

Bausteine Forschungsdatenmanagement
Empfehlungen und Erfahrungsberichte für die Praxis von
Forschungsdatenmanager*innen

Langzeitarchivierung in der NFDI. Überblick und Perspektiven

Anna Lisa Schwartzⁱ

2025

Zitiervorschlag

Anna Lisa Schwartz. 2025. Langzeitarchivierung in der NFDI. Überblick und Perspektiven.
*Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die
Praxis von Forschungsdatenmanager*innen* Nr. 3/2025. DOI: [10.17192/bfdm.2025.3.8909](https://doi.org/10.17192/bfdm.2025.3.8909).

Dieser Beitrag steht unter einer
[Creative Commons Namensnennung 4.0 International Lizenz \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

ⁱORCID: [0000-0003-0391-3879](https://orcid.org/0000-0003-0391-3879)

Abstract

Die Nationale Forschungsdateninfrastruktur (NFDI) verfolgt seit 2020 das Ziel, Forschungsdaten systematisch zu erschließen, nachhaltig zu sichern und zugänglich zu machen. Langzeitarchivierung (LZA) spielt dabei eine zentrale, wenn auch heterogen umgesetzte Rolle. Dieser Beitrag untersucht anhand von Positionspapieren, wie die LZA als Aufgabe in der NFDI angegangen wird. Zunächst werden grundlegende Konzepte und Herausforderungen der LZA im NFDI-Kontext erläutert. Anschließend werden strategische Überlegungen zur LZA vor Förderbeginn sowie aktuelle Entwicklungen in den Konsortien und der Arbeitsgruppe Langzeitarchivierung dargestellt. Der Beitrag zeigt, dass trotz konzeptioneller Fortschritte die praktische Umsetzung der LZA in der NFDI weiterhin vor erheblichen Herausforderungen steht, insbesondere hinsichtlich der Heterogenität von Datentypen, fehlender einheitlicher Standards und unklarer Verantwortlichkeiten.

1 Forschungsdaten langfristig archivieren – eine Herausforderung für die NFDI

Seit 2020 wird in Deutschland der Aufbau der Nationalen Forschungsdateninfrastruktur (NFDI) durch die Deutsche Forschungsgemeinschaft (DFG) gefördert. Deren Ziel ist die Auffindbarkeit und Vernetzung von Forschungsdaten über Disziplinen hinweg. Die NFDI soll „Datenbestände von Wissenschaft und Forschung systematisch erschließen, nachhaltig sichern und zugänglich machen sowie (inter-)national vernetzen. Sie wird in einem aus der Wissenschaft getriebenen Prozess als vernetzte Struktur eigeninitiativ agierender Konsortien aufgebaut“ („Nationale Forschungsdateninfrastruktur“, 2025). Die Umsetzung dieser Ziele erfolgt in derzeit 26 Konsortien, einem Zusammenschluss verschiedener Institutionen und Initiativen eines oder mehrerer zusammenhängender Fachgebiete (NFDI, 2025b). Zusätzlich werden Basisdienste in Base4NFDI entwickelt und Querschnittsthemen in fünf Sektionen behandelt (NFDI, 2025a).

Während die NFDI mit ihren Aktivitäten und Sektionen interdisziplinär ausgerichtet ist, sind die Forschungsdaten der einzelnen Konsortien nicht notwendigerweise disziplinenübergreifend. Zu den großen Herausforderungen der NFDI zählt die Heterogenität der Daten und Konsortien. Einige vertreten Fachwissenschaften, die seit Jahrzehnten datengetrieben arbeiten, während in anderen Disziplinen erst ein Bewusstsein dafür geschaffen werden muss, was Forschungsdaten überhaupt sind (Blaser-Meier, 2019, S. 11–18). Unterschiedliche Datentypen und das Verständnis von Roh-, Primär- und Metadaten erschweren diese Aufgabe. All diese Aspekte sind für die nachhaltige Sicherung von Forschungsdaten relevant, wie sie als eines der Ziele der NFDI formuliert wurde („long-term data storage“) (Deutsche Forschungsgemeinschaft, 2024, S.2).

