

Archiving Archaeological Research Data. On Requirements, Objectives and First Experiences from a German Respectively Saxon Point of View

REINER GÖLDNER, Saxon Archaeological Heritage Office, Germany

Preserving digital data over long periods of time is an ambitious task. While most analogue documents are best stored without touching them, digital data need permanent care and curation. Regarding some differences between analogue and digital data—lifecycle, complexity, functionality and so on—it can be seen that digital archiving is a new challenge. This paper will deal with the requirements and objectives of archiving archaeological research data, as discussed in a working group of the Association of State Archaeologists in Germany (Verband der Landesarchäologen in Deutschland – VLA). Interesting themes of the discussions were “Archive Objectives,” “Worthiness of Archiving,” “Suitability of Archiving” and “Future Use Scenarios.” This paper will also deal with first experiences in building up a digital archive for archaeological research data in Saxony. The archive follows the principles of the “Open Archival Information System” (OAIS). It is based on professional software that has to be adapted to the special requirements of archaeological data such as excavation documentation. OAIS defines various complex functions to be performed, e.g., data management and preservation planning. Preservation strategies are presented here to give an idea of how to preserve the content and function of archived material while the systems for archiving are changing over time. This is an alternative to the idea that knowing file formats will solve all preservation problems. Archiving is the most intensive and most expensive way to store data. But are there no alternatives? This paper will address some small steps that lead towards proper archiving without large budgets. Such steps will help to reduce the risk of losing important digital information, and will also bridge the gap until proper preservation tools become available at a reasonable price.

This article may be helpful for archaeologists that are about to build up an archive to preserve digital data for the long term. The concept of preservation strategies presented here may also be of interest for specialists in digital archiving.

Key words:

Archiving, long term preservation, archaeological research data.

SDH Reference:

Reiner Göldner, Archiving Archaeological Research Data. On Requirements, Objectives and First Experiences from a German Respectively Saxon Point of View, 2017. SDH, 1, 2, 165-181.

DOI:10.14434/sdh.v1i2.23205

Author's address: Reiner Göldner, Landesamt für Archäologie Sachsen, Zur Wetterwarte 7, 01109 Dresden, Germany; email: reiner.goeldner@lfa.sachsen.de.

Permission to make digital or hardcopies of part or all of this work is granted without fee according to the open access policy of SDH.

© 2017 SDH Open Access Journal

1. INTRODUCTION

This paper is about archiving digital archaeological research data that typically emerge from archaeological excavations and research projects. Archiving is considered to be long term preservation that is not limited in time. Archiving means to permanently preserve the content and function of digital data and to provide access to it. In contrast to other views about archiving, the main difference is that curation (realization of preservation planning, preservation strategies) is included. Because of the curation effort, it can be stated that archiving is the most intensive and most expensive way to store and to preserve data. This paper results from two main working processes in which I participated, that took place almost parallel to each other and inspired each other in many ways. First, a working group within the Association of State Archaeologists in Germany held an intensive discussion on archiving digital data, starting with several details about file media and file formats and later reaching more general aspects like objectives and requirements. This discussion was very important because until now there has been no full digital archaeological archive at work in Germany, neither centralized nor regional. Thus, all participants were very interested in finding ways to better preserve the unique archaeological data for the long term. Secondly, the Saxon Archaeological Heritage Office decided to take the first practical steps in the field of digital archiving, in order to get practical experience in this field. Dealing with a real digital archiving system quickly revealed problems, e.g., the effort to basically run the system and the lack of specific preservation strategies for our data. The work on appropriate solutions led to my ideas about preservation strategies and about small steps towards archiving.

This article may be helpful for archaeologists who are about to build up an institutional archive to preserve digital archaeological research data for the long term; the paper offers some helpful ideas about basic objectives and requirements. If they have rather small budgets, some advice is given regarding small solutions. There is also a part that may be of interest to specialists on digital archives. This section about preservation strategies presents a new conceptual framework designed to deal with several kinds of archival material, and it is an alternative to the idea that knowing file formats will solve all preservation problems.

The first topic to be discussed, "Objectives and Requirements," contains results from a working group of the Association of State Archaeologists in Germany. It can be seen as a consolidated view of several specialists with a focus on some more general aspects of archiving, i.e. archival objectives and requirements.

The next topic, "Preservation," contains a very short introduction to the well-accepted reference model OAIS (Open Archival Information System) and some ideas about preservation strategies that describe how to curate and preserve the content and function of archived material while systems are changing over time (assuming a migration method).

Two years ago, the Saxon Archaeological Heritage Office started the project "Archiving Archaeological Research Data" with the aim of studying the opportunities and risks of operating a digital research archive. The third topic to be discussed concerns "First Experiences with Implementation" of this archive.

