Archiving Digital Archaeological Data - Evaluation of a Survey in Germany

Reiner Göldner, David Bibby and Henriette Senst

'Archaeological archiving is a dynamic process, which begins the moment a project is planned.' EAC 2014

Archiving of digital research data according to established standards and methods is an important part of the research data lifecycle. It guarantees the long-term transparency and reusability of research results. This is especially important in the case of archaeological data where the preservation of unique and irreplaceable cultural heritage is paramount. To gain insight into the current state of affairs regarding long-term archiving in Germany, the SEADDA (Saving European Archaeology from the Digital Dark Age) COST Action and the Commission on Archaeology and Information Systems in the Association of German State Archaeologists, with the support of the IANUS Research Data Center of the DAI, conducted a survey in spring/summer 2021. This article presents the results of the survey along with recommendations for further and supporting measures aimed at improving the long-term archiving of digital archaeological data in Germany.

1. Introduction

Archiving digital research data is an essential component of 'good scientific practice' and, in archaeology, it serves to preserve a unique and irretrievable cultural heritage. In 2018, the COST Action SEADDA (Saving European Archaeology from the Digital Dark Age) was established to foster knowledge exchange on best practices for preservation, dissemination, and reuse of archaeological data in Europe. SEADDA Working Group 1 'Stewardship of Archaeological Data' attempted to assemble an overview of the current state of the art. Unfortunately, the German situation is fragmented because of the cultural sovereignty of German states. To determine the current state of affairs, SEADDA and the VLAK-AIS conducted a survey in the spring/summer of 2021. While the survey results and the recommendations arising from them are specific to Germany (see Göldner et al 2022; and Umfrage zur

[Archivierung digitaler archäologischer Daten](#)), the authors believe the problems to be universal. We therefore offer this contribution in a more internationally accessible English language version. It is intended to be of interest to everyone concerned with digital archaeological archiving.

The survey received 30 valid response sets, offering an adequate picture of a considerable sector of archaeology in Germany. One-third of the response sets can be attributed to state and local archaeological agencies (affiliated with the [Association of German State Archaeologists](#)) and another third to the academic sector (universities, colleges and research institutes), with the final third unattributable to specific organisations.

The survey results confirm an awareness of archiving and a motivation to preserve digital archaeological data for future generations and subsequent research because these data often contain unique and irreplaceable information. However, the responses also highlight the need to implement more adequate archiving processes, e.g.

- 75% are without an archiving policy,
- 75% are unfamiliar with the relevant OAIS standard,
- > 50% are not making a daily backup, and
- > 50% are not aware of the FAIR principles.

Some massive obstacles are also mentioned, e.g.

- insufficient consideration within work and research processes,
- insufficient expertise,
- a considerable lack of resources.

A self-assessment results in:

- 20% low-level awareness (aware of basic principles but need to act) and
- < 20% higher level (archiving processes that go beyond that awareness).

Given the state of the art in digital archiving, the authors present some recommendations for further and supporting measures. The survey shows that awareness and the importance of digital archaeological archiving must be more clearly and even assertively promoted. Data management skills must be improved. Awareness of standards such as OAIS and FAIR and measures for their implementation must be raised. A further focus must be placed on the resources available for an adequate and sustainable implementation of digital archaeological archiving. The survey results clearly indicate a need for both human and financial resources. Making use of pre-existing facilities could provide a cost-effective solution, but the subject is complex. Qualified guidance would be useful for decision-making processes, but this is not yet available. Digital archaeological archiving demands specialised knowledge, which can be acquired through training, further education, and via the sharing of practical experience. While OAIS is highly relevant, metadata standards, legal practice and the FAIR principles are also important. A countrywide 'competence centre' could also save resources in terms of specialised knowledge.

2. The Survey

A growing awareness of the importance of digital data archiving is indicated by a number of national and international projects, such as IANUS, NFDI and SEADDA.

The National Research Data Infrastructure project (NFDI) supported by the German federal government aims to ensure the broad and sustainable availability and reuse of research data across all areas of science and the humanities. This is achieved through professional archiving and provision following the FAIR principles (Wilkinson et al. 2016). The NFDI4Objects consortium plans to provide researchers working on the material heritage of human history (including archaeologists) with methods, standards and interfaces for long-term archiving. The results of the survey reported here are intended to inform NFDI4Objects on the present situation on the ground, and provide it with useful information for its own strategy and policy-making decisions.

In recent years, the SEADDA COST action, with 35 European and 4 international partner organisations, has played an important role as a platform for the international exchange of best practice for digital archaeological archiving. SEADDA's mission is to improve international knowledge exchange about innovations in the field, to promote best practices in the preservation, dissemination and reuse of archaeological data, and to support and establish the topic as a European priority by developing a common understanding of digital archaeological data management and building professional networks.

