenticity, integrity, confidentiality and availability” (nestor, 2009, p. 1). These technical aspects are relevant issues for trust, but beyond that organizational aspects are just as important. As pointed out in Audit and Certification of Trustworthy Repositories (CCSDS, 2011), “[c]onstant monitoring, planning, and maintenance, as well as conscious actions and strategy implementation will be required of repositories to carry out their mission of digital preservation” (p. 2-1). These quotations show that building trust depends on more than one factor. A trusted digital repository has to ensure that the digital objects it preserves are not corrupted by accident or intentionally, and that access is given – not only physically, but also in appropriate digital formats. Another criterion of trust is if and how the organization demonstrates its know-how in digital preservation and, for example, if succession plans exist for the case that the institution ceases to exist. Thus, transparency is very important in the context of trust. All stakeholders should have the opportunity to ascertain the statements made by the institution.

Accordingly, to appear trustworthy, the GESIS data archive has to provide stakeholders – data depositors, data users and funders – with sufficient information to demonstrate that their data is treated according to the agreed standards of the social sciences and digital preservation communities. Because existing certification standards and audit tools support archives in the building of trust, we decided to start a process of audit and certification within the European Framework for Audit and Certification (see below).

Our decision to do so coincided with similar efforts initiated by the Council of European Social Science Data Archives (CESSDA), of which GESIS is a member. As CESSDA has been transformed into a new organization and legal form, CESSDA AS, and is on its way to becoming a European Research Infrastructure Consortium (ERIC), it is necessary that all member institutions agree on the same standards regarding trustworthiness. To start off this process, during 2013 all archives carried out a self-assessment based on the guidelines of the Data Seal of Approval (see below).

All of this illustrates that the need to prove trustworthiness is not only an issue for new players, but also for established ones like the GESIS Data Archive. However, the challenges such “established players” face are somewhat different from those that new archives have to deal with: It is a different kind of procedure to set up a completely new service or to conduct a certification process in an existing system. Thus, when setting up a new archive it is possible to take into account the requirements for trusted digital repositories from the outset. What is more, new archives can benefit from other institutions and their experiences and avoid mistakes. In contrast, an existing archive may have gained a lot of expertise and know-how over time, but it can be very complex and challenging to adapt established workflows to new requirements.

European Framework for Audit and Certification of Trusted Repositories

Over the years, many different approaches and standards have been developed in the field of audit certification for trusted digital repositories. The most established among them are:

- the Data Seal of Approval (DSA), originally initiated by DANS,
- the nestor Catalogue of Criteria for Trusted Digital Repositories, which became a German DIN standard (DIN 31644) in 2013 and will also be available in English, and
- the Trusted Repository Audit Checklist (TRAC), which is also an ISO standard (ISO 16363).

To achieve greater harmonization between these different initiatives and criteria catalogues, a Memorandum of Understanding (MoU) was signed in 2010 for a European Framework for Audit and Certification. This process was accompanied by the European Commission and the Alliance for Permanent Access to the Records of Science (APARSEN). The MoU defines three levels of certification (see figure 2):

1. Basic Certification is granted by obtaining the DSA.
2. Extended Certification requires completing the DSA and an externally reviewed self-audit based either on ISO 16363 or DIN 31644.
3. Formal Certification requires completing the DSA and a full external certification based either on ISO 16363 or DIN 31644.

Data Seal of Approval

The target audience of the DSA are repositories committed to long-term preservation. Working from the assumption that data quality is dependent on aspects related to the creation, storage and (re-)
The use of digital data (DSA, 2013, p. 5), the DSA contains 16 Guidelines reflecting different roles: Data producers, data repository and data users. Although the main focus of the DSA is on data repositories, it is open to other digital archives as well. There are different levels of compliance for each guideline:

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Not Applicable</td>
</tr>
<tr>
<td>1</td>
<td>No. We have not considered it yet</td>
</tr>
<tr>
<td>2</td>
<td>Theoretical. We have a theoretical concept</td>
</tr>
<tr>
<td>3</td>
<td>In Progress. We are in the implementation phase</td>
</tr>
<tr>
<td>4</td>
<td>Implemented. This guideline has been fully implemented for the needs of our repository</td>
</tr>
</tbody>
</table>

To be awarded the DSA, the minimum compliance level as stated in the DSA guidelines has to be reached by the applicant (see DSA, 2013, p. 6). The Archaeology Data Service has published a best practice report to support other institutions in obtaining the DSA (Mitcham and Hardman, 2011).