1.1 Begriffliche Grundlagen

Für das Verständnis der Langzeitarchivierungs-Strategien in der NFDI sind zunächst begriffliche Erläuterungen notwendig. Bei der digitalen Langzeitarchivierung handelt es sich bei dem zu archivierenden Objekt um ein digitales Objekt, was nach dem Open Archival Information System (OAIS) Standard definiert ist als „object composed of a set of bit sequences“ (Funk, 2010a, Kap. 7:3). Dieses Referenzmodell beschreibt den Aufbau eines digitalen Langzeitarchivs mit seinen verschiedenen Komponenten und deren Zusammenwirken mit dem Ziel des Informationserhalts über einen langen Zeitraum hinweg. Im OAIS-Modell werden drei Informationspakete unterschieden: ein Submission Information Package (SIP) mit den digitalen Informationen, die z.B. an ein Archiv abgegeben werden. Werden diese Informationen aufbereitet, z.B. durch Metadaten ergänzt, spricht man vom Archival Information Package (AIP), was im digitalen Langzeitarchiv verbleibt. Wird ein Teil des AIPs für spezifische Nutzengruppen verfügbar gemacht, entsteht das Dissemination Information Packages (DIP) (Brübach et al., 2010, Kap. 4:7–4:8).

Wie unterscheiden sich aber nun Langzeitspeicherung und -archivierung, zwei Begriffe, die oft und fälschlich synonym verwendet werden (Lindlar et al., 2020)? Erstens ist der zeitliche Faktor entscheidend: „Long Term is long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long Term may extend indefinitely“ (Funk, 2010b, Kap. 8:10). Zweitens gibt es verschiedene Erhaltungsstrategien, die die Speicherung (z.B. ein Backup von Forschungsdaten) von der Archivierung unterscheiden:

- **Bitstream Preservation:** Hier steht der Erhalt der physischen Information im Vordergrund, d.h. die Sicherung der Bitströme ohne Verlust (Funk, 2010c).
- **Content Preservation:** Sie sorgt für den Erhalt der logischen (logical preservation) und semantischen Struktur (semantic preservation) des digitalen Objekts, um die Nutzbarkeit und Interpretierbarkeit der Informationen langfristig sicherzustellen (Heseler et al., 2024, S. 82).

Zudem grenzt die LZA sich durch signifikante Eigenschaften (*significant properties*) ab. Das sind jene Merkmale eines digitalen Objekts, die notwendig sind, um seine Interpretierbarkeit auch zukünftig zu gewährleisten. (Heseler et al., 2024, S. 42.48).

Mit diesem begrifflichen Rahmen lässt sich nun untersuchen, wie Langzeitarchivierung in der strategischen Planung der NFDI konzeptualisiert wird. Hierfür wurden Positionspapiere vor und während des Aufbaus der Konsortien untersucht.

2 Langzeitarchivierung als Aufgabe in der NFDI

2.1 FAIR- und TRUST-Prinzipien

Die Umsetzung der *FAIR-Prinzipien* (Findable, Accessible, Interoperable, Reusable) (Wilkinson, Dumontier et al., 2016) ist wesentlicher Bestandteil der „Kultur des Daten-Teilens“, wie sie die NFDI aufbauen möchte (NFDI e.V., 2020). Die FAIR-Prinzipien betreffen verschiedene Aspekte der Langzeitverfügbarkeit von Forschungsdaten:

- **Findable:** Verwendung persistenter Identifikatoren (PIDs)
- **Accessible:** Dokumentation von Workflows und Zugangsprotokollen
- **Interoperable:** Nutzung standardisierter (Metadaten)Formate
- **Reusable:** Lizenzierung

Ergänzend zu den FAIR-Prinzipien wurden die *TRUST-Prinzipien* (Transparency, Responsibility, User focus, Sustainability, Technology) für vertrauenswürdige digitale Repositorien entwickelt (Lin, Crabtree et al., 2020). Während sich die FAIR-Prinzipien auf Daten beziehen, adressieren die TRUST-Prinzipien die Infrastrukturen selbst und betreffen damit direkt die LZA. In diesem Kontext sind das CoreTrustSeal (CoreTrustSeal Standards and Certification Board, 2025) und das *nestor-Siegel* relevant, das der Zertifizierung vertrauenswürdiger digitaler Archive dient und auf dem internationalen Standard ISO 16363 basiert (Group, 2025). Außerdem hat das Kompetenznetzwerk *nestor* wertvolle Beiträge für die LZA von Forschungsdaten geleistet, indem es z.B. relevante Metadatenstandards für verschiedene Disziplinen zusammengetragen hat (Neuroth et al., 2010, Kap. 6:1–6:18).

