Long-term Archiving in the NFDI

Results of a survey exploring current trends and requirements

Katharina Markus (ORCID:0000-0002-9316-8982), Kai Naumann (ORCID:0000-0002-2799-1030), Markus Schmalzl (ORCID:0009-0005-7574-2699), Juliane Watson, Dagmar Triebel (ORCID:0000-0003-1980-3148)

Introduction

Shortcomings in long-term archiving have the potential to become the Achilles heel of German National Research Data Infrastructure (Nationale Forschungsdateninfrastruktur, NFDI). This is because digital research data is exposed to numerous risks, including missing licences and usage rights[1], a lack of organisational and financial guarantees for institutional services[2], unclear responsibilities and processes[3], human error, hardware or storage media failure[4], hardware obsolescence[5], software obsolescence[6], research software that no longer works and many other issues. In such cases, data recovery is either impossible or extremely costly and resource-intensive – and reconstruction of lost data again is often out of the question[7]. These risks may materialise even before the end of the ten-year retention period recommended by the German Research Foundation (Deutsche Forschungsgemeinschaft,

---

[1] "In the intervening 15 years, new issues and approaches to rights management in the digital environment – such as CC licenses – have arisen. These issues, along with content licenses, copyright and intellectual property laws, will inevitably continue to evolve; highly specific and legally binding terms may prove out of sync with long-term preservation goals." (Sprout and Jordan 2018)

[2] "We also found databases whose search function had either disappeared or was nonfunctional or had lost other key functionality. We counted these as dead even if they still existed on the Web. […] according to our criteria, a total of 203 (62%) were dead […]. (Attwood, Agit and Ellis 2015)

[3] "[…] leaving a very large portion (~3 TB) of legacy data in need of organization and integration. Routine data management tasks such as finding records, identifying data types and production dates, sorting through multiple copies, culling corrupted and redundant files, and reorganizing data have been conducted manually, placing a significant burden on research staff." (Arora, Esteva and Trelogan 2014)

[4] "Many of the tapes had suffered "stiction' problems and were literally falling apart.[…] Several data sets were unrecoverable – some waterfowl data and land use data in the Eastern Townships." (Bleakly 2002)

[5] (Digital Preservation Coalition (DPC) 2023b; Bleakly 2002)

[6] "Much of the software in which this data was created is now obsolete, and some of the data formats are challenging or impossible to read in modern software." (Pinnick et al. 2018). "In the early 2000s, the data recovery company was unable to make sense of the recovered bit files – they were in a proprietary format and could not be accessed." (Doig 2020)

[7] "Some geoscience data acquisition projects are too expensive to repeat, such as drilling deep boreholes to the depth of many kilometres costing tens of millions of pounds […] Seismic data originating from earthquakes are another example of unique and unrepeatable data […]." (Pinnick 2017)

DFG)[8], thus jeopardising the reproducibility, reusability and traceability of the data and undermining good research practice. The only way to ensure research data remains FAIR (findable, accessible, interoperable and reusable)[9] for decades to come is by adopting a carefully crafted long-term archiving solution[10].

The German Council for Scientific Information Infrastructures (Rat für Informationsinfrastrukturen, RfII) has also acknowledged this situation by declaring long-term archiving to be one "of the most important tasks of a national research data infrastructure"[11]. In 2021, the NFDI carried out an internal assessment of its consortia to develop an initial understanding of their needs. This revealed that the consortia had little knowledge of long-term archiving[12] and suggested there was a need for a solution. Thus, long-term archiving became a candidate for a potential NFDI basic service and was included as part of the Base4NFDI application[13]. At the same time, a Long-Term Archival (LTA) working group[14] was established in the NFDI Section Common Infrastructures[15] to tackle the topic of long-term archiving across all consortia.

In order to lay the foundations for creating a suitable LTA basic service, the LTA working group carried out a detailed survey of the needs of NFDI consortia. The results of this survey are discussed below.

Methodology and limitations of the analysis

The survey on the status of long-term archiving (LTA) in the consortia was designed to identify needs, shortcomings and existing knowledge in the LTA context. Due to the highly diverse nature of the consortia, the members of the LTA working group opted to allow the survey participants to formulate their responses as free text. The survey questions are preceded by an introductory paragraph which explains and defines long-term archiving on the basis of relevant standards (ISO 14721[16]). This is followed by a section of questions on the consortium's existing or planned services and a section exploring their requirements for a new LTA service or for the expansion of an existing LTA service (see appendix).

   Once the survey was ready, the members of the LTA working group identified contact partners in each consortium (26 at the time of the survey according to the NFDI's list[17]) and asked them to complete it. Since some of the working group members are also members of one or more consortia, they were in some cases able to complete the survey themselves or to discuss it within their respective consortia. No consent was sought or obtained to publish

---

[8] (Deutsche Forschungsgemeinschaft. 2022b; 2022a)
[9] (Wilkinson et al. 2016)
[10] The EOSC Long-Term Data Preservation Task Force also links long-term archiving and FAIR (Andreu et al. 2023).
[11] (German Council for Scientific Information Infrastructures 2016; Rat für Informationsinfrastrukturen 2016)
[12] Unpublished results. A general need was also mentioned in the Base4NFDI application (Bernard et al. 2023)
[13] (Bernard et al. 2023)
[14] (Bach et al. 2022)
[15] (Diepenbroek, Schimmler and Ebert 2021)
[16] ('ISO 14721:2012. Open Archival Information System (OAIS) Reference Model' 2012)
[17] (NFDI 2023)

the detailed results of the survey: this was a deliberate decision aimed at encouraging the individual consortia to provide straightforward, factual responses and to prevent them from being swayed by strategic considerations of how they were presenting themselves. As a result, this publication only includes detailed information on specific consortia if that information is openly available.

The responses were evaluated by a team within the LTA working group. The option of adding free-text comments to responses limited the ability to perform a quantitative analysis, so the primary results of this evaluation were qualitative. Where responses were unclear (for example if a respondent answered a question such as "Does your consortium have [...]" with a description of a research data service rather than a simple yes or no), the team made their own interpretation of the response (as yes, no, or ?) for the purposes of the evaluation. If, however, a respondent provided a clear yes or no and then went on to qualify their answer in the comments, this clear yes or no was nonetheless taken as the basis for evaluation. One consequence of this is that statements such as "Yes, x is available from one data provider" from one consortium and "No, x is only available from one data provider" from another consortium were treated differently in the analysis despite their similar content. This imprecision was tolerated in order to avoid skewing the statements made by the consortia, and it did not affect the observed trends that are discussed in this paper.

Results and observed trends

Amount of responses received

Even though all the consortia received the survey, it was not possible to obtain responses from all of them. At the time of publication, 17 of the 26 consortia (65 percent) had submitted a response.