The specialist commission 'Archaeology and Information Systems' of the Association of German State Archaeologists (VLA) engages in promoting digital archaeological archiving in Germany and also deals with this topic. VLA played a significant role in designing and analysing this survey, which took place seven years after the last questionnaire.

The following general instructions were given for completing the questionnaire:

- The term >archiving should be understood to mean permanent preservation of usability for an unlimited period of time.
- The questions relate to the archiving of digital archaeological data within the respective responding institutions.
- The focus is on archive-worthy research data (e.g. project reports, excavation documentation, digital images, geospatial and surveying data, 3D scans etc.), but not administrative documents.

The questionnaire was distributed as follows

- via the VLA Commission 'Archaeology and Information Systems' to all archaeological state offices in Germany,
- via the NFDI4Objects distribution list to all consortium stakeholders,
- to the German Chapter of CAA (Computer Applications and Quantitative Methods in Archaeology).

The distribution thus covered all major players working with digital archaeological data in Germany.

An important feature of the survey was confidentiality. Some aspects were judged to be (very) sensitive, so it was deemed that the anonymity of the participants would result in more realistic answers. However, as a consequence, there was no opportunity to clarify individual points with specific participants or to publish answer details that could potentially identify a participant.

A total of 30 valid responses were received (meaning that they included one or more answers) and these represent a wide cross-section of those active in German archaeology. One response seemed to originate from an Austrian University but it was decided to keep it in the evaluation. These responses comprise the 100% basis for the percentage evaluation of the survey.

3. Evaluation of the Questionnaire

This detailed evaluation is based on the structure of the questionnaire and the questions asked. For most questions we include:

- A Premise: a short explanation about the background of the question (representing the view of the authors),
- A Summary: a compact summary of the answers (as simple count and percentage of the 30 valid answer sets in case of single or multiple-choice answers, or carefully summarised and compressed in case of free text answers),
- A Recommendation: suggestions for improving the situation indicated by answers (representing the view of the authors).

3.1. General aspects

1. **To estimate the representativeness of the survey, we would like to know what type of institution your answers refer to?**

   Summary:

   About two-thirds of all relevant answers can be narrowed down institutionally. Ten answers (33%) can be assigned to the field of state and municipal archaeology (State heritage departments affiliated with the Association of German State Archaeologists). Ten (33%) responses can be attributed to the academic sector (universities, colleges and research institutions), with the final 33% representing a variety of other bodies.

   Though the response to the survey was not overwhelming, the anonymous participation of a high proportion of German state and municipal archaeological offices and the academic sector provides an adequate picture of a considerable part of archaeology in Germany (or in the German-speaking world).

2. What is your institution's internal (or overarching) policy on archiving digital archaeological data (data policy)?

Premise:

A regular policy for archiving digital archaeological data is required for proper archiving.

Summary:

Only seven answers (23%) can be interpreted to indicate that a policy is in place; a further response (3%) referred to 'recommendations'. For 74% of the responses, it must be assumed that no appropriate policy is known or applied.

Recommendation:

Examples of such policies can be found in DANS, ADS, DAI, and relating to research data in general, for example, at FDMentor, the Leibnitz Institute or the Technical University Berlin. Support should be offered for the creation (and reuse) of institutional policies for archiving digital archaeological data.

3. Which body is responsible for in-house digital archiving?

Premise:

A central and permanently established department (or organisational unit) should be responsible for archiving.

Summary:

Ten answers (33%) indicate this to be the case. A further four responses (13%) name individuals who are perhaps only employed on a project-related/temporary basis. In 54% of the responses, no responsible body/individual is indicated.

Recommendation:

The archiving of digital archaeological data should be firmly (permanently) anchored within an organisation (business allocation plan, competence profiles for job descriptions, …). Successful models should be made available.

4. **What qualifications do the workers involved in digital archiving have?**

Training with a focus on digital archiving (OAIS) is considered an appropriate qualification.

Summary:

There were only two such cases (7%). In 21 other cases (70%), reference is made to further training, interest and knowledge of local guidelines; 23% of the responses were blank. There is a clearly identifiable shortage of competent and skilled workers.

Recommendation:

To increase qualifications, appropriate training and further education should be offered, especially orientated toward OAIS but not forgetting other aspects of research data management such as FAIR (see questions 31-33) and open science. In addition, corresponding competence profiles (cf. question 3) should be considered when distributing archival tasks. Also, positions in digital archaeological archiving must be made more attractive than they are at present.

5. **How long should digital data be usable in your institution?**

Premise:

Archiving means *permanent* usability without a time limit, represented here by the statement 'at least 100 years'.

Summary:

This was confirmed in seven responses (23%). In the academic context, shorter 'retention periods' of, for example, ten years are often applied, but strictly speaking (in the OAIS sense) this is not considered to be archiving. Even so, four responses (13%) considered ten years to be adequate. 60% of the answers were blank.