For all those who work with small budgets “Small Steps towards Archiving” are presented at the end. Taking such steps will help to reduce the risk of losing important digital information and will bridge the gap until proper preservation tools become available at a reasonable price. The aspects presented in this chapter are based on ideas I collected from my own work, from the discussions at the working group of the Association of State Archaeologists and from third parties, e.g., Gollins 2009.



Figure 1. Project Teaser “Archiving Archaeological Research Data” (© photo/assemblage R. Göldner)

2. OBJECTIVES AND REQUIREMENTS

The information in this section is the result of intensive conferencing and discussion at the thematic working group of the Association of State Archaeologists in Germany [Archiving Group of VLA]. Thus, it can be seen as a consolidated view of several specialists (from six German states) that actively address digital archiving in their work. The main goals of that working group were to exchange knowledge and experiences of its members in the field of digital archiving, to bundle archival expertise of the decentralized system of German state archaeology and to support those archaeologists or institutions that are willing to take some first steps towards digital archiving. After dealing with many special questions of digital archiving, like data formats and secure storage, the working group focussed on some less technical and more general aspects that are important in running a digital archive as an organization, i.e. archival objectives and requirements.

The following sections present important results of the discussion. These results and some other interesting questions (e.g., should additions, corrections, and editing of archaeological archive material be allowed? What does authenticity mean in relation to digital archive material? What about usage rights for the archive?) can be found in “Ein Erfahrungsaustausch zur Archivierung digitaler archäologischer Daten ...” [Göldner 2016].

2.1 Archive Objectives

The Archiving Group of VLA stated the objective of an archaeological archive as follows: The aim of digital archiving is to permanently preserve digital content and digital function of archived material and to provide access to it [Göldner, R. et al. 2015, 5]. To clarify the difference between “content” and “function” in this context we assume that “content” is information that is used by people, visual or sensory, and that “function” is information that is used by machine, operational or algorithmic. This objective can be detailed a little more:

- to securely store the material (prevent damage and loss);
- to regularly curate the material and to restore or renew it, if needed;

- to index and register the material;
- to present the material and to make it accessible.

With regard to the use of archived material, archaeological archives are considered as information sources for:

- the present work of archaeological offices
- heritage protection and conservation
- publication and presentation
- future study and research

2.2 Archival Value

Archive material is material of archival value that is suitable for archiving. Archival value answers the question “which material, which content, which function is worth of archiving?” and suitability for archiving answers the question “which material is suitable for archiving from a technical point of view?”

In general, material is worth archiving if it has ageless, steady value to society. In most cases, “steady value” is interpreted to mean that the material is of particular importance, i.e., it is more prominent than the average material. Archive material is a selection of material, because in general not all existing material is worth archiving. In practice, this requirement has to overcome two challenges. First it provokes the opposition of people that cannot throw things away, and second, it demands discussion and establishment of criteria to appraise the potential of the archive material.

This general statement can be applied to typical archaeological institutions; the Archiving Group of VLA mentions the following criteria:

- if it supports future study and research;
- if it documents heritage that cannot be preserved itself;
- if it documents ephemeral archaeological situations or activities (e.g. excavation documentation);
- if it supports heritage protection and conservation;
- if it is a result of archaeological research.

On the other hand, examples of material that is not worth archiving are given: drafts, duplicates, intermediate states, incomplete, faulty or unfinished material.

Archivists have to check the archival value of the material under consideration. To do so they need archaeological expertise and they have to negotiate specific criteria for specific materials with the producers (i.e., archaeologists). Archiving is the most intensive and most expensive way to store and to preserve data. Thus, criteria for archival value are needed in order to use existing but limited archival resources in a most effective and reasonable manner.

2.3 Suitability for Archiving

Besides archival value there is another aspect to preselection of archive material: which material is suitable for archiving? The answer depends on formal, technical aspects of the archive material, such as type, complexity, connectivity, interoperability and other factors. All these factors affect the preservation strategy, so we can state that material is suitable for archiving if the archive has a reasonable (and specific) preservation strategy for it. From my point of view, an archive must not accept any material for which it does not have any basic preservation strategy. This is a critical aspect, because it requires the archive to integrate competencies of both information technology and the application field (archaeology), a requirement that archives usually are not familiar with.

2.4 Future Use Scenarios

More than traditional analogue archives (because digital material ages faster), digital archives have to be concerned about scenarios of future use. But interestingly, there seems to be a lack of imagination; my question about future use was very often answered with the words “just like now.” But this answer does not help much to establish criteria for preservation strategies that have to guarantee adequate use in 100 years. But let us have a look at some basic ideas we discussed at the working group, with the result that the question of future use scenario was answered as follows.

How will the archived material be used in 100 years?