The consortia that returned the completed survey are spread across all four of the NFDI's major scientific disciplines: humanities and social sciences, engineering sciences, life sciences and natural sciences[18] (Fig. 1).
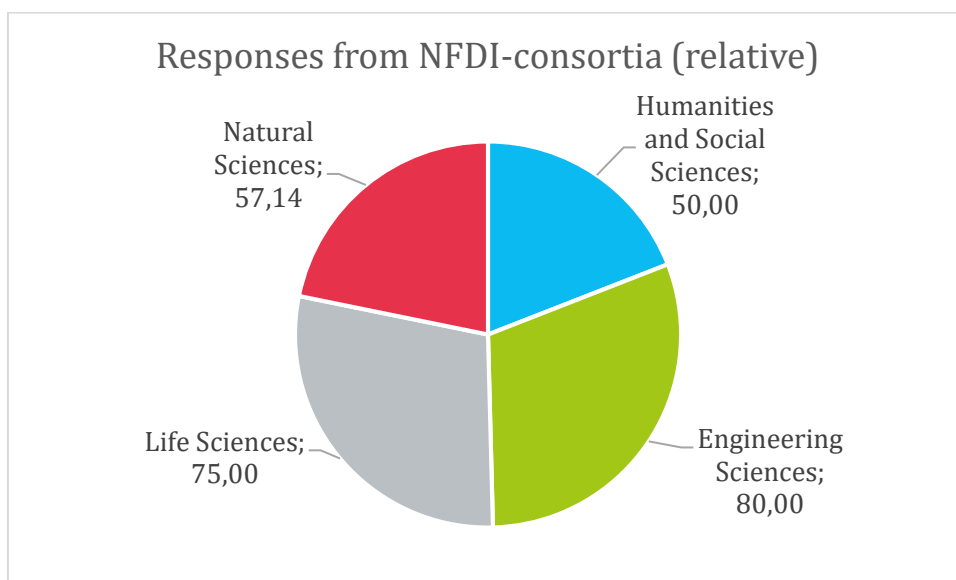
---

[18] (NFDI 2023)
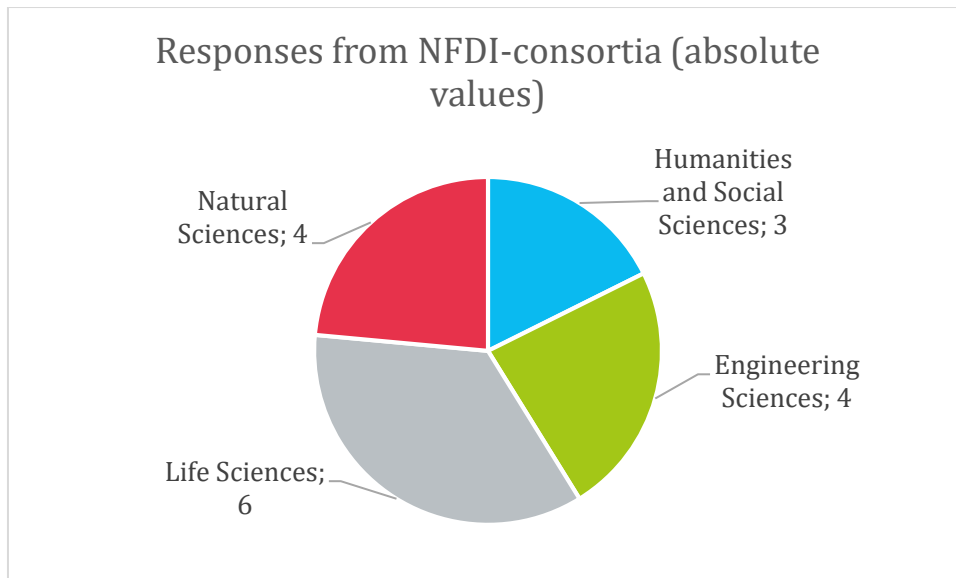
Fig. 1 Responses from consortia by scientific discipline.
Consortia were assigned to the respective disciplines based on the categories published by the NFDI.[19]

The discipline that yielded the most responses was the life sciences (responses from six consortia), while the humanities and social sciences yielded the fewest (responses from three consortia). Thus, the survey mainly reflects responses from the life sciences, natural sciences and engineering sciences. However, since the humanities and social sciences are likelier to work with legacy data (e.g. archival holdings), it can nonetheless be assumed that people working in this field have an above-average awareness of long-term archiving.

Fig. 2 Proportion of total consortia in each scientific discipline that responded to the survey.
Figures calculated as a percentage, with all the consortia in a scientific discipline equalling 100 percent. Consortia were assigned to the respective disciplines based on the categories published by the NFDI.[20]

---

[19] (NFDI 2023)
[20] (NFDI 2023)

The absolute numbers (Fig. 1) do not provide sufficient information on how response behaviour differs between scientific disciplines. Accordingly, the response rate was calculated as a percentage of all consortia within a discipline (Fig. 2). This reveals an 80 percent survey response rate for consortia in the engineering sciences, falling to just 50 percent for consortia in the humanities and social sciences.

Understanding of long-term archiving, LTA terminology and standards

In this paper, the term "long-term archiving" is synonymous with long-term preservation as defined in the Open Archival Information System (OAIS) standard: "Long Term Preservation: The act of maintaining information, Independently understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term."[21]. The following terms are used synonymously in this paper: digital long-term archiving, long-term archiving[22], long-term availability, digital preservation and long-term preservation. Long-term archiving is not subject to a time limit and may, for example, significantly exceed ten years. In this respect, it differs from the Guidelines for Safeguarding Good Research Practice published by the German Research Foundation (DFG), which define "archiving" as the retention of data in institutions "for a period of ten years", though the DFG also states that shorter or longer periods may also be appropriate[23].

In order to clarify the scope of long-term archiving and the responsibility for preservation actions in institutions and consortia, the Object Levels of Preservation[24] model was used both in the survey and in its evaluation. It defines three different preservation levels:

- bitstream preservation,
- logical preservation and
- semantic preservation.

These three levels build one upon the other.

Bitstream preservation refers to the preservation of digital objects on the bitstream level. Logical preservation uses techniques such as migration to new file formats and emulation to prevent information from being lost as a result of e. g. obsolete formats, obsolete software or outdated technologies. Semantic preservation aims to preserve objects on the conceptual level, for example by preserving context, documenting authenticity, object structure, relationships between objects and sufficient descriptive metadata.

The survey conducted by the LTA working group combined the two levels of logical preservation and semantic preservation into a single category called preservation on the functional level (see survey in appendix), thus allowing all consortia which stated they offer

---

[21] ('ISO 14721:2012. Open Archival Information System (OAIS) Reference Model' 2012; The Consultative Committee for Space Data Systems (CCSDS) 2012)

[22] Long-term archiving is the direct translation of the German term „Langzeitarchivierung", which is predominently used in Germany for the tasks and responsibilites the OAIS standard covers ('ISO 14721:2012. Open Archival Information System (OAIS) Reference Model' 2012; The Consultative Committee for Space Data Systems (CCSDS) 2012). The international community uses digital preservation and long-term preservation instead.