2.2 Strategische Empfehlungen vor und während der NFDI-Gründung

Bereits vor Gründung der NFDI widmeten sich Förderinstitutionen und wissenschaftliche Infrastrukturen in Deutschland in verschiedenen Papieren dem Thema Langzeitarchivierung. In den 2015 von der DFG herausgegebenen *Leitlinien zum Umgang mit Forschungsdaten* wurde ein Aufbewahrungszeitraum von 10 Jahren vorgeschlagen sowie die Archivierung in einer institutionellen, fachlich einschlägigen oder überregionalen Infrastruktur empfohlen (Deutsche Forschungsgemeinschaft, 2015). In den DFG-*Praxisregeln „Digitalisierung“* von 2022 ist die Unterscheidung zwischen der Langzeitsicherung (LZS) „in stabilen, migrationsfähigen Formaten auf einem technisch wie organisatorisch sicheren Speichersystem [...] (Bitstream Preservation)“ und der LZA, die darauf aufbaut und „noch weitreichendere technische und organisatorische Fragestellungen“ umfasst, ausschlaggebend (Altenhöner et al., 2022, S.35).

Der Rat für Informationsinfrastrukturen (RfII) empfahl 2016 in dem Papier *Leistung aus Vielfalt* den Aufbau einer nationalen Forschungsdateninfrastruktur und betonte die Notwendigkeit nachhaltiger Speicher- und Archivierungslösungen (RfII, 2016). Die Gemeinsame Wissenschaftskonferenz (GWK) formulierte in ihren Beschlüssen zur

NFDI ähnliche Empfehlungen und hob die Bedeutung von Langzeitarchivierung als Grundvoraussetzung für die Nachnutzbarkeit von Forschungsdaten hervor (Becker et al., 2018).

2019 vereinten sich mehrere Konsortialinitiativen in der *Berlin Declaration*, um gemeinsame übergreifende Themen wie Governance, Nutzendeneinbindung, rechtliche Aspekte und technische Infrastruktur in Arbeitsgruppen zu bearbeiten (Glöckner, Diepenbroek et al., 2019). Die *Leipzig-Berlin-Erklärung* von 2020 konkretisierte diese Handlungsfelder und schlug Umsetzungsmaßnahmen vor, wobei der nachhaltige Datenzugang und die Langzeitarchivierung zentrale Elemente der gemeinsamen technischen Infrastruktur bildeten (Bierwirth et al., 2020, S.4). Im Jahr 2021 entstand aus einem Workshop die NFDI-Sektion *Common Infrastructures*, in der u.a. eine Gesamtstrategie zur Langzeitarchivierung entwickelt wird (Diepenbroek et al., 2021, S.6).

Diese frühen Positionspapiere zeigen das Spannungsfeld, in dem sich die LZA in der NFDI bewegt: Einerseits wird ihre zentrale Bedeutung anerkannt, andererseits bleiben konkrete Umsetzungsstrategien, Verantwortlichkeiten und Finanzierungsmodelle offen.

3 Langzeitarchivierung in den NFDI-Konsortien und der Arbeitsgruppe LTA

3.1 Die Sektion *Common Infrastructures* und die Arbeitsgruppe LTA

Ein Ziel unter mehreren der Sektion *Common Infrastructures* ist die Entwicklung gemeinsamer Infrastrukturen für die NFDI, einschließlich Lösungen für die Langzeitarchivierung (Diepenbroek et al., 2021). Die Arbeitsgruppe *Long-term access and preservation* (LTA) wurde 2022 als Teil dieser Sektion etabliert, um LZA-Strategien über die Konsortien hinweg zu koordinieren. Eine zentrale Forderung besteht in der dauerhaften Archivierung von Forschungsdaten über die von der DFG geforderten 10 Jahre hinaus.

Die AG LTA führte 2024 eine Umfrage zu Tendenzen und Bedarfen der LZA unter den Konsortien durch. Die Ergebnisse zeigen verschiedene Problemfelder (Markus et al., 2024, S.5-13):

1. **Fehlende Gesamtstrategie:** gemeinsame Koordination der LZA-Strategie mit nachhaltigen Finanzierungsmodellen
2. **Heterogenität der Daten:** Vielzahl an Datentypen, für die es in den Fachcommunities selbst noch keinen Ansatz zur LZA gibt
3. **Fehlende Standards:** Einheitliche Formate, Metadatenschemata und Zertifizierungen
4. **Policies:** Fehlende Richtlinien (z.B. hinsichtlich Erhaltungsstrategien) für die LZA

5. **Rechtliche Aspekte:** Klärung von Zugriffs- und Nutzungsrechten
6. **Wissensaufbau als gemeinsame Basis:** Anbindung an Expertise aus dem Bereich der LZA, Aufbau einer einheitlichen Terminologie (v.a. bzgl. LZA/LZS und Erhaltungsstrategien, (siehe Tabelle Markus et al., 2024, S.7)).