Recommendation:

Attention must be paid to an archaeologically uniform understanding of the term 'archiving'. Guidance can be found, for example, in the [VLA Guide to Archiving](#) or the [EAC Guidelines for Archaeological Archiving in Europe](#).

6. **Are you familiar with the OAIS (Open Archival Information System) standard?**

Premise:

OAIS (Reference model/Referenzmodell) is considered the most important standard for digital archiving. Knowledge of it is mandatory. Understanding, skills and abilities are not what is asked here, only *knowledge* of OAIS.

Summary:

Only seven responses (23%) confirmed at least a basic knowledge of OAIS. It must therefore be assumed that 23 cases (77%) do not have a basic understanding of the most important standard for digital archiving (OAIS).

Recommendation:

To increase qualifications, appropriate training and further education should be offered/located, orientated in particular towards the OAIS (cf. question 4).

---

7. **Which areas of the OAIS have been implemented sustainably in your institution?**

Premise:

For the operation of a digital archive according to OAIS, all six areas must be implemented appropriately and sustainably:

- Ingest (data transfer),
- Storage (archive storage),
- Management (data management),
- Access,
- Preservation planning,
- Administration (system administration).

Summary:

Only two institutions (7%) confirm this implementation. Another answer (3%) only mentions 'ingest (data transfer)'. 90% of the answers cannot confirm an implementation according to OAIS.

Recommendation:

The importance of the OAIS standard for digital archiving should be clearly emphasised.

---

8. What archiving software is used to support the tasks?

Premise:

Specialised archiving software could/should be used to support the archiving task.

Summary:

Responses mention 'Rosetta' (OAIS-compliant), d.3 (document management), Fedora (repository) and database-based specialised procedures.

Recommendation:

Successful practice (including the software used) in existing digital archives should be described, collected and made available for reuse. In general, the use of free and open-source software (FOSS) and its orientation towards OAIS should be recommended; individual solutions/systems should be avoided.

---

9. What annual resources are directly available for digital archaeological archiving (without recording) in your institution?

Premise:

Adequate resources should be made available for the archiving task. The intention was to get an idea of the volume of resources that are used for archiving tasks in practice.
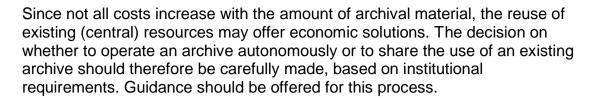
Summary:

Available are:

- Full-time positions: 7 responses: 0, 0, 0, ¼, ¼, ½, 2 positions,
- Material resources: 6 responses: 0, 0, 0, 1 500 20 000, 60 000 €,
- Number of Servers: 5 responses: 0, 1, 1, 3, 'as needed',
- Archive storage: 10 responses: 0, 4, 5, 10, 22, 35, 60, 60, 100 TB, 'scalable as required',
- (Yearly) Increase: 7 responses: 0, 0, 1, 1, 1, 4, 10 TB.

Recommendation:

The answers give only rough indications. More tailored questions would have to be asked to establish more details. Thus the following recommendation is not only derived from the answers given but is of general relevance.

For digital archaeological archiving, adequate human and financial resources must be guaranteed, based on the quantity and types of archival material.

Since not all costs increase with the amount of archival material, the reuse of existing (central) resources may offer economic solutions. The decision on whether to operate an archive autonomously or to share the use of an existing archive should therefore be carefully made, based on institutional requirements. Guidance should be offered for this process.

3.2. Motivation

10.   What is the legal basis in your federal state for digital archaeological archiving?

Premise:

The motivation for archiving is often primarily based on a legal obligation/state laws.

Summary:

(State) monument protection laws and/or archiving laws are mentioned in five responses.

Recommendation:

Further training on the legal framework of digital archiving is recommended. In doing so, both overarching principles and the legal situation in each federal state must be taken into account.

---

11.   What is the internal motivation for digital archiving?

The purpose of this question was to find out what motivation exists in addition to the legal obligation. Summary:

The following information is highlighted in this regard

- o   Data should be usable by future generations.
- o   Data should be preserved for subsequent research.
- o   Data are required for monument protection.
- o   The functionality of the original data should also be guaranteed in the future.
- o   Data contain irreplaceable information documenting on the no longer existent original situation that was destroyed by excavation.

Motivation is consistently present but contrasts with the implementation status. Obstacles are clearly named (question 12) and the implementation level is predominantly described as 'very low' (question 34).

Recommendation:

To improve implementation, digital archiving must be understood as a common management task (see also question 13).

12.   Which aspects may prevent proper digital archiving?

Premise:

The purpose of this question was to find out what problems archiving encounters or why it has not yet been adequately implemented.

Summary:

The following information is highlighted in this regard

- o   Archiving is not yet recognised/assessed as relevant.
- o   Task and responsibility unclear (business distribution plan).
- o   Insufficient human and financial resources.
- o   Insufficient expertise.
- o   Lack of support from central IT service providers.
- o   Lack of interest, as not archaeological.