[23] (Deutsche Forschungsgemeinschaft. 2022b; 2022a)

[24] (Lindlar et al. 2020)

functional preservation to cover both logical preservation and semantic preservation (see also Fig. 3). These two levels are also grouped into a single category by the authors of the NFDI4Culture guidelines, who favour the term "content preservation"[25]. This definition of content preservation is adopted below and used synonymously with the term functional preservation (Fig. 3).

Currently, hardly any of the consortia are addressing the issue of digital long-term archiving in any depth. Of the 17 consortia that responded to the survey, 11 said they practice bitstream preservation, while just six practice content preservation. Some responses that listed content preservation did not specifically mention bitstream preservation, but since the former builds on the latter, it was assumed that both forms of preservation are practised in these cases.

None of the consortia explicitly said that they do not practice any form of bitstream preservation, but six of the respondents failed to respond to this question. In each of these cases, the failure to answer this question correlated with the absence of an LTA service. Some of these consortia made reference to existing research data services (data storage solutions, specialist databases, subject repositories, generic repositories, university publishing and storage services, research data centres, archives, software directories, workflow platforms, etc.), while others declared that long-term archiving was not their responsibility or appeared to have no experience or connection with the topic of long-term archiving.

In contrast, the six consortia that say they practise content preservation often point to a specific research data service that preserves data on a higher level than the bitstream level, though this does not necessarily imply that all of the discipline's data is digitally preserved in this way. The challenges of incorporating a variety of services and distributed LTA solutions are examined in the next section.

Lindlar et al. 2020 / Thibodeau 2002[26]	WissGrid project 2011[27]	RfII 2016[28]	EOSC TF Long Term Data Preservation	NFDI4Culture 2023[30]	LTA WG survey 2023

---

[25] (Heseler, Büttner and Arnold 2024)

[26] "These classes are the Physical, Logical, and the Conceptual Object, which themselves have unique properties […] and bear distinct risks, e.g. lack of robust storage on the physical level, unsuitable file formats or dependencies on externally linked resources on the logical level, and incomplete accompanying descriptive metadata or semantic drift on the conceptual level." (Lindlar et al. 2020). (Rudnik and Lindlar 2021; Thibodeau 2002)

[27] "Content Preservation: For citations it is not sufficient that the bits of the object are still present. The used technology has to reproduce the content ensuring its authenticity even if the original technical environment might be no longer available. This is a shift of perspective from technical stability to technical reusability. […]" (Aschenbrenner et al. 2011)

[28] "The longer the data needs to be retained, the more likely it will be necessary to take measures that not only preserve the data technically, but also preserve its content (referred to as content preservation). Necessary measures could include migrating the data to new formats or making changes to adapt to new hardware and software environments. […]" (German Council for Scientific Information Infrastructures (RfII) 2016; Rat für Informationsinfrastrukturen 2016)

[30] "Eine weitere Strategie der digitalen Langzeitarchivierung ist der Erhalt der Nutzbarkeit gespeicherter Informationen. Dieser wird durch die Content Preservation erreicht und schließt

2023[29]					
"Bitstream Preservation"	"Bitstream Preservation"	-	-	"Bitstream Preservation"	"Bitstream Preservation"
"Logical Preservation"	"Content Preservation"	"Content Preservation"	"Logical-Technical Preservation"	"Content Preservation"	"preservation on the functional level"
"Semantic Preservation"	"Data Curation"	"Data Curation"	"Conceptual preservation for understanding and reuse"		

Fig. 3 Table of terms used in the literature to describe preservation levels or curation levels and how they relate to the Object Levels of Preservation. The terms in each column stem from the source named in the first row. Each term was mapped to the Object Levels of Preservation by the authors of this paper based on the definition provided in the respective source. This mapping of terms was only explicitly stated by a few of these sources. It is not always possible to make a clear distinction between semantic preservation and content preservation, so this table is only an approximation.

It can be assumed that data loss at levels higher than the bitstream level will be relevant for most research data services, even within the ten-year timeframe recommended by the DFG (cf. the corresponding reports from the DPC 2023[31]). Research software may cease to be sold just a few years after its launch, or it could end up being incompatible with formats used ten years previously. Given this situation, there is a need to link existing research data services to long-term archiving solutions or to upgrade these services – including a corresponding increase in personnel – in order to prevent the loss of important data and information. At the same time, research funders need to provide the financial support required to make these improvements. Thus, the survey shows that the NFDI lacks three key things: a common understanding of long-term archiving, a common set of terminology and an active implementation of long-term archiving.

The NFDI's LTA working group therefore recommends:
- creating a highly visible glossary of key terminology (e.g. on forschungsdaten.info);

---

neben der technischen Interpretierbarkeit (logische Perspektive) auch die inhaltliche Interpretierbarkeit (semantische Perspektive) ein." (Heseler, Büttner and Arnold 2024)
[29] "Logical-Technical Preservation. In addition to D and/or C above the repository takes long-term responsibility for ensuring that the data and metadata are updated over time to newer standards and formats in response to technical risks (e.g. file format or software obsolescence), the changing needs of the designated community (e.g. newer alternate formats become necessary for reuse)." (Andreu et al. 2023)
[31] "Digits are born vulnerable. Every single byte of data depends on a global infrastructure of technology, process and people for its meaning and purpose to be realized. Much data serves the moment: it is quickly forgotten in a continuous flow of process and interaction." (Digital Preservation Coalition (DPC) 2023b)

- ensuring that the NFDI also gets involved in efforts to standardise terms in the English-speaking world (cf. the 2020 paper by Lindlar and colleagues[32]), with EOSC channels offering a suitable tool to achieve this;
- ensuring that consortia have access to consultation and information on LTA;
- creating funding options for LTA solutions[33].

The survey responses also point to different perceptions of long-term archiving standards (OAIS[34], ISO 14721[35], ISO/IEC 16363[36], DIN 31644:2012-04[37], DIN 31645:2011-11[38], PREMIS[39]). When a standard clearly benefits all the parties involved or helps avoid risks, compliance tends to be high. But standards in other areas often tend to be ineffective, or are rendered obsolete by changing circumstances. In the case of video tapes, it took from the 1970s to the late 1980s for the VHS standard to become dominant. In contrast, it took just a few years to establish standardised internet protocols. Examples in the LTA context include growing efforts to create an internationally recognised standard for software-independent archiving of database structures and content (see DILCIS Board[40]). Such initiatives are important for various research communities that practice long-term data archiving in large open source databases. The SIARD concept and format is based on creating information packages (SIPs) that conform to the ISO standards Unicode, XML and SQL:2008 as well as the URI Internet standard and the ZIP industry standard[41]. SIARD is one of the standards used by DILCIS for eArchiving[42], and SIARD tools are employed by subject repositories and national archives. SIARD is currently being used in an NFDI4Biodiversity pilot project[43].