- It is passively used as information (overview and detail), that is, for instance, in the fashion of articles in popular or scientific journals.
- It is actively used for scientific re-analysis, i.e., to deeply analyze the archive material with scientific methods and tools with the goal of creating new scientific results.

Which future use types are possible?

- Visual use, like viewing and reading, maybe also printing, is easy to imagine.
- Machine analysis of texts is quite clear, and may be supplemented by image analysis or similar tasks.
- Functional use with special application software (CAD, GIS, etc.) is hardly predictable even for few years, but it is and it will be the most interesting use type of digital data.

Imagining future use scenarios varies according to three hypotheses:

1. Future use scenarios will be limited to present use scenarios, regarding the presently collected information.
2. Questions of future generations are hardly predictable. Probably the data will appear to future archaeologists as a heterogeneous data mound.
3. With a maximum of collected information (more than presently can be used), there will arise new use scenarios in future that we cannot foresee today.

3. PRESERVATION

In general, there are four preservation methods for digital data, and they all have advantages and disadvantages. The hardcopy method tries to generate a copy of the data at a durable medium like paper or PET microfilm. It is very durable (LE500 material is guaranteed to survive at least 500 years), but it needs appropriate storage areas, and it is not well suited for functional data. The technology preservation method tries to preserve current hardware (and software) as long as possible for future use. Similar to this approach, the emulation method emulates old computer hardware with modern software and modern hardware. Both methods provide an adequate system environment for old data but they are rather complicated and expensive. The migration method uses regular adaptations of the archive data to new computer systems. It is a very flexible and powerful method, hence it is often preferred.

The migration method depends on appropriate ways to migrate the archived data without loss of significant content and function. Most information in this field is available for preferred data formats, so migration seems to depend on data formats. But this strategy does not fit all needs of archived materials; complex and functional data especially need more complex preservation strategies. As an alternative, a small conceptual framework of four preservation strategies is presented below. These preservation strategies depend on archive material types and lead to several types of archive representations. They also use preferred data formats, but not as a key issue.

This section addresses two important aspects of preservation. The first is the well accepted reference model OAIS (Open Archival Information System), with a short introduction to its functional model. The second aspect refers to preservation strategies that give an idea on how to preserve the content and function of archived material while systems are changing over time. These preservation strategies are the heart of every digital archive, because they guarantee the survival of the archived material.

3.1 Open Archival Information System - OAIS

The most important concept for building up and running a digital archive is called "OAIS", which stands for "Open Archival Information System" [OAIS Magenta Book]. OAIS is a reference model (framework) for an archive to provide long term preservation of archived material. OAIS describes an archive as an organization that combines people and systems. It provides an information model and a functional model and it determines responsibilities. OAIS is an ISO standard (ISO 14721) which is well accepted worldwide.

An overview of the OAIS functional model is shown in Figure 2. The functional entities are:

- Ingest: transfer of the archive material from the producer as SIP (submission information package) containing both data and metadata to the archive
- Archival Storage: securely stores material as AIP (archiving information package) and provides data for access and for preservation processes (e.g. migration)
- Data Management: organizes all kinds of metadata and descriptive information regarding the archived material

- Administration: manages the operation of the whole archive
- Preservation Planning: supports all tasks to keep the archived material accessible and usable over the long term, e.g. preservation strategies, technology monitoring and risk analysis
- Access: allows users to retrieve data and relevant information from the archive as DIP (dissemination information package)