Recommendation:

Efforts should be made to remove the obstacles indicated. The topics 'organisation and management' (see also questions 3 and 13), 'qualification' (see also question 4) and 'resources' (see also question 9) again play a major role and should therefore be promoted.

13.   How present is the topic of digital archaeological archiving in the management of your institution?

Premise:

The positioning of digital archiving within the institution's business processes is a management task.

Summary:

Ten answers (33%) state that this topic is present in management (very/relatively strongly etc.). Three answers (10%) state that this topic is not very present in management, and the others (57%) make no statement.

Recommendation:

Like 'digitisation' itself ([Roadmap for a digital transformation](#)), ([Is the right Group Leading Your Digital Initiatives?](#)) digital archiving is a clear management task, because starting with the first creation of digital data, the business processes of different areas must be integrated into the archiving. This is also underlined by the [requirements of research data management](#). Management support in this regard through motivation and advice is essential, applying the principles of digitisation to archiving.

---

14. What do you think are the reasons for the lack of presence of this topic?

Premise:

We intended to seek qualitative information but during the evaluation of the survey, the leading nature of this question was recognised in hindsight. This should be taken into account. Nevertheless, the answers do not contradict the general trend (see questions 12 or 35).

Summary:

It could be that a substantial lack of resources and the complexity of the topic are reasons for the absence of this topic.

Recommendation:

The issue of 'resources' gains further weight here.

3.3. Data transfer

15. Is archiving done in-house or by an external service provider?

Premise:

The purpose of this question was to learn more about where archiving is currently carried out.

Summary:

In ten cases (33%) archiving is carried out by the institution questioned, and in two cases (7%) by an external service provider. One institution is a service provider itself.

Recommendation:

As already indicated in question 9, the reuse of central resources from external service providers promises economic archive solutions. The answers here point to a rather fragmented situation. Potential guidance should take institutions' typical requirements for archiving their digital material into account.

16. What measures are taken when transferring digital material to the archive (ingest)?

Premise:

The transfer of digital archival material is a critical process; it is here that the quality of the archive is initially decided. Important steps in this process are adequate and appropriate metadata and critical judgement of both the archival value of the content of the material as well as the technical achievability of the formats, based on subject-specific guidelines.

Summary:

The archival value of the content is only tested in two cases (7%). Technical archivability is only checked in three cases (10%). In five cases (17%) the archival material is listed in detail. Some (vague) reference is made to both quality management and data transfer agreements.

The statement 'Everything is archived that is offered to the archive' is given in five responses (17%).

Recommendation:

A common understanding should be developed on how digital archaeological archive material is transferred so that it can be archived professionally. Proposals for transfer criteria or a collection strategy in the field would be helpful. The categories of archival value (lasting content value for society) and technical archivability (optimistic prognosis for long-term preservation) are important here. Material is already available, e.g. DANS, ADS, IANUS and at the VLA and should be actively promoted.

Archiving everything that is offered must be viewed critically because it ignores the high cost of permanently archiving material that may not merit this process and ignores the risk of failure of subsequent archiving tasks should it not prove technically possible to do so. The acquisition of archival material should be coupled with quality management. Criteria in this regard should be discussed and agreed upon within the archaeological community and made available centrally for subsequent use.

17.   What preservation strategies do you use for digital archival material?

Premise:

The term 'content-related curation', may need clarification. What is meant here is the supervision of migration processes to preserve the subject content of the archival material as the creator/author intended it. Technically significant criteria serve this purpose. For example: the connection between an archaeological find, its representation in photos, plans etc. and references to data about it in lists, databases, geodata etc. must be preserved at all costs. Or, put another way: 'The entire archive must be compiled in a way that preserves relationships between each element and facilitates access to all parts in the future.'

During the evaluation, there was a clear consensus on the need for archivable data formats and regular format-related curation. However, there was no clear consensus on regular content-related curation.

Summary:

The answers are as follows:

- o   3 (10%) only archivable data formats (e.g. PDF/A) are used.
- o   3 (10%) regular format-related curation (conversion of obsolete formats) takes place.
- o   0 (0%) regular content-based curation takes place

Recommendation:

Curation of archival material can be understood as supervision of migration processes and counteracts risks that may limit or prevent future use of data (Mayerinik 2020). To reach a professional consensus, the characteristic of curation and especially the importance and necessity of 'content-based curation' as described above should be widely and repeatedly discussed (see also question 22).

18.   What kind of subject data do you digitally archive (content, not data format)?

Premise:

The answers give a rough overview of the content of archive material.

Summary:

Data are archived in various combinations, i.e. excavation photos, written descriptions (plans/sections, finds/samples, etc.), 3D scans (raw data, objects etc.), and more. Remote sensing data, building surveys, raw survey data, scans of historical inventory books and other materials are also mentioned. Both 'born digital' and digitised analogue originals are present.