During the introductory process, standards need institutions that are willing to comply with them and thus to exert an influence on other organisations. This is why large research data institutions such as DESY chose to participate in an EU-level project on long-term archiving systems while adhering to LTA standards[44].

Against this backdrop, the LTA working group proposes the following:
- The NFDI should commit to upholding, expanding and implementing existing LTA standards in the context of FAIR principles.

---

[32] (Lindlar et al. 2020)
[33] See also recommendations 96 and 97 of the EOSC Long Term Data Preservation Task Force (Andreu et al. 2023)
[34] (The Consultative Committee for Space Data Systems (CCSDS) 2012; 'ISO 14721:2012. Open Archival Information System (OAIS) Reference Model' 2012)
[35] ('ISO 14721:2012. Open Archival Information System (OAIS) Reference Model' 2012)
[36] ('ISO 16363:2012. Audit and Certification of Trustworthy Digital Repositories' 2012)
[37] ('DIN 31644:2012-04. Information and documentation – Criteria for trustworthy digital archives' 2012)
[38] ('DIN 31645:2011-11. Information and documentation – Guide to the transfer of information objects into digital long-term archives' 2011)
[39] (PREMIS Editorial Committee 2015; Guenther, Dappert and Peyrard 2016)
[40] (DILCIS Board 2024a)
[41] (DILCIS Board and Swiss Federal Archives 2021; SFA 2024)
[42] (DILCIS Board 2024b)
[43] (Weibulat et al. 2023)
[44] (Fernandes et al. 2020; ARCHIVER Project 2019)

- Working on the basis of FAIR data principles, the NFDI should, as one of its core tasks, promote the development of expertise in LTA of research data and support training programmes at institutions that practice LTA of research data.
- The NFDI should encourage experts to get involved in national and international initiatives such as Research Data Alliance (RDA) Interest Groups[45], Digital Information LifeCycle Interoperability Standards Board (DILCIS Board) working groups and the EOSC Long Term Data Preservation Task Force (LTDP-TF). By participating in such initiatives, experts can help create concepts and develop new versions of LTA specifications and standards (at all three levels: bitstream, logical-technical and semantic).
- The NFDI should work with bodies such as CoreTrustSeal[46], the nestor Certification working group[47], the nestor OAIS Review working group[48] and the PREMIS Maintenance Activity and Editorial Committee[49] to support the documentation, conceptual adaptation and ongoing development of existing standards in order to increase awareness and boost acceptance of LTA in the NFDI.

Distributed or centralised LTA solutions in consortia

Asked whether their consortium offers long-term archiving service(s), the survey respondents listed a large number of research data services (see also institutions that collaborate with the NFDI[50]), all with their own individual characteristics. These services differ in various ways, for example in the scope of their content preservation/LTA strategy, as well as in regard to whether or not they have a preservation policy. As a consequence, and in order to gain an overview of actual and potential LTA services, an analysis was carried out to determine whether each consortium is pursuing a distributed or centralised LTA solution. In this context, centralised LTA means that the consortium offers a central LTA service for all, or nearly all, of the data in which the consortium specialises. This tends to be a central organisational unit with, for example, centralised standards and a business model, though hosting may still be carried out at a variety of institutions.

When asked whether they had an LTA service, almost all the consortia pointed to various existing research data services, thereby implying that they use a distributed LTA solution. This is in line with NFDI efforts to expand, reuse and connect existing services and solutions[51]. In many cases, however, it is unclear whether these research data services really offer LTA in the sense of content preservation. Unless a commitment to content preservation is clearly stated, it cannot be assumed that any deliberate, planned preservation is taking place beyond the bitstream level. In many cases, it is also unclear

---

[45] See also (Taylor-Grant, Molloy and Rhee 2015), for example

[46] (de Leeuw and Crabtree, n.d.)

[47] (Schoger and Keitel 2020; 'Normenausschuss 009-00-15 AA Records Management und Langzeitverfügbarkeit von Informationsobjekten', n.d.)

[48] ('Nestor – AG OAIS Review' 2022; 'ISO/TC 20/SC 13 – Space Data and Information Transfer Systems' 2021)

[49] ('PREMIS Maintenance Activity and Editorial Committee (Library of Congress)' 2022)

[50] (Miller, Rißler-Pipka and Weimer 2022)

[51] (Rat für Informationsinfrastrukturen 2016; German Council for Scientific Information Infrastructures 2016)

whether these research data services form part of the consortium and are official partners, or whether they are simply involved in a logistical role, for example because they are well-known in the discipline as a service provider.

One way to clarify LTA coverage within a discipline would be for research data services to obtain comprehensive certification via CoreTrustSeal[52] and/or nestor seal[53]. These research data services should also offer clear information on the scope of the preservation actions that they apply in a needs-based manner to their own data[54]. There should also be transparency regarding the relationships between different services, particularly in the context of long-term archiving[55].

The use of LTA tools to identify and validate formats (e.g. DROID, Siegfried, JHOVE) could become standard practice when it comes to processing research data and publishing research work. The results of this tool-based file analysis could be made publicly available as LTA metadata. The same applies to metadata that supports emulation.

There is a strong case for consortia to take on a coordinating and monitoring role for LTA issues in their specialist areas. Although a German consortium can only exert a limited influence on international subject databases and data providers, this impact should not be entirely dismissed – especially since long-term archiving is becoming increasingly visible on an international level[56].

The LTA working group therefore recommends
- that consortia should coordinate the implementation of LTA for data in their specialist area and monitor the status of this implementation;
- that research data services should be certified with the CoreTrustSeal and/or the nestor seal;
- that consultation and funding should be made available to consortia and participating institutions in the LTA context.

Policies and responsibilities

The importance of clarifying responsibilities in the context of long-term archiving should not be underestimated, since the failure to allocate roles and assign tasks is yet another risk factor for data loss.

It is advisable to decide which data in a discipline or subject area should be digitally preserved and to record this decision in a policy. Not all of the data generated through research needs to be preserved, but it is vital that the selection process is both systematic and transparent. In its 2016 recommendations on long-term archiving, the RfII calls this selection process "relevance assessment" and suggests that this task should primarily be undertaken within the disciplines themselves, since they have the requisite expertise[57]. The

---

[52] (CoreTrustSeal Standards and Certification Board, n.d.)
[53] (Bähr et al. 2019)
[54] See also recommendation 19 of the EOSC Long Term Data Preservation Task Force (Andreu et al. 2023)
[55] See also recommendation 40 of the EOSC Long Term Data Preservation Task Force (Andreu et al. 2023)
[56] (Andreu et al. 2023)
[57] (German Council for Scientific Information Infrastructures 2016; Rat für Informationsinfrastrukturen 2016)

impact of these RfII recommendations was measured in the survey by a question on whether the consortium has a "collection policy" in place. The key here is that the criteria used to select data for long-term preservation (i.e. the assessment or appraisal criteria) should be easy to communicate and presented as transparently as possible[58]. Examples of such criteria might include the data's relevance for the particular subject area, as well as copyright and licensing, the size of a data set, its uniqueness and so on. If a conscious decision has been made that the researchers in a research organisation or department can and should be the ones to decide on which data to include for long-term archiving, then this should also be documented in a policy. However, most consortia who responded to the survey said that they do not have a collection policy. There is therefore still a need to provide advice and support to the consortia, or to the consortia's data repositories, to help them choose which data to preserve.