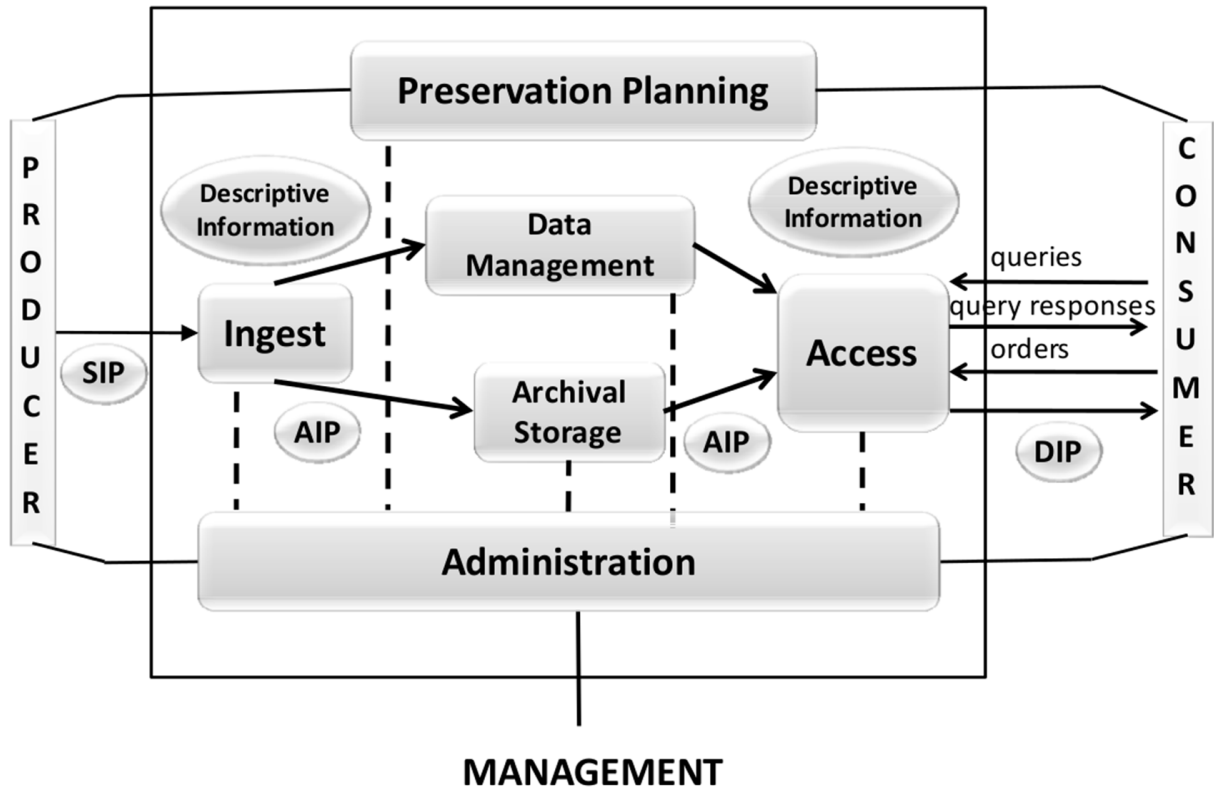


Figure 2. Functional entities of OAIS ([OAIS Magenta Book] page 4-1, figure 4-1)

3.2 Preservation Strategies

Preservation strategies give ideas on how to preserve content and function of archived material while systems are changing. An archive needs an initial preservation strategy for each data type that will be archived. From my point of view, there are four abstract types of preservation strategies that are listed below on Table 1.

The first data type is used for simple visual data. The main aim of this data type is to be perceived (viewed or read) by people. To preserve these data types (and only these) it is sufficient to consider the file format and to comply with approved preservation formats.

The second data type includes functional aspects. The purpose of such data is to be interpreted by machine and to provide some (simple, common) algorithmic and operational capabilities. Preservation of simple functional data is also orientated by approved file formats, but it additionally needs a careful description of the data structure and context. Beyond that, an informative visual representation is recommended.

The third data type involves all kinds of systematic, structured information (e.g. lists of archaeological finds, descriptions, geodata). Such information will be mapped into standardized datasets and stored in permanently operated server databases. The operation of a server database includes regular data migrations, so it can be considered a good archive.

The fourth data type typically contains several files with diverse interconnections and functional and interoperable characteristics. Such data typically needs a specific system environment and specific software to be used (e.g. a combination of CAD-drawing and georeferenced raster images, used in archaeological excavation documentation). In order to preserve such data, it is recommended to build a representation that is abstracted from software-specific overhead. Of course, a careful description is needed and a visual representation should be stored as well. In addition, the original must be kept as long as the required system environment exists.

Table 1. Types of Preservation Strategies

DATA TYPES	EXAMPLES	PRESERVATION STRATEGY	ARCHIVE REPRESENTATIONS
SIMPLE VISUAL DATA	text, image, drawing, ...	depending on file format: use approved preservation formats (like PDF/A, TIFF baseline, SVG, ...)	visual
SIMPLE FUNCTIONAL DATA	spreadsheet, geodata, 3D, ...	use simple, common standard formats (like CSV, XML, GML, DXF, Shape, ...) carefully describe data structure and context create an informative visual representation	functional visual
STRUCTURED DATA	datasets, geodata	use permanently operated server databases (or geodatabases) to store systematic, structured information carefully describe data structure and context	datasets
COMPLEX DATA	interconnected data, interoperable files	build a functional representation that is independent from software-specific overhead carefully describe data structure and context create an informative visual representation keep the original	functional visual original

Archive representations support various future use scenarios. In the context of preservation strategies, we use the following archive representations:

- Visual Representation: for long term use with standard viewing.

- Abstracted Functional Representation: for long-term functional use after individual customizing.
- Original Data: for short term use with an original live system.
- Datasets: for special database use.

To use this concept in practice, more specific preservation strategies are needed. The above-mentioned types have to be specified in detail according to real data. Only the strategy for simple visual data types is nearly ready to use. Be aware that without specific preservation strategies curation will collapse. "Archiving" is differentiated from "storing" by having sufficient preservation strategies (or preservation planning). From this perspective, preservation strategies are the most important aspect of archiving.