3.4. Metadata

19. What metadata on the archival material are collected?

Premise:

Metadata are important to identify and find materials. Technical metadata are needed to perform archiving tasks, so it would be interesting to see how far this is respected in practice.

Summary:

- o 10 (33%), overview information about the fieldwork.
- o 6 (20%), table of contents of the digital documentation.
- o 5 (17%), detailed metadata for individual components.
- o 1 (3%), none.

---

20. For which components are detailed metadata collected?

Premise:

The answers give a rough overview of the focus of collected metadata.

Summary:

Metadata are collected in various combinations, i.e. especially for excavation photos, but also for written descriptions (plans/sections, finds/samples, etc.), 3D scans (raw data, objects), and more.

---

21. Which metadata standards are used?

Premise:

Metadata is more valuable (findable, interchangeable) if it is standardised.

Summary:

- o 6 (20%), Dublin Core.
- o 6 (20%), EXIF.
- o 3 (10%), IPTC.
- o 2 (6%), METS.
- o 2 (6%), DataCite.
- o 1 (3%), ISO 19115
- o 2 (6%), none.

Recommendation:

Awareness of the importance of metadata standards for archiving should be promoted through continuous discussion and further training.

3.5. Archive material types

The categorisation required in this section was taken from the topic sheet on archiving digital data 'Erhaltungsstrategien 1 – Grundlagen' ('Preservation Strategies 1 - Fundamentals') issued by the [VLA Commission on Archaeology and Information Systems](), in which increasing demands on preservation strategies and thus increasing expenditure are derived from the increasing complexity of the data. As a short summary, four types of materials are distinguished:

- Simple 'visual' material: mainly used for viewing and reading (e.g. texts, images, drawings),
- Simple 'functional' materials: used like visual materials but with the addition of some aspects in the computer background (e.g. spreadsheets, tables, geodata),
- Systematic information: with clear, formal and repeatable data structures using standard data types and standard data interconnections, that are represented in datasets and are systematically queryable (e.g. databases),
- Complex materials: comprise collections of interoperable data that together outclass the above-mentioned demands (e.g. CAD plans with integrated photogrammetric images, GIS map documents with multiple layers combining several internal and external data sources, structured file systems with multiple inter-dependent files of different types, excavation documentation).

The well-known and often used (only) format-based preservation strategy is helpful in the case of simple visual and functional material but, in particular for complex materials, individual and complex preservation strategies are needed.

22. **What types of archival material accrue for digital archiving?**

Premise:

The question was structured to cover the four material types mentioned above and provided a multiple choice of some very typical materials of each type. (From the answers may be derived a type of preservation strategy.)

Summary:

All the types of archival material named in the answers are mentioned with comparable frequency. In addition to the simple types, complex types of material are also present, for which complex preservation strategies are also required.

Recommendation:

During the evaluation, it became clear that the question of preservation strategies for complex types of archival material and, related to this, the necessity of content-related curation (see question 17) has, alongside purely format-related curation, so far not been sufficiently discussed.

Specialised knowledge is required in particular for the realisation of the more complex conservation strategies. This specialised knowledge can be provided more economically in competence centres than would be possible on a decentralised basis.

---

23. What formats are used for simple 'visual' materials (texts, photos, drawings)?

Premise:

The intention of this and the next question was to learn more about which file formats are used for capturing, archiving and reusing archaeological data. The answers give an overview.

Summary:

Typical formats for simple visual materials are: CSV, DOCX, JPG, MP4, ODT, PDF, PDF/A, PNG, TIF, TXT etc. as well as RAW for photos. It is not apparent that different formats are used for capture, archiving (archive storage) and reuse.

Recommendation:

A centrally maintained register with recommendations on archivable data formats for archaeological archive material would be helpful and could summarise, consolidate, and replace local recommendations.

The separate consideration of data formats for acquisition, archiving and subsequent use has the advantage of being able to use an optimally suitable format in each case (e.g. a shapefile for the offline acquisition of geodata, a server geodatabase with simple features for archiving and a WMS/WFS for subsequent use). However, for such a system to function, a procedure fixed in advance is a prerequisite. This possibility should certainly be communicated and made known. Here too, specialised knowledge of the advantages and disadvantages of different formats for different tasks is required.

24. Which formats are used for simple 'functional' materials (spreadsheets, shapefiles etc.)?

Summary:

Typical formats for simple functional materials are: 3D PDF, CSV, DWG/DXF, Geopackage, KML, MS-Access, OBJ, ODS, PDF, PLY, Shapefile, TSV, XLS/X, XYZ. It is not clear that different formats are used for capture, archiving (archive storage) and subsequent use.

Recommendation:

The recommendations suggested for question 23 also apply here.