General long-term archiving decisions – such as the scope of long-term archiving, the question of which preservation actions can be carried out by whom and the specification of the target group for which the archived data is being preserved – are typically documented in a preservation policy[59]. A consortium can create a general policy on which data they consider themselves responsible for. They can also encourage data storage institutions to obtain certification (CoreTrustSeal, nestor seal) and assess whether the data in their subject area is being preserved in an adequate and appropriate way. All these matters can be documented and published in a preservation policy at the consortial level[60]. However, very few consortia answered in the affirmative when asked whether they had a general preservation policy in place, or even a policy for individual institutions. Similarly, in a survey conducted across Europe, 37 percent of all repositories stated that they do not have a preservation policy. Consequently, the lack of good practice is cited as one of the main challenges in long-term archiving[61].

The LTA working group recommends
- establishing clear roles and responsibilities, including for preservation actions and for the monitoring of technical developments[62], and
- setting up an consultation service to support consortia with this issue.

Data and data types not covered by LTA

The survey of NFDI consortia shows that many of the document classes and data types used in the scientific communities represented by the consortia continue to lack the kind of LTA solutions described above. Thus, 11 of the 17 consortia expressed the need for an LTA service, two answered "no" to this question and four consortia failed to clarify whether they needed such a service or not. In this evaluation, both the general need for an LTA service and a specific need for an LTA solution for certain types of data were interpreted as a need

---

[58] See also recommendation 61 of the EOSC Long Term Data Preservation Task Force (Andreu et al. 2023)

[59] (nestor-Arbeitsgruppe Policy 2014; nestor working group Policy 2014; Digital Preservation Coalition (DPC) 2023a)

[60] See also recommendation 114 of the EOSC Long Term Data Preservation Task Force (Andreu et al. 2023)

[61] (Shearer et al. 2023)

[62] See also recommendations 71 to 82 of the EOSC Long Term Data Preservation Task Force (Andreu et al. 2023)

for an LTA service. This need relates to data from high-performance research and citizen science projects, as well as information from original systems that have been evolving for decades in this or similar forms in research, administration or other contexts. Such systems may already have standards and best practices in place at regional, national and international levels, such as the archiving of text documents or database content.

The gaps and shortcomings revealed by the survey for certain types of data are therefore also a reflection of the low priority granted to long-term archiving and the lack of engagement with and knowledge of this topic within the NFDI. The survey identified a lack of LTA solutions for the following groups of documents and data and/or data-product classes (based on the need for an LTA service expressed by consortia in the survey and supplemented by the authors based on experience):

- large to very large data sets (terabyte and petabyte scale), for example from data modelling or from multidimensional evaluation approaches for integrated data (data cubes),
- software archiving solutions (research software),
- citizen science data,
- databases and data tables,
- documents and information from project management,
- measurements and measurement series,
- streaming data,
- social media data,
- "long tail data" (cf. an assessment of this at the EU level[63]).

In most cases, the solutions that are needed tally with observations made at the international level. According to the Digital Preservation Coalition (DPC), the media – and especially the data formats and infrastructures – that are used for research data are extremely vulnerable. The DPC's international survey classifies the following 'digital species' as critically endangered or endangered[64]:

- obsolete optical and magnetic media (e.g. hard disks, USB sticks, DVD ROMs), also when provided as a supplement to printed research results,
- research data on GitHub and in wikis,
- manuals, instructions, informal private correspondence about research (e.g. personal network drives, e-mail accounts, niche social networks),
- project-specific data portals on the web that are succumbing to technological change, especially in medicine, but also in other areas of science (e.g. with citizen science participation),
- commercial software involved in the research process published by companies that no longer exist, in particular cloud services,
- open source intelligence from non-governmental sources and data from NGOs, including results of their analysis,
- audiovisual materials, 3D architectural and engineering drawings, data sets from the field of virtual reality,

---

[63] "A long-tail of data exists that has not been brought into a managed storage, curation or preservation system." (Andreu et al. 2023)
[64] (Digital Preservation Coalition (DPC) 2023b)

- objects of all kinds whose unresolved legal status (uncertainty over intellectual property rights) impedes preservation.

Conclusions and recommendations

Long-term archiving, which enables valuable research data to be preserved for ten years or more, has already been acknowledged as both necessary and desirable by the RfII[65]. Yet the survey on long-term archiving in NFDI consortia reveals that few consortia have actually engaged with the topic of long-term archiving beyond the bitstream level, and even fewer have come up with any appropriate solutions. Thus, it is currently impossible for the NFDI to fully achieve its goal of making data available according to FAIR principles, and it seems reasonable that this aspiration will remain out of reach in the near future. The survey results also highlighted that most consortia have a need for an LTA service.

Given this situation, the LTA working group considers consultation for the consortia about the following topics necessary:
- What is LTA? And what are the benefits of LTA?
- Which LTA standards exist and how can they be implemented?
- What is the established, internationally recognised LTA terminology?
- Which LTA solutions are available – whether structural, technical, organisational, or personnel-related – and how can they be connected?
- Which responsibilities need to be clarified?
- Which aspects of research data management need to be considered to achieve the desired LTA preservation level?

There is a need to establish or improve LTA awareness and functionalities within many of the consortia, as well as in the research data services they refer to in the survey.

Meeting this need would require a consulting service, either situated within the NFDI or working in close collaboration with the NFDI. It is important to incorporate the nestor competence network[66] in this context, since it already has the relevant expertise.

The LTA working group also makes the following recommendations:
- LTA projects in institutions and consortia should receive financial support.
- Research funders should provide funding for long-term archiving in research projects.
- Research data services should acquire certification via CoreTrustSeal and/or nestor seal.
- Efforts should be made to help shape LTA standards.
- Consortia should take steps to monitor the status of long-term archiving in the consortia's specialist areas, to identify gaps and shortcomings and to come up with solutions.

---

[65] (German Council for Scientific Information Infrastructures 2016; Rat für Informationsinfrastrukturen 2016)
[66] (nestor, n.d.)

References

Andreu, Toni, Lluís Anglada, David Antos, Thomas Bähr, Maciej Brzeźniak, Pierre-Yves Burgi, Cécile Cavet, et al. 2023. 'Recommendations Consultation. EOSC-A Long Term Data Preservation Task Force', October. https://doi.org/10.5281/zenodo.10014698.

ARCHIVER Project. 2019. 'PETRA III / EuXFEL DATA ARCHIVING'. ARCHIVER. 19 July 2019. https://web.archive.org/web/20230308132824/https://www.archiver-project.eu/sites/default/files/Technical%20Summary%20%20-%20PETRAIII_EuXFEL%20Data%20Archiving.pdf.