Zusammenfassend sieht die AG LTA einen generellen Aufklärungsbedarf in der NFDI, die Notwendigkeit einer einheitlichen Terminologie und die auf Seiten der Konsortien liegende Verantwortung, Desiderate in Bezug auf die Langzeitarchivierung aus den Fachcommunities heraus auf die Ebene der NFDI zu tragen (Markus et al., 2024, S.4-12). Dieser Umfrage folgte 2025 ein überarbeitetes AG-Konzept, in dem die Notwendigkeit betont wird, sowohl technische als auch organisatorische Aspekte der LZA zu adressieren. Ein zentraler LZA-Service für die NFDI sei aufgrund bereits bestehender Angebote zum aktuellen Zeitpunkt jedoch nicht zielführend (Markus, Leinen & Stäcker, 2025). Ein Workshop zu „Long Term Data Issues“ (Markus, Leinen, Naumann & Valena, 2025) identifizierte weitere Herausforderungen für die LZA-Strategien in der NFDI.

4 Zusammenfassung und Perspektiven

Die Untersuchung der LZA-Strategien in der NFDI zeigt ein ambivalentes Bild: Einerseits ist ihre Bedeutung seit der Gründungsidee konzeptionell verankert. Andererseits stehen der praktischen Umsetzung erhebliche Herausforderungen entgegen.

Zentrale Erkenntnisse:

1. **Begriffliche Heterogenität:** Die Verwendung einer uneinheitlichen Terminologie erschwert die Kommunikation und Koordination zwischen und über die Konsortien hinaus.
2. **Zuständigkeiten:** Es fehlt an einer klaren Struktur hinsichtlich Verantwortlichkeiten, Zuständigkeiten und Schnittstellen zwischen den Konsortien und Archivierungseinrichtungen.
3. **Technische Heterogenität:** Die Vielzahl an Datentypen, Formaten und Nutzungsszenarien erfordert differenzierte LZA-Strategien.
4. **Nachhaltige Finanzierung:** Die langfristige Archivierung von Forschungsdaten benötigt langfristige Finanzierungsmodelle, über Projektlaufzeiten hinaus.

Perspektiven:

Die AG LTA der Sektion *Common Infrastructures* bietet eine Plattform für den notwendigen Austausch und die Koordination. Ihre Arbeit sollte sich auf folgende Schwerpunkte konzentrieren:

- **Standardisierung:** Identifizierung gemeinsamer Archivierungsstandards für Metadaten, Formate und Zertifizierungen unter Berücksichtigung disziplinspezifischer Anforderungen.

- **Governance-Strukturen:** Klärung von Verantwortlichkeiten und Anlaufstellen für die gemeinsame Erarbeitung nachhaltiger Archivierungsstrategien.
- **Kompetenzaufbau:** Aufbau von LZA-Kompetenzen in den Konsortien unter Anwendung von Best-Practices.
- **Pilotprojekte:** Erprobung konsortienübergreifender LZA-Lösungen.

Einhergehend mit den oben genannten Schwerpunkten könnte eine Meta-Analyse der Arbeitsprogramme der Konsortien dazu dienen, Schnittmengen im Bereich LZA zu identifizieren. Um ein konkretes Beispiel zu nennen: Fast alle Konsortien haben die Zertifizierung von Repositorien als Ziel in ihren Arbeitsprogrammen formuliert (Abb. 1). Eine solche Zertifizierung erfordert eine Erhaltungsplanung (*preservation planning*) (nestor Certification working group, 2025), die sich in eine Policy integrieren ließe. Eine solche ist in vielen Konsortien ebenfalls in Arbeit.