25. What software is used for simple visual materials?

Premise:

The intention of this and the next question was to learn more about which software is used for capturing, archiving and reusing archaeological data. The answers give an overview.

Summary:

The software specified for simple visual materials is: Adobe, Gimp, Irfanview, Libre-Office, MS-Office, MS-Word and Photoshop. A distinction between capture and after-use is not evident.

26. What software is offered for simple functional materials?

Summary:

The software specified for simple functional materials includes: Agisoft Metashape, Aicon (hardware-related), ArcGIS, Artec, AutoCAD, Cloudcompare, Cyclone, Geomagic Studio, Gigamesh, Konica Minolta (hardware-related), Meshlab, MS-Excel, MS-Office, QGIS, Survey2GIS, TroveSketch and Vessel Reconstructor (TU Chemnitz). A distinction between acquisition and after-use is not evident.

Recommendation:

The software used is closely related to the data formats used. FOSS (Free and Open Source Software) supports the existing open standards well in this

respect. This relationship should be investigated more closely for archaeology-specific software and result in corresponding recommendations to be made available centrally to the specialist community.

3.6. Archive storage

27.   Where is the digital archive material stored?

Premise:

To meet the basic requirements for archive storage according to OAIS, archiving must be operated internally or externally on a central server.

Summary:

For 16 responses this is the case (53%). Local PCs, external hard drives, USB memory and CDs/DVDs are not suitable as permanent archive storage (8 stated, 27%).

Recommendation:

Information on suitable archival storage and the advantages of centralised solutions should be provided.

---

28.   How is the archival material secured against loss and falsification (data integrity)?

Premise:

Data backups are part of regular IT operations and are a matter of course for archiving. To protect against unintentional changes during archiving and to guarantee integrity and authenticity, procedures such as bit-stream preservation and electronic signatures (certificates) must be used.

Summary:

This is reported by one response (3%). Bit-stream preservation is used in three cases (10%). Daily backups are reported 12 times (40%) (but this is insufficient for archiving).

Recommendation:

Information should be provided on suitable procedures to secure archival material against loss and falsification. The difference between data protection and archiving must be made clear.

3.7. Conservation planning

29. How are the data prepared for deposition in the archive?

Premise:

Appropriate metadata are a prerequisite in preparation for deposition in the archive. If necessary, the data must also be reformatted to become archivable.

Summary:

In the responses, the following important preparations were named:

- Preparation and collection of metadata.
- Quality control.
- Selection according to archival value.
- Removal of duplicates.
- Conversion to archive formats (originals are cancelled).
- Check of standardised data structures.
- Form checksums.

Recommendation:

'Archaeological archiving is a dynamic process from the beginning of project planning'. The preparation of data for inclusion in the archive is the last and most important step that data creators must take. Production and archiving must be considered as an integrated process. Because the data creators are generally not archivists, special support is needed, starting at the beginning of data acquisition. In this respect, archaeology-related recommendations and best practices should be centrally available.

---

30. What archiving methods are used in your institution for digital data?

Premise:

Hardcopy (printing on media with a long shelf life), hardware preservation (keeping systems running as long as possible), migration (transforming from older to newer formats and systems regarding significant features) and software emulation (re-creating old computer systems on current systems) are regarded as principal archiving methods.

The answer option 'the data is preserved as it is' should initially indicate an inadequate method, but this terminology was found to be misleading because it can be interpreted in different ways. If applied to the migration method, the

statement would be insufficient, whereas it would be appropriate for the hardcopy method.

Summary:

- o   10 (33%) 'the data is preserved as it is' (please note the hint above),
- o   4 (13%) hardcopy,
- o   5 (17%) hardware preservation,
- o   2 (7%) migration,
- o   4 (13%) software emulation.

Besides migration also hardcopy, hardware preservation and software emulation were mentioned in the answers.

Recommendation:

The discussion of archiving methods should also include experiences with hardcopies, hardware preservation and software emulation as they are actually used, beyond any OAIS standards. This also includes weighing up the advantages and disadvantages of the various methods. Advice on this can be found, for example, in the VLA's 'Ratgeber Archivierung' ('Guide to Archiving').

3.8. FAIR data

31.   Are you familiar with the FAIR (Findable, Accessible, Interoperable, Reusable) principle?

Premise:

The FAIR principles (Wilkinson et al. 2016) are not part of OAIS but are increasingly used in the academic sector (universities, colleges and research institutions) as research data management requirements.

Summary:

Eleven responses (37%) indicate knowledge. It remains an open question whether 7% 'no' answers or 63% (incl. 56% non-answers) indicate a lack of knowledge of FAIR.

Recommendation:

The FAIR principles should be publicised and promoted more widely! The 'FAIR Data Maturity Model' should be used as a basis (FAIR Data Maturity Model Working Group 2020).