Arora, Ritu, Maria Esteva, and Jessica Trelogan. 2014. 'Leveraging High Performance Computing for Managing Large and Evolving Data Collections'. *International Journal of Digital Curation* 9 (2): 17–27. https://doi.org/10.2218/ijdc.v9i2.331.

Aschenbrenner, Andreas, Harry Enke, Thomas Fischer, and Jens Ludwig. 2011. 'Diversity and Interoperability of Repositories in a Grid Curation Environment'. *Journal of Digital Information* 12 (2). https://jodi-ojs-tdl.tdl.org/jodi/article/view/1896.

Attwood, Teresa K., Bora Agit, and Lynda B. M. Ellis. 2015. 'Longevity of Biological Databases'. *EMBnet.Journal* 21 (0): 803. https://doi.org/10.14806/ej.21.0.803.

Bach, Felix, Andreas Degkwitz, Wolfram Horstmann, Peter Leinen, Michael Puchta, and Thomas Stäcker. 2022. 'Concept for Setting up an LTA Working Group in the NFDI Section "Common Infrastructures"', April. https://doi.org/10.5281/zenodo.6451456.

Bähr, Thomas, Thomas Gerdes, Henk Harmsen, Christian Keitel, Christoph Schmidt, Astrid Schoger, Sabine Schrimpf, et al. 2019. *Erläuterungen Zum Nestor-Siegel Für Vertrauenswürdige Digitale Langzeitarchive, Version 2.1. Nestor-Arbeitsgruppe Zertifizierung*. Frankfurt am Main: nestor c/o Deutsche Nationalbibliothek. http://nbn-resolving.de/urn:nbn:de:0008-2019062507.

Bernard, Lars, Reinhard Altenhöner, Franziska Böhm, Michael Diepenbroek, Barbara Ebert, Juliane Fluck, Sonja Herres-Pawlis, et al. 2023. 'Base4NFDI - Basic Services for NFDI', December. https://doi.org/10.5281/zenodo.10245518.

Bleakly, Denise R. 2002. 'Long-Term Spatial Data Preservation and Archiving: What Are the Issues?' Report SAND2002-0107. Sandia National Laboratories. https://doi.org/10.2172/793225.

CoreTrustSeal Standards and Certification Board. n.d. 'CoreTrustSeal Requirements 2023-2025'. Accessed 29 October 2023. https://doi.org/10.5281/zenodo.7051012.

Deutsche Forschungsgemeinschaft. 2022a. 'Guidelines for Safeguarding Good Research Practice. Code of Conduct. V. 1.1'. https://doi.org/10.5281/zenodo.6472827.

———. 2022b. 'Leitlinien Zur Sicherung Guter Wissenschaftlicher Praxis. Kodex. V. 1.1'. https://doi.org/10.5281/zenodo.6472827.

Diepenbroek, Michael, Sonja Schimmler, and Barbara Ebert. 2021. 'Sektionskonzept Common Infrastructures Zur Einrichtung Einer Sektion Im Verein Nationale Forschungsdateninfrastruktur (NFDI) e.V.' *Zenodo*. https://doi.org/10.5281/zenodo.5607490.

Digital Preservation Coalition (DPC). 2023a. 'Preservation Policy Toolkit - Digital Preservation Coalition'. Digital Preservation Policy Toolkit. Version 2. 2023. https://web.archive.org/web/20231031005940/https://www.dpconline.org/digipres/implement-digipres/policy-toolkit.

———. 2023b. 'The Bit List 2023. The Global List of Endangered Digital Species. Fourth Edition Revised November 2023'. Digital Preservation Coalition (DPC). http://doi.org/10.7207/dpcbitlist-23.

DILCIS Board. 2024a. 'DILCIS Board Home'. 2024. https://web.archive.org/web/20240219163250/https://dilcis.eu/.

———. 2024b. 'Used Standards'. 2024.

https://web.archive.org/web/20240219163555/https://dilcis.eu/about/used-standards.

DILCIS Board, and Swiss Federal Archives. 2021. 'SIARD 2.2'. SIARD (Software Independent Archiving of Relational Databases). 2021. https://web.archive.org/web/20240219162844/https://dilcis.eu/content-types/siard.

'DIN 31644:2012-04. Information Und Dokumentation - Kriterien Für Vertrauenswürdige Digitale Langzeitarchive'. 2012. DIN-Norm. Beuth. https://doi.org/10.31030/1854419.

'DIN 31645:2011-11. Information Und Dokumentation - Leitfaden Zur Informationsübernahme in Digitale Langzeitarchive'. 2011. DIN-Norm. Beuth. https://doi.org/10.31030/1818277.

Doig, James. 2020. 'Preserving the Bits: A Salutary Tale from the National Archives of Australia'. *DigitalPreservationCoalition Blog* (blog). 6 November 2020. https://web.archive.org/web/20240103082849/https://www.dpconline.org/blog/james-doig-preserving-the-bits.

Fernandes, João, Bob Jones, Marion Devouassoux, Dominique Buyse, Ruben Van Calenberg, Vaggelis Motsenitalis, and Jakub Urban. 2020. 'ARCHIVER D2.1- State of the Art, Community Requirements and OMC Results'. Zenodo. https://doi.org/10.5281/zenodo.3974087.

German Council for Scientific Information Infrastrukteres. 2016. 'Enhancing Research Data Management: Performance through Diversity. Recommendations Regarding Structures, Processes, and Financing for Research Data Management in Germany'. Göttingen. urn:nbn:de:101:1-20161214992.

Guenther, Rebecca Squire, Angela Dappert, and Sébastien Peyrard. 2016. 'An Introduction to the PREMIS Data Dictionary for Digital Preservation Metadata'. In *Digital Preservation Metadata for Practitioners: Implementing PREMIS*, edited by Angela Dappert, Rebecca Squire Guenther, and Sébastien Peyrard, 23–36. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-43763-7_3.

Heseler, Jörg, Alexandra Büttner, and Matthias Arnold. 2024. *Grundlagen Der Digitalen Langzeitarchivierung: Eine Handreichung Zur Digitalen Langzeitarchivierung Aus Perspektive Der NFDI4Culture Community. Aufgabenbereich Datenpublikation Und Langzeitarchivierung (TA4)*. Arbeitsgruppe Digitale Publikationen und Daten (NFDI4Culture). https://doi.org/10.5281/zenodo.10423858.

'ISO 14721:2012. Open Archival Information System (OAIS) Reference Model'. 2012. ISO-Standard. https://web.archive.org/web/20231211011858/http://www.iso.org/standard/57284.html.

'ISO 16363:2012. Audit and Certification of Trustworthy Digital Repositories'. 2012. ISO-Standard. https://web.archive.org/web/20231223120511/https://www.iso.org/standard/56510.html.

'ISO/TC 20/SC 13 - Space Data and Information Transfer Systems'. 2021. ISO. 27 August 2021. https://web.archive.org/web/20231215060030/https://www.iso.org/committee/46612.html.