4. FIRST EXPERIENCES WITH IMPLEMENTATION

After some years of theoretical discourse, in 2014 the Saxon Archaeological Heritage Office started the project "Archiving Archaeological Research Data" with the aim of implementing its own digital research archive. Endowed with rather small resources, it took some time to go into production with the first archive service for general material (this is simple and unspecific visual material without any institutional policies and workflows). The current work is to configure the second archive service: excavation documentation. This is not so much a technical problem, but rather a problem of preparing data to be suitable for archiving. The following section presents the first experiences with the implementation of this archive.

We began the project with some ambitious ideas about digital archiving. This was preceded by several theoretical discussions about digital archiving (within the Archiving Group of VLA), leading to some reasonable drafts and papers [Göldner et al. 2015]. From this starting point, we wanted to gain practical experience and so we decided to acquire an OAIS compliant archiving system, to implement it with the help of the vendor and to configure it to archaeological material. But soon we reached limits. Contrary to the idea of having an easy to use software system that would archive almost all of our data with reasonable effort, we failed in archiving simple TIFF images. During the next months, we learned much about technical aspects of archiving, like TIFF file tags.

But we also had to revise our theoretical concept. After capturing the structure and requirements of our present excavation documentation we had one big problem: there is quite a lot of digital data that can't be archived by a simple preservation strategy based on file formats. A solution is the concept of preservation strategies that is presented in the previous section.

Initially the project combined the skills of archaeology and information technology. But soon it became clear that archiving skills are very helpful as well, for instance, to understand the archiving system workflows.

Often, I have been asked for some basic advice to start digital archiving. Hence, I try to outline two aspects in the following sections, based on the first experiences and important from my point of view: how to become an archivist (collecting skills and performing first tasks) and to follow OAIS (select reasonable tasks).

4.1 To Become an Archivist

Always start with a clear archiving task. Archive objectives should be discussed with the stakeholders. There should be clear ideas about the archived material, producers and future users. And of course, there should be some basic resources to start archiving.

Archives require competent staff. To build up and run a digital archaeological archive you need three skills, archaeological, archival and IT-related skills. Normally it is not a problem to access archaeological skills at archaeological institutions. But there are also quite a number of tasks for trained archivists, like supporting producers and end users, managing data and metadata, organizing curation and preservation and much more. If there is no professional archivist available, try to get some training. And there are also quite a number of specific IT tasks, e.g. to understand data and data formats, to securely store the archived material and to professionally run migration processes if needed.

Take into account the fact that archaeological and archival aspects interact and that they may be contrary. For example, many archaeologists like to preserve all produced data indefinitely. But they have to discuss with archivists the issues of archival value and the suitability of data for archiving.

To get training in digital archiving, look at global players, such as Digital Preservation Coalition [DPC], Open Preservation Foundation [OPF], the “nestor” network of competence [NESTOR], The National Archives [TNA], National Library of New Zealand [NDHA] or the SLUBArchive Dresden [SLUBArchiv]. Read some guidebooks, such as “Digital Preservation Handbook” [DPC], “Guides to Good Practice” of the Archaeological Data Service [ADS-G2GP], “IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften” of the German Archaeological Institute [IANUS-IT], “nestor Handbuch” [Neuroth 2010] or “Ratgeber zur Archivierung digitaler Daten” [Göldner et al. 2015] and be sure to be inspired by the ideas of Open Archival Information System [OAIS]. Also be aware of the archaeological archives and related projects like ADS, ARCHES, DANS, IANUS and the Archiving Group of VLA.

4.2 To Follow OAIS

The OAIS model is not a simple approach. As presented above, OAIS provides a broad framework with many archiving tasks. So it is necessary to adopt this framework to the needs of the particular archive. According to our experiences there are some hints and warnings regarding this process:

- You cannot buy a ready-made archive. You only can buy some components that help you to operate an archive.
- Do not underestimate the effort to run a real archiving system. Have a look at those who run a real archival system, like ADS or DANS.
- Archiving is the most intensive and most expensive way to store and to preserve data. That is mainly because of the need for preservation planning and continuous curation.
- It is unlikely that you will be able to implement an archive all at once. That would be too complex, so try to build small and simple partitions.

So, my recommendation is to gradually implement the OAIS:

- Apply criteria of archival value to select archive material.
- Reduce the number of data formats and use data types that are as simple as possible.
- Be sure to have an initial preservation strategy for each data type that will be archived.
- Use the available infrastructure as well and intensively as possible (corporate storage, tape archive, virus check, etc.).
- Keep an inventory with descriptive (e.g., subject), technical (e.g., format) and organizational (e.g., access rights) metadata.