32.   **How are the FAIR principles taken into account?**

Premise:

The following subordinate questions are taken from the requirements of FAIR. (FAIR is described in far greater detail in the [FAIR Data Maturity Model](#)). The following answers provide a general overview of how intensively FAIR is implemented in the respective institutions.

Summary:

- How are the metadata and data from the digital archive publicly discoverable?
  - 1 (3%) via the Internet
  - 1 (3%) internal only
  - 2 (6%) not yet (planned)
- How does one get access to archival data of interest?
  - 1 (3%) open access
  - 4 (13%) on request
- How are interoperable (standardised, machine-readable) (meta-)data available?
  - 1 (3%) via the internet
  - 2 (6%) planned
- What vocabularies are used that are also FAIR?
  - 2 (6%) Getty AAT
  - 1 (3%) ÖFOS
  - 1 (3%) WNK
- How are the provenance and the professional quality of the (meta-)data described?
  - 1 (3%) Verbal
- What types of after-use are possible?
  - 1 (3%) CC BY 4.0 Licence
  - 4 (13%) case-dependent, licence-dependent, ad hoc, on request

Evaluation of the six questions per dataset:

- 1 (3%) Answered all questions with a yes
- 1 (3%) Answered all questions with yes, but with contradictions in the explanatory text (e.g. 'not at all').
- 6 (20%) partly answered with yes

Recommendation:

The answers indicate that the full scope and consequences of the adoption of the FAIR principles are not yet recognised. In this respect, the FAIR principles should be made more widely known. It is also essential that detailed and discipline-specific requirements are included in the discussion. These detailed criteria must be used for a realistic evaluation. More information and tools for self-evaluation are available [FAIR-Aware: Assess Your Knowledge of FAIR](#), [Your first step towards your FAIR data(set)/](#), [FAIR Principles](#) and [Automated FAIR Data Assessment Tool](#).

33. What certifications have the digital archive achieved?

Premise:

Certifications are the basis for trust in a digital archive; this is particularly important for external services (see question 27). They are awarded in an intensive and complex procedure, and the requirements are high.

Summary:

Only one answer indicated a certification ('Core Trust Seal'). However, this answer seems to be from an Austrian source and so no certification is mentioned from Germany.

Recommendation:

Certificates should be sought in general and by archive service providers. Owing to the complex requirements, support should be offered for this. In this respect, too, the high costs can be better absorbed by central providers.

3.9. Closing

34. Self-assessment: at what level is your institution currently implementing digital archiving?

Premise:

The question is based on the Business Process Maturity Model with the BPMN levels: Initial, Managed, Standardised, Predictable, Innovative. The levels of implementation of digital archiving given in the survey increase from the top to the bottom of the statement set.

Summary:

The responses indicate the following levels of maturity:

- 2 (7%) No awareness of the requirements,
- lowest BPMN level:
  6 (20%) Knowledge of basic principles and recognition of the need for action,
- 0 An implementation plan has been defined,
- 3 (10%) Basic processes have been introduced,
- 2 (7%) A comprehensive management process has been initiated to avoid losses,
- highest BPMN level:
  0 Proactive measures have been initiated to continuously improve implementation.

Six institutions know basic principles and recognise the need to act (lowest BPMN level) and five institutions can be identified with implemented archiving processes (higher BPMN level).

Recommendation:

Since a robust catalogue of criteria was missing, the self-assessment was probably intuitive. However, the trend does not contradict the answers to other questions, so it can be deduced that the maturity of the implementation of digital archaeological archiving can be significantly improved. For this, support is needed: support for the processes themselves, as well as to raise awareness in some institutions of the need to archive digital archaeological data.

---

35. **In which areas would (external) support be helpful for you? Or do you have any comments or advice on the survey?**

Premise:

The last question provided the opportunity to give suggestions.

Summary:

The following suggestions were named accordingly:

- o Push the issue more aggressively.
- o Competence enhancement (training etc.).
- o Resources (personnel, material resources).
- o Create central standardised resources/archives!

Recommendation:

The desire for support should be taken seriously and given due consideration.

4. Overview of the responses

With 30 valid response sets, the survey provides an adequate picture of the current state of archiving digital archaeological data in Germany. In three-quarters of the institutions, there is no regular policy for archiving digital archaeological data, and in half there is no permanently established responsibility for digital archaeological archiving. There is a clearly identifiable lack of competence and specialist staff. In three-quarters of the responses, the OAIS standard relevant to digital archiving is unknown. Only three institutions work according to this standard.

The state institutions base their motivation on the relevant (specific) state monument protection and archive laws. In addition, they are motivated by the fact that the data, some of which contain irreplaceable information, are necessary for the protection of monuments and should also remain usable for future generations and subsequent research. About one-third of the responses confirm that the topic is strongly present at the management level.

However, this motivation is countered by obstacles that can be roughly summarised as follows:

- Insufficient consideration within both working and research processes.
- Tasks and responsibilities are not clearly defined.
- There is a clear lack of resources.
- The complexity of the issue.
- Existing expertise is insufficient and adequate support is lacking.