Leeuw, Lisa de, and Jonathan Crabtree. n.d. 'CoreTrustSeal Maintenance WG'. Resrach Data Alliance (RDA). Accessed 20 December 2023. https://web.archive.org/web/20230610213245/https://www.rd-alliance.org/groups/repository-audit-and-certification-dsa%E2%80%93wds-partnership-wg.html.

Lindlar, Michelle, Pia Rudnik, Sarah Jones, and Laurence Horton. 2020. '"You Say Potato, I Say Potato" Mapping Digital Preservation and Research Data Management

Concepts towards Collective Curation and Preservation Strategies'. *International Journal of Digital Curation* 15 (1): 26. https://doi.org/10.2218/ijdc.v15i1.728.

Miller, Bernhard, Nanette Rißler-Pipka, and Lukas Weimer. 2022. 'Infrastructure Networks within NFDI: Survey among Institutions Partnering with NFDI Consortia'. Zenodo. https://doi.org/10.5281/zenodo.6474644.

nestor. n.d. 'Nestor - Startseite'. Accessed 21 December 2023. https://web.archive.org/web/20231221110257/https://www.langzeitarchivierung.de/Webs/nestor/DE/Home/home_node.html.

'Nestor - AG OAIS Review'. 2022. Nestor – Kompetenznetzwerk Langzeitarchivierung Und Langzeitverfügbarkeit Digitaler Ressourcen in Deutschland e.V. 3 January 2022. https://web.archive.org/web/20230602143603/https://www.langzeitarchivierung.de/Webs/nestor/DE/Arbeitsgruppen/AG_OAIS_Review/ag_oais_review_node.html.

nestor working group Policy. 2014. 'Guidelines for the Creation of an Institutional Policy on Digital Preservation'. Frankfurt am Main : nestor – Network of Expertise in Long-Term Storage c/o Deutsche Nationalbibliothek. https://nbn-resolving.de/urn:nbn:de:0008-2014111006.

nestor-Arbeitsgruppe Policy. 2014. 'Leitfaden Zur Erstellung Einer Institutionellen Policy Zur Digitalen Langzeitarchivierung'. Frankfurt am Main : nestor - Kompetenznetzwerk Langzeitarchivierung c/o Deutsche Nationalbibliothek. https://nbn-resolving.de/urn:nbn:de:0008-2014052004.

NFDI. 2023. 'Konsortien | NFDI'. Verschiedene Disziplinen – Ein Ziel. 5 October 2023. https://web.archive.org/web/20231005183731/https://www.nfdi.de/konsortien/.

'Normenausschuss 009-00-15 AA Records Management und Langzeitverfügbarkeit von Informationsobjekten'. n.d. din-de. Accessed 21 December 2023. https://web.archive.org/web/20231221153546/https://www.din.de/de/mitwirken/normenausschuesse/nid/nationale-gremien/wdc-grem:din21:54774796.

Pinnick, Jaana. 2017. 'Exploring Digital Preservation Requirements: A Case Study from the National Geoscience Data Centre (NGDC)'. *Records Management Journal* 27 (2): 175–91. https://doi.org/10.1108/RMJ-04-2017-0009.

Pinnick, Jaana, Andrew Riddick, Robert McLaverty, and Garry Baker. 2018. 'A Case Study: Management and Exploitation of the Nuclear Decommissioning Agency Geoscience Data Archive'. *New Review of Information Networking* 23 (1–2): 99–110. https://doi.org/10.1080/13614576.2018.1544089.

PREMIS Editorial Committee. 2015. 'PREMIS'. https://web.archive.org/web/20231209121615/https://loc.gov/standards/premis/.

'PREMIS Maintenance Activity and Editorial Committee (Library of Congress)'. 2022. 1 February 2022. https://web.archive.org/web/20231227174834/https://www.loc.gov/standards/premis/ma.html.

Rat für Informationsinfrastrukturen. 2016. 'Leistung Aus Vielfalt. Empfehlungen Zu Strukturen, Prozessen Und Finanzierung Des Forschungsdatenmanagements in Deutschland'. Göttingen. urn:nbn:de:101:1-201606229098.

Rudnik, Pia, and Micky Lindlar. 2021. 'Foliensatz: Crashkurs Digitale Langzeitarchivierung - Ebenen-Modelle der LZA (3-Ebenen-Modell von K. Thibodeau & Levels of Digital Preservation der NDSA)'. February 26. https://doi.org/10.5281/zenodo.4761975.

Schoger, Astrid, and Christian Keitel. 2020. 'Nestor - AG Zertifizierung'. Nestor – Kompetenznetzwerk Langzeitarchivierung Und Langzeitverfügbarkeit Digitaler Ressourcen in Deutschland e.V. 12 February 2020. https://web.archive.org/web/20231220125132/https://www.langzeitarchivierung.de/Webs/nestor/DE/Arbeitsgruppen/AG_Zertifizierung/ag_zertifizierung_node.html.

SFA, Swiss Federal Archives. 2024. 'SIARD Suite'. 3 January 2024. https://web.archive.org/web/20240219162542/https://www.bar.admin.ch/bar/en/

home/archivierung/tools---hilfsmittel/siard-suite.html.

Shearer, Kathleen, Silvia Mirlene Nakano Koga, Eloy Rodrigues, Natalia Manola, martine Pronk, and Vanessa Proudman. 2023. 'Current State and Future Directions for Open Repositories in Europe'. Zenodo. https://doi.org/10.5281/zenodo.10255559.

Sprout, Bronwen, and Mark Jordan. 2018. 'Distributed Digital Preservation: Preserving Open Journal Systems Content in the PKP PN'. *Digital Library Perspectives* 34 (4): 246–61. https://doi.org/10.1108/DLP-11-2017-0043.

Taylor-Grant, Rebeccah, Laura Molloy, and Hea Lim Rhee. 2015. 'Archives and Records Professionals for Research Data IG'. RDA. 26 November 2015. https://web.archive.org/web/20240221074127/https://www.rd-alliance.org/groups/archives-records-professionals-for-research-data.html.

The Consultative Committee for Space Data Systems (CCSDS). 2012. 'Reference Model for an Open Archival Information System (OAIS). Recommended Practice CCSDS 650.0-M-2'. https://web.archive.org/web/20231230090114/https://public.ccsds.org/Pubs/650x0m2.pdf.

Thibodeau, Kenneth. 2002. 'Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years'. In *The State of Digital Preservation: An International Perspective*, 4–31. Washington, D.C.: Council on Library and Information Resources. https://web.archive.org/web/20230822111705/http://www.clir.org/wp-content/uploads/sites/6/pub107.pdf.