**nestor Seal/ DIN 31644**

The DIN 31644/nestor Seal is based on the nestor Catalogue of Criteria for Trusted Digital Repositories (2009). In 2012 it was accepted as the German Standard DIN 31644. It contains 34 criteria covering the following thematic areas: organizational framework, handling of information objects and their representations, infrastructure and security. The level of compliance is measured on the following scale:

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No concepts are in place</td>
</tr>
<tr>
<td>3</td>
<td>A concept exists</td>
</tr>
<tr>
<td>6</td>
<td>Well-elaborated concept</td>
</tr>
<tr>
<td>10</td>
<td>Implemented</td>
</tr>
</tbody>
</table>

**RAC/ISO 16363**

The Repositories Audit Checklist (RAC) was developed from the Audit Checklist for the Certification of Trusted Digital Repositories (2005) and Trustworthy Repositories Audit and Certification: Criteria and Checklist (2007). It became an ISO standard in 2011 and information about the standard can be found at the Primary Trustworthy Digital Repository Authorisation Body (ISO-PTAB). RAC consists of 50 main criteria and has 109 criteria in total. Their structure is orientated towards the nestor Catalogue and accordingly RAC criteria cover the following areas: organizational infrastructure, digital object management, infrastructure and security risk management.

**Our Approach**

As stated above, the GESIS Data Archive decided to commence its certification activities with the DSA. Several reasons contributed to this decision. First of all, the DSA is a good starting point for certification activities because it addresses all basic aspects of trust but is not as detailed as either DIN 31644 or ISO 16363. Thus the DSA is an immensely helpful tool to gain an overview of the processes within our archive. It therefore helps us lay the groundwork for follow-up activities in audit and certification within the context of the European Framework.

**Getting started**

As the DSA is a self-assessment, the applicant completes a self-assessment statement for each of the guidelines including links to the relevant documentation or evidence (see DSA, 2013, p. 6). This will then be reviewed by a peer reviewer appointed by the DSA-Board.

As a first step we evaluated the existing workflows with regard to the requirements stated in the guidelines. In this manner we obtained an overview of those workflows already in place and supported by sufficient documentation. This initial evaluation showed that the majority of our workflows comply with the DSA guidelines. However, a lack of appropriate documentation, especially on our website, became apparent.

In consequence, the main tasks at this stage were:
- The detection and subsequent creation of missing documentation and documents, e.g. policies or recommendations.
- The revision of existing documentation and documents if those were not up to date or not yet ready for publication. E.g. a versioning policy had to be updated.

An example for a newly created document is our preservation policy: It contains information on the organizational context of the GESIS Data Archive, states our mission, and describes the main
principles of our approach to the digital preservation of data (see also Friese in this issue). The process of developing the policy not only meant creating the content, it was also a process requiring a good deal of coordination: staff members from different teams had to be involved as well as the head of the department. As it is insufficient for the certification process to publish the policy only in German, an English translation was also created.

As explained above, an important requirement for trustworthy digital repositories is transparency: all relevant information should be available to the public. In our case this meant redesigning the archive’s website and deciding which information it should include. In this process, we reconsidered the whole structure of the website. It is now organized along the steps of the data lifecycle and refers to the functional entities in the OAIS reference model. Creating the website content was another challenge and involved more than simply adding some documents. We had to answer the question of how to make the website helpful for the different stakeholders and user groups: data producers, data users, funders, and other interested persons – both in terms of content and the language used.

**Experiences and Benefits**

One of the (first) benefits of the preparatory work on the DSA was that we gained an overview of what we already have in place and what will have to be amended or improved. It became clear that we would have to reconsider some of our workflows.

Traditionally, the focus of the GESIS Data Archive has been the indexing and processing of empirical social research data. Over the past years, more and more digital preservation issues became relevant – for example, questions of how to implement procedures to ensure authenticity and integrity, or the need to define different archival packages referring to the OAIS model etc. But not only did the archive have to update workflows; an important part of this process (which is not completed yet) was the development of a common understanding of digital preservation: Is it an “an added value” that we somehow create on top of everything else that the archive does? Or is it not rather the sum of everything we do in the archive: the bundle of measures, that is, that we employ to ensure access and use of the data in the long term, including, for example, the creation of metadata and documentation, DOI registration, etc.?

In addition, the strengthened focus on digital preservation has consequences not only with regard to processes and procedures but also for the level of transparency. Stakeholders have an increasing interest in learning how their data is curated and preserved, and the archive has to provide the respective evidence in order to maintain its stakeholders’ trust.