Because technology does not solve all archiving problems (neither does OAIS certified software), there is a notable “manual” effort:

- to establish appropriate archive objectives and requirements (remember the aspects of preserving content and function, of archival value and of future use scenarios outlined above in section 2);
- to manage the “OAIS” and to organize all tasks (think about data ingest, maintaining descriptive information, quality checks, storage management, preservation planning, risk analysis, data curation and migration, providing search and access capabilities, daily system operation and so on; see also section 3.1);
- to develop and maintain acceptable preservation strategies (as described in section 3.2; such strategies depend on the type of archival material and have to be specified according to real data);
- and to apply them (concerning migration methods it must be guaranteed and proved that all significant features and functions are preserved with regard to future use scenarios. These operations have to be planned, performed and documented with much care and competent staff is needed for all curation tasks).

5. SMALL STEPS TOWARDS ARCHIVING

As already stated, archiving is the most intensive and most expensive way to store data. But are there no alternatives? This section will address some small steps in the field of digital archiving without having large budgets. Taking such steps will help to reduce the risk of losing important digital information. They will also bridge the gap until proper (certified) preservation tools become available at a reasonable price.

I would like to combine this section with a warning. David Rosenthal said: “Especially in the current difficult funding situation, it is important NOT to give the impression that we can ‘preserve’ digital information with ever-decreasing resources, because then what we will get is ever-decreasing resources. ... I’m sorry, but preserving this stuff costs this much. Less money, no preservation, just storage.” [Rosenthal, 2017]. I endorse Rosenthal’s view; however, I know some institutions that are far

from any adequate funding situations, but which want to take small steps. So, keep in mind that this is not a finished solution but a way to approach the goal.

5.1 Responsibility and Competency

No matter whether you plan a project or a permanent business, start with a clear archiving task that is supported by relevant stakeholders. After that you need someone who is responsible for this task. And, of course, there have to be some basic resources like working time and IT infrastructure or some money to pay for appropriate services. Moreover, you need competencies in archaeology, in archiving and in IT. Think about training and self-study to gain the necessary expertise.

There are many projects and institutions that share their broad experiences of archiving; please see section 4.1 "To become an archivist" to find suggestions. Be sure to have the above-mentioned competencies or at least be sure to know someone with these competencies.

5.2 Archival Objectives

A leading principle in the field of archiving is to permanently preserve (digital) archaeological material with enduring value to science and society. This implies a selection of material. Think about these questions: which records have archival value? Which criteria could be used to prove that? Therefore, minimal quality management is needed. Some simple criteria may be:

- Unique scientific results are worth archiving.
- Documentation of heritage that cannot be preserved is worth archiving.
- Not worth archiving are materials like drafts, duplicates, intermediate states, incomplete, faulty or unfinished material.

Never aim to archive "all" data right from the beginning, because this most probably would lead to a curation collapse due to overwhelming curation efforts.

5.3 Secure Storage

A very basic task is to securely store digital data. Fortunately, storage technologies are now advanced enough, so it is not very complicated to obtain some kind of secure storage, if you consider the following tips.

- Don't use any removable media, because the chance to detect and fix errors is slim.
- Use an available corporate storage, data center or online storage that is managed by professional IT staff. Such systems should use redundancy (e.g. RAID), and backup (e.g. tape backup). To achieve even more security, you may ask for geo redundancy with two or more storage systems at distinct locations, but that is a rather high-level requirement.
- Think about signatures or checksums to check and prove integrity.
- Think about virus checks to prevent malware from being archived.

Within the OAIS reference model these aspects are addressed in the section “Archival Storage Function.”

5.4 Complete Inventory

Another important element is an archive inventory that helps to find and manage data. Without any inventory, searching would be complicated or impossible. Some basic metadata has to be recorded for every entity of archived material; usually it is differentiated into descriptive, technical and organizational information.

First, you need descriptive metadata. See the core elements of the Dublin Core Standard [DC] to get inspired. Think about creator, title and subject as minimum. You may need some additional key attributes to link to important resources outside the archive. Second, technical metadata are required. This includes the original filenames (and possibly a structure map of a set of files and directories), the data format (best to use the PRONOM identifiers) and the dates of creation and modification. Some tools to detect data format and checksum are freely available (e.g. DROID). And third, you need some organizational metadata with access rights, copyrights and retention policies.

To store these metadata, any kind of available document management system would be helpful. But in many cases a structured file system would be sufficient, if file names and directory names include basic meta-information. This could be complemented by some details stored as spreadsheets (e.g. the inventory of a directory).

In regard to the OAIS, the inventory addresses the “Data Management Function” and the “Access Function.”