Weibulat, Tanja, Anna Lisa Schwartz, Stefan Seifert, Markus Schmalzl, and Dagmar Triebel. 2023. 'Was Haben Mehltaupilze Mit Digitaler Langzeitarchivierung Zu Tun?' July 25. https://doi.org/10.5281/zenodo.8268840.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. 'The FAIR Guiding Principles for Scientific Data Management and Stewardship'. *Scientific Data* 3 (1): 160018. https://doi.org/10.1038/sdata.2016.18.

Appendix

Survey (German)

Die Langzeitarchivierung (und –verfügbarkeit) von Forschungsdaten bedarf einer kontinuierlichen Aufmerksamkeit und Weiterentwicklung und gewinnt mit einem ständig überproportionalen Wachstum der Datenmenge sowie der Ausweitung auf immer neue Datentypen an Dringlichkeit, sei es durch Anforderungen von Förderern - die DFG verlangt 10 Jahre Aufbewahrung - oder der Aufgabe digitale Kultur- und Wissensobjekte auf unbestimmte Zeit zu speichern und verfügbar zu halten. Dabei ist zu unterscheiden zwischen dem Erhalt auf der Ebene des Bitstream und der funktionalen Erhaltung. Letzteres umfasst die Archivierung der digitalen (Daten-)Objekte in ihrem lesbaren, funktionalen, i.e. logischen, semantischen und relationalen Kontext, was je nach Datentyp und -format einen weitaus größeren Aufwand bedeutet als die reine Erhaltung des Bitstream. Mögliche Erhaltungsmaßnahmen sind z. B. Migration von Datenfiles in ein neues Dateiformat, Migration von Systemen oder Emulation der ursprünglichen Systemumgebung über 10 und mehr Jahre hinweg. Langzeitarchivierung strebt Erhalt sowohl auf Bitstream als auch auf funktionalem Level an.

Innerhalb der Sektion Common Infrastructure der NFDI hat sich die AG LTA formiert, die sich u.a. zum Ziel gesetzt hat:

- Bedarfe und Anforderungen der Konsortien an LZA-Services zu erheben
- Bestehende Service-Angebote und entwickelte Prozesse zur LZA zu dokumentieren
- Angebote zu vernetzen und Bedarfe und Angebote wo immer möglich zusammenzubringen (Broker)
- Übergreifende Services bedarfsorientiert zu entwickeln
- Geschäftsmodelle zu entwickeln und zu etablieren

Fragen zu **Angeboten** des Konsortiums

1. Bietet das Konsortium Langzeitarchivierungsdienst(e) im obigen Sinne an?

   Wenn ja:

   a. An wen richtet sich dieser Dienst, von wem kann/darf er genutzt werden?
   b. Gibt es Auswahlverfahren für Daten zur Langzeitarchivierung (collection policy)?
   c. Welche Archivierungsansätze (bitstream, funktionale Ebene) werden angeboten?
   d. Welche digitalen Objekte (digital entities) sind zur Archivierung vorgesehen?
   e. Gibt es zu diesem Dienst Geschäftsmodelle (Kosten, Verträge, …)?
   f. Hat der Dienst oder die anbietende Institution eine Preservation Policy?
   g. Wurde der Dienst zertifiziert (nestor-Siegel, CoreTrustSeal bzw. Data Seal of Approval)?
   h. Gibt es zur Nutzung des Archivs abgestimmte Prozesse und Rollenmodelle?
   i. Welche LZA Standards und Protokolle werden dabei berücksichtigt bzw. verwendet (OAIS, PREMIS, ggf. Community spezifische LZA-Standards, …)?
   j. Welche technischen Rahmenbedingungen gibt es zu berücksichtigen?
      i. Gibt es Spezialisierungen auf bestimmte Datentypen (Texte, Videos, Messdaten,…)?
      ii. Gibt es Begrenzungen auf z. B. spezifische Dateiformate?
      iii. Gibt es Begrenzungen der Datenmenge?

   Wenn nein:

   k. Ist ein solcher Dienst geplant?
   l. Nutzen Sie Angebote außerhalb des Konsortiums?
   m. Planen Sie, ein vorhandenes Repositorium zu einem LZA-Dienst auszubauen? Wenn ja, welches?

Fragen zu **Bedarfen** des Konsortiums

1. Besteht im Konsortium ein Bedarf nach einem LZA Dienst bzw. Ausbau eines vorhandenen LZA-Dienstes?

Wenn ja:

     a. Welche technischen Rahmenbedingungen gibt es zu berücksichtigen?
         i. Datentypen
        ii. Dateiformate
      iii. Datenmenge
     b. Welche anderen Rahmenbedingungen wie z.B. Zugangsbeschränkungen und Sicherheitsanforderungen sind zu beachten?

Survey (English)

Long Term Archiving (LTA) of research data needs continuous attention and development. Developing an approach and solutions regarding LTA becomes also increasingly urgent due to fast growth of data volume and use of new data types but also due to requirements of funders – DFG requires retention of data for 10 years – or due to the responsibility of storing and making cultural objects accessible for an indeterminate timespan. Preservation of data differs between preservation on the bitstream level and preservation on the functional level. The functional level contains archiving of data objects in their readable, functional, i. e. logical, semantic and relational context, which requires more effort than preservation on the bitstream level. Possible preservation actions are for example migration to new file formats, migration to new systems or emulation of system environments for 10 years or more. Long Term Archiving aims to preserve data on bitstream level as well as on the functional level.

Within the Section Common Infrastructure of NFDI the Working Group LTA was formed to work on:

· Collecting needs and requirements of consortia regarding a LTA-service

· Documenting existing services and developing processes for LTA

· Connecting services and needs where possible (Broker)

· Developing and establishing business models

Questions about **services** of the consortium:

1. Does the consortium offer a LTA-service (as described above)?

   If yes

     a. Who are the intended users of the service, who can/is allowed to use the service?
     b. Are data selected for long term preservation (collection policy)?
     c. Which preservation levels are offered (bitstream, functional level)?
     d. Which digital objects (digital entities) are intended for archiving?

e. Does the service have a business model (costs, contracts, …)?
f. Does the service or the institution where the service is located have a preservation policy?
g. Is the service certified (nestor seal, CoreTrustSeal, Data Seal of Approval)?
h. Are there processes and role models for use of the LTA-service?
i. Which LTA-Standards and protocols are used (OAIS, PREMIS, possibly community-specific LTA-standards, …)?
j. Are there technical factors that need to be taken into consideration?
    i. Specialization on specific data types (text, videos, measurement data,…)?
    ii. Limitation regarding e. g. specific file formats?
    iii. Limitations regarding data volume?

If no

k. Is the consortium planning a LZA-service?
l. Is the consortium using LTA-services outside the consortium?
m. Is the consortium planning to develop an existing repository towards a LTA-service? If yes, which repository?

Questions about the **needs** of the consortium

1. Is there a need in the consortium for a LTA-Service or further development and expansion of an existing LTA-Service?

If yes:

a. Which technical factors need to be taken into consideration?
    i. File types
    ii. File formats
    iii. Data amounts
b. Which other factors, like e. g. access restrictions and security demands need to be taken into account?

Imprint

Licensing