However, the DSA application not only required us to address external stakeholder communication; it also prompted a review and evaluation of our internal communication and documentation procedures. For example, the archive staff uses a wiki for internal documentation purposes. It was built and filled with content over the last years, but our review in relation to the DSA application showed that it was not up to date – neither with regard to the contents nor to the structure of our workflows. We are now in the process of restructuring and updating it, which also includes agreeing on the current procedure of maintaining the wiki and keeping it up to date.

The processes set in motion by our decision to apply for the DSA have helped to create awareness of the capabilities and strengths as well as of weaknesses or gaps within our archive. The systematical compilation of existing and relevant information and documentation required by the DSA is a helpful step in itself, and this gap analysis has helped us gain a more concrete idea of our workflows and their documentation instead of the vague feeling that “surely, everything works as it is supposed to.”

Preparing the DSA application served as an incentive to continue some projects – e.g. for new services or the adoption of standards – that had been planned for some time but had been neglected in the face of “more pressing” problems arising in our day to day work. Some of the required measures are easily created and implemented, but others need more time and discussion to be realized. The fact that so many activities are linked to each other entails that it may take some time to implement new processes: the necessary changes concern different applications or workflows cutting across different teams and cannot be made without implications for other parts of the system.

The process of (preparing) an audit or certification is very time consuming. But in our experience, the process of preparing the DSA self-assessment statement produced many valuable effects in that it helped us establish a common understanding for the mission and goals of our archive. The first steps of complying with the guidelines of the DSA have been made and we have gained an overview of our capabilities as well as existing gaps. Accordingly, we now know where we stand and what our tasks for the near future are. Our next step will be to hand in the DSA application. After this is completed and the DSA will have been granted, we are already planning to take the next step in the European Framework for Audit and Certification: the extended certification, which will take the form of conducting a self-audit for the nestor Seal, based on DIN 31644.

**References**

tions/archive/652x0M1.pdf> [Accessed 03 December 2013]

Consortium of European Social Science Data Archives (CESSDA-AS). Homepage. [online] Available at: <http://www.cessda.net/> [Accessed 03 December 2013]


DIN, 2012. DIN 31644: Kriterien für vertrauenswürdige digi-
tale Langzeitarchive. Available at: <http://www.nabdin.de/
cmd?level=tpl-art-detaillansicht&committeeid=54738855&arti-
d=147058907&languageid=de&bc уровне level=3> [Accessed 03 December 2013]

Friese, Y., 2014. How to develop a preservation policy? Guidelines from the nestor working group. [pdf] In: IASSIST Quarterly [include inform-
ation for this issue].


NOTES

1. Natascha Schumann is affiliated at the Data Archive for the Social Sciences at the GESIS Leibniz Institute for the Social Sciences in Cologne. The main focus of her work is on digital curation of social science research data and audit and certification in this area. Her contact email is natascha.schumann@gesis.org.

2. This paper is an updated version of a presentation given at the IASSIST 2013 conference in Cologne.

3. GESIS Constitution (in German): http://www.gesis.org/das-institut/der-verein/satzung/


Heike Neuroth, Stefan Strathmann, Achim Ößwald, Jens Ludwig (Eds.)

Digital Curation of Research Data
Experiences of a Baseline Study in Germany
Toronto Welcomes You To The 40th Anniversary Conference of IASSIST
The International Association for Social Science Information Services & Technology

June 3 - 6, 2014

The “official” hashtag is #iassist2014 #iassist40 — spread the word!
The International Association for Social Science Information Service and Technology (IASSIST) is an international association of individuals who are engaged in the acquisition, processing, maintenance, and distribution of machine readable text and/or numeric social science data. The membership includes information system specialists, database librarians or administrators, archivists, researchers, programmers, and managers. Their range of interests encompasses hard copy as well as machine readable data.

Paid-up members enjoy voting rights, benefit from reduced fees for attendance at regional and international conferences sponsored by IASSIST. Join today by filling in our online application:

http://www.iaassistdata.info/

Online Application

IASSIST Member ($50.00 (USD))
Subscription period: 1 year, on: July 1st
Automatic renewal: no

Please fill in the information our Online Form

The application is in USD, however, we do accept Canadian Dollars, Euro, and British Pounds as well.

The membership rates in all currencies as well as the Regional Treasurers who manage them are listed on the Treasurers page.