5.5 Smart Curation

As outlined above, preservation strategies are very important to long term preservation. But how to keep them simple and inexpensive? The best way is to avoid complex data types and to restrict archival material to only a few simple data types (see Table 1 above). You may follow this advice for smart curation:

- Reduce the number of your data formats to reduce curation costs.
- Only use simple data formats, because there are simple preservation strategies for them.
- Understand your file formats as well as possible so that you can detect preservation risks.
- Observe changes in technology regarding the archived data formats.
- Migrate obsolete data to contemporary formats. Remember to test all migration steps and to prove the preservation of content and function. Document all migration steps.

Do not forget the different types of archival representations (see above), as you may benefit from using more than one: original data, visual representation, abstracted functional representation and dataset. For example, think about a MS Access database file; with some tables of data, some data entry forms and some reports for presentation, this file is the original data representation. A set of reports with all important data may be printed or exported to files using PDF/A format, which is a

visual representation. All tables may be exported to files using XML/XSD, completed by a structural description, that is an abstracted functional representation (that does not care about software dependent user interface).

The OAIS preservation strategies are part of the "Preservation Planning Function." A basic set of preservation strategies follows:

1. Use well accepted archive formats for simple files at visual standard formats:
 - store text files at PDF/A format,
 - store image files at uncompressed baseline TIFF format,
 - store drawings at SVG.
2. Use less well accepted archive formats for more special files at functional standard formats, add a visual representation and a clear description of the data structure
 - store spreadsheets at CSV or XML format using UTF-8 encoding,
 - store CAD data at DXF format and add characteristic plans at PDF/A or TIFF format (and maybe store the original),
 - store geodata as Shapefile or GML and add characteristic maps at PDF/A or TIFF format (and maybe store the original).
3. Put systematic information into systematic data structures (datasets, geodata) and use permanently operated server databases to store systematic, structured information.

The idea of preserving digital data without large budgets is not new. My first ideas were presented at CHNT 14 [Göldner 2010], not so far from the revision presented here. Another interesting approach is "Parsimonious Preservation" [Gollins 2013, Gollins 2009], that argues against overestimation of risks like storage disaster and format obsolescence. This is very helpful, but carefully read the chapter "Something to Note" at the end of Gollins 2013. Gollins' ideas are about a "foreseeable future (5 to 10 years)" and do not include any preservation strategies. That is not sufficient from my point of view, because problems are delayed but not solved. My clear recommendation is to have appropriate preservation strategies from the beginning.

The operation of archive systems is costly, even if you try to save money as mentioned above. Thus it may be interesting to regard alternatives: use shared or cloud systems and look for service providers for system parts like storage or for a whole digital archive. In this case pay attention to certifications like the "Data Seal of Approval" [DSA], "Trustworthy Repositories Audit & Certification" [TRAC] and "nestor-Siegel für vertrauenswürdige digitale Langzeitarchive" [nestorSeal] and in addition pay attention to an existing exit strategy (to preserve your data beyond the life of the service provider). The hope remains that professional and subject specific (or adjustable) archive services will come into being in the near future, so that smaller institutions could benefit from this development.

6. RÉSUMÉ

Considering requirements, objectives and first experiences in the field of archiving archaeological research data, I would like to put a warning first: There are three critical ways to cause a "Curation Collapse" at the archive:

- Ignore archival value.
- Accept unsuitable material.
- Operate without a preservation strategy.

On the other hand there is the OAIS reference model, a basic framework for long term preservation. There is a lot of expert knowledge and experience and there is even OAIS compliant software available now. If you intend to start digital archiving without a big budget, remember these useful and small steps towards archiving:

- set responsibility and gain competency;
- establish archival objectives;
- care about secure storage;
- build a complete inventory;
- develop preservation strategies;

Perhaps that is not a perfect archive, but these steps prevent the occurrence of the main preservation risks with reasonable effort. These steps are also a very good preparation for migrating available data to a professional archive system.

So, start now, find someone responsible, build up an archive step by step and preserve your digital assets for the long term.

7. REFERENCES

ADS. Archaeology Data Service. Homepage. Retrieved January 2017:

<http://archaeologydataservice.ac.uk/>

ADS-G2GP. Guides to Good Practice. Homepage. Retrieved January 2017:

<http://guides.archaeologydataservice.ac.uk/g2gp/Main>

ARCHES. Archaeological Resources in Cultural Heritage: a European Standard. Homepage.

Retrieved January 2017:

<http://archaeologydataservice.ac.uk/arches/Wiki.jsp?page=The%20Standard%20and%20Guide%20to%20Best%20Practice%20in%20Archaeological%20Archiving%20in%20Europe>

Archiving Group of VLA. Kommission Archäologie und Informationssysteme - AG Archivierung.