Most of the archiving takes place in-house; only two institutions employ an external service provider. One institution is a service provider itself. When digital archival material is transferred, it is rarely checked for archival value and archival suitability. Detailed indexing is seldom employed. In many cases, metadata is recorded either as an overview, as a table of contents or, if in detail, for individual documents or images. Common standards such as Dublin Core, EXIF, IPTC and METS are used for this purpose.

The archival material covers all areas of archaeological documentation. The range of material types extends from simple (texts, images etc.) to complex (interoperable linked data, increasingly used in digital excavation documentation). It should be possible to develop simple format-based or more complex preservation strategies. The formats used are wide-ranging. It is not obvious that distinct formats are being used for recording, archiving (archival storage) and reuse.

About half of the responses confirm their use of suitable central server storage for archiving. However, just under a third refer to unsuitable archive storage (local PCs, external hard drives, USB storage, CDs/DVDs). Less than half of the responses report daily backups. Only a few cases refer to bit-stream preservation and electronic signatures to maintain data integrity and authenticity.

Interestingly, hard copy, hardware preservation and emulation also play a role as principal archiving methods in addition to migration. With regard to migration, a format-related preservation strategy is only referred to five times. There was no feedback on 'content-related curation' (meaning the supervision of migration processes to preserve the subject content of the archival material as the creator/author intended).

Less than half of the respondents are aware of the FAIR principles. The responses indicate that some people have heard about FAIR but have not yet recognised its full scope and the necessary consequences.

Certifications are the basis for trust in a digital archive; they are awarded in complex procedures with very high requirements. But there were no certifications mentioned among the respondents from Germany.

The self-assessment question about the level of implementation was inspired by the Business Process Maturity Model. It reveals that six institutions (20%) have knowledge of basic principles and recognise the need to act (lowest BPMN level) and five institutions (<20%) can be identified with implemented archiving processes (higher BPMN level).

5. Recommendations

Recommendations for further and supporting measures were generated for almost all questions and answers of the survey and are summarised here once again.

The participants in the survey themselves would like to see the issue addressed more aggressively, more adequate resources made available, knowledge and skills imparted, and standardised central solutions promoted.

From a strategic point of view, awareness of the importance of digital archaeological archiving should be promoted more clearly and aggressively. To this end, tasks and responsibilities must be more clearly defined. Standards such as OAIS and FAIR should be made more widely known and there should be more focus on their implementation. Finally, competence must be improved at the management level. Ultimately the inception and implementation of the archiving processes are management tasks.

Another focus is on resources. Archiving cannot take place 'on the side', even digitally. There is a clear need for human and financial resources for an adequate and sustainable implementation of digital archaeological archiving, resulting primarily from the very high complexity of the topic. For the effective use of the available resources, there is no way of 'avoiding' the concept of 'centralisation'. Since not all costs increase with the amount of archival material, sharing or reusing existing (central) resources promises economic solutions. Orientation aids - also for management - would be useful for consideration and decision-making processes in this regard.

'Archaeological archiving is a dynamic process from the beginning of project planning' (EAC [2014], 19). This process must be understood as integrative, starting with the first data collection, and supported accordingly through recommendations and best practices during a complete project life cycle. In addition, quality and risk management play a major role; criteria in this regard should be coordinated in the specialist community and made available centrally.

Digital archaeological archiving requires specialised knowledge; existing knowledge should be extended for this purpose. An increase in skills through education and training is necessary, especially on the topics of OAIS, metadata standards, legal principles and FAIR. The exchange of experience in specialist networks can help. A 'competence centre' would also save resources spent on specialised knowledge.

Acknowledgements

Bibliography

Europae Archaeologiae Consilium (EAC) Working Group Archaeological Archives 2014 *EAC Guidelines 1 - The Standard and Guide to Best Practice in Archaeological Archiving in Europe*. https://www.europae-archaeologiae-consilium.org/eac-guidlines[Last accessed: 17 February 2022]

Göldner, R., Bibby, D. and Senst, H. 2022 'Archivierung digitaler archäologischer Daten : Auswertung einer Umfrage', *Forum for Digital Archaeology and Infrastructure* **1**. https://doi.org/10.34780/edw6-f6ea

Mayernik, M.S., Breseman, K., Downs, R.R., Duerr, R., Garretson, A. and Hou, C-Y. 2020 'Risk Assessment for Scientific Data', *Data Science Journal* **19**(1), 10. http://doi.org/10.5334/dsj-2020-010

RDA FAIR Data Maturity Model Working Group 2020 *The FAIR Data Maturity Model. Specification and Guidelines*, https://zenodo.org/record/5834115

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* 2016 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data* **3**, 160018(2016). http://www.nature.com/articles/sdata201618