Homepage. Retrieved January 2017:

<http://www.landesarchaeologen.de/verband/kommissionen/archaeologie-und-informationssysteme/projektarbeitsgruppen/archivierung/>

DANS. Data Archiving and Networked Services. Homepage. Retrieved January 2017:

- <https://dans.knaw.nl/en>
- DC. Dublin Core Metadata Element Set. Retrieved January 2017:
<http://www.dublincore.org/documents/dces/>
- DPC. Digital Preservation Coalition. Homepage. Retrieved January 2017: <http://www.dpconline.org>
- DPH. Digital Preservation Handbook. Digital Preservation Coalition. Retrieved January 2017:
<http://dpconline.org/handbook>
- DROID. Digital Record Object Identification. The National Archives. Retrieved January 2017:
<https://www.nationalarchives.gov.uk/information-management/projects-and-work/droid.htm>
- DSA. Data Seal of Approval. Homepage. Retrieved June 2017: <https://www.datasealofapproval.org>
- R. Göldner. 2016. *Ein Erfahrungsaustausch zur Archivierung digitaler archäologischer Daten – Zusammenfassung 2015/16*. Retrieved January 2017:
http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Archaeologie-Informationssysteme/Dokumente_AIS_Archivierung/Erfahrungsaustausch_Archivierung_2015-11_Zusammenfassung.pdf
- R. Göldner et al. 2015. *Ratgeber zur Archivierung digitaler Daten*. Retrieved January 2017:
http://www.landesarchaeologen.de/fileadmin/Dokumente/Dokumente_Kommissionen/Dokumente_Archaeologie-Informationssysteme/Dokumente_AIS_Archivierung/Ratgeber_Archivierung_V1.0.pdf
- R. Göldner. 2010. Preserving Digital Data Step by Step – Small Solutions for Small Institutions. In: *Proceedings of the 14th International Congress "Cultural Heritage and New Technologies"*. Wien: eBook Edition, 2010. Retrieved January 2017:
http://www.stadtarchaeologie.at/wp-content/uploads/eBook_WS14_Part3_AT_1.pdf (pp. 117-125)
- T. Gollins. 2013. *Putting Parsimonious Preservation into Practice*. Retrieved January 2017:
<https://www.nationalarchives.gov.uk/documents/information-management/parsimonious-preservation-in-practice.pdf>
- T. Gollins. 2009. *Parsimonious Preservation: Preventing Pointless Processes*. The National Archives. December 2009. Retrieved January 2017:
<http://www.nationalarchives.gov.uk/documents/information-management/parsimonious-preservation.pdf>
- IANUS. Forschungsdatenzentrum Archäologie & Altertumswissenschaften. DAI. Homepage. Retrieved January 2017:
<http://www.ianus-fdz.de/>
- IANUS-IT. IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften. DAI. Retrieved January 2017: <http://www.ianus-fdz.de/it-empfehlungen/>
- NDHA. National Digital Heritage Archive. National Library of New Zealand – Digital Preservation Programme. Homepage. Retrieved January 2017: <http://digitalpreservation.natlib.govt.nz/about-2/programme-overview/>
- NESTOR. Kompetenznetzwerk zur Langzeitarchivierung. Homepage. Retrieved January 2017:
<http://www.langzeitarchivierung.de>
- nestorSeal. nestor-Siegel für vertrauenswürdige digitale Langzeitarchive. Online. Retrieved June 2017: http://www.langzeitarchivierung.de/Subsites/nestor/DE/Siegel/siegel_node.html

- H. Neuroth et al. 2010. nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung. Online. Retrieved June 2017: http://www.nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf
- OAIS. Open Archival Information System. Wikipedia article. Retrieved January 2017: https://en.wikipedia.org/wiki/Open_Archival_Information_System
- OAIS Magenta Book. Open Archival Information System. CCSDS Washington, 2012, Retrieved January 2017
<http://public.ccsds.org/publications/archive/650x0m2.pdf>
- OPF. Open Preservation Foundation. Homepage. Retrieved January 2017:
<http://openpreservation.org>
- PRONOM. Public File Format Registry. The National Archives. Retrieved January 2017:
<https://www.nationalarchives.gov.uk/PRONOM>
- D. Rosenthal. 2017. Threats to stored data. Homepage. Retrieved June 2017:
<http://blog.dshr.org/2017/03/threats-to-stored-data.html>
- SLUBArchiv. Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden – Archiv. Homepage. Retrieved January 2017:
<http://www.slub-dresden.de/ueber-uns/slubarchiv/>
- TNA. The National Archives. Homepage. Retrieved January 2017:
<http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/>
- TRAC. Trustworthy Repositories Audit & Certification. Homepage. Retrieved June 2017:
<https://public.ccsds.org/publications/archive/652x0m1.pdf>