Dokumentname	Core Trust Seal	nestor Seal	ICSU World Data System
NFDI4Culture_2020	■	■	
NFDI4Memory_2022	■	■	
NFDI4Objects_2023			
Text+_2022	■	■	
Base4NFDI_2023			■
DAPHNE4NFDI_2023			
DataPLANT_2021		■	
FAIRagro_2023			
KonsortSWD_2020	■	■	
MaRDI_2022	■		
NFDI4Bio_2020	■	■	■
NFDI4Chem_2020	■	■	■
NFDI4Earth_2021	■	■	■
NFDI4energy_2022		■	
NFDI4Health_2021		■	
NFDI4Immuno_2021	■		
NFDI4Ing_2020		■	
NFDI4Microbiota_2023			■
NFDI4CS_2024			
PUNCH4NFDI_2020	■		

Abbildung 1: Verteilung der Kategorien Core Trust Seal, nestor Seal und ICSU WDS, zusammengestellt mit MAXQDA (unveröffentlichte Masterarbeit; Bildnachweis: Anna Lisa Schwartz)

Langzeitarchivierung in der NFDI ist kein eindimensionales Problem, sondern erfordert ein Zusammenspiel von technischen, organisatorischen, rechtlichen und finanziellen Lösungen. Die nächsten Jahre werden zeigen, ob es der NFDI gelingt, die konzeptionellen Grundlagen in tragfähige, nachhaltige Archivierungsinfrastrukturen zu überführen. Einen ersten Schritt in diese Richtung geht das an der ZB MED initiierte Projekt FAIR4ever, das ab Frühjahr 2026 startet und Bedarfserhebungen und eine detaillierte Umfeldanalyse zur LZA von Forschungsdaten in der NFDI durchführen möchte ("FAIR4ever", 2026)

Literaturverzeichnis

- Altenhöner, R., Berger, A., Bracht, C., Klimpel, P., Neuburger, T., Stäcker, T., & Stein, R. (2022). DFG-Praxisregeln „Digitalisierung“. Aktualisierte Fassung 2022. <https://doi.org/10.5281/zenodo.7435724>
- Becker, P.-C., Blümel, I., Bruch, C., Förster, A., Hahn, M., & Horstmann, W. (2018). Stellungnahme zum Aufbau einer Nationalen Forschungsdateninfrastruktur (NFDI). <https://doi.org/10.5281/zenodo.1405424>
- Bierwirth, M., Glöckner, F. O., Grimm, C., Schimmler, S., Boehm, F., Busse, C., Degkwitz, A., Koepler, O., & Neuroth, H. (2020). Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung. <https://doi.org/10.5281/zenodo.3895209>
- Blaser-Meier, S. (2019). Forschungsdatenmanagement in den Geisteswissenschaften. *Bibliothek Forschung und Praxis*, 43(1), 11–18. <https://doi.org/10.1515/bfp-2019-2002>
- Brübach, N., Queitsch, M., Liegmann, H., & Oßwald, A. (2010). Das Referenzmodell OAIS. In H. Neuroth, A. Oßwald, R. Scheffel, S. Strathmann & K. Huth (Hrsg.), *Nestor-Handbuch* (4:3–4:14).
- CoreTrustSeal Standards and Certification Board. (2025). CoreTrustSeal Trustworthy Digital Repositories Requirements 2026–2028: Extended Guidance. <https://doi.org/10.5281/zenodo.17659853>
- Deutsche Forschungsgemeinschaft. (2015). Leitlinien zum Umgang mit Forschungsdaten. Zugriff 28. Dezember 2025 unter <https://www.dfg.de/resource/blob/172112/leitlinien-forschungsdaten.pdf>
- Deutsche Forschungsgemeinschaft. (2024). Guidelines for Consortia National Research Data Infrastructure (NFDI). <https://www.dfg.de/resource/blob/168700/768e3541bb773bf86229d8a57e57db65/nfdi100-en-data.pdf>
- Diepenbroek, M., Schimmler, S., & Ebert, B. (2021). Sektionskonzept Common Infrastructures. <https://doi.org/10.5281/zenodo.5607490>
- FAIR4ever. (2026). <https://www.zbmed.de/forschen/laufende-projekte/fair4ever>
- Funk, S. E. (2010a). Digitale Objekte und Formate. In *Nestor-Handbuch* (7:3–7:8).
- Funk, S. E. (2010b). Migration und Emulation. In *Nestor-Handbuch* (8:10–8:15).

- Funk, S. E. (2010c). What Is Bit Preservation? Zugriff 28. Dezember 2025 unter <https://blogs.loc.gov/thesignal/2010/11/what-is-bit-preservation/>
- Glöckner, F. O., Diepenbroek, M., et al. (2019). Berlin Declaration on NFDI Cross-Cutting Topics. <https://doi.org/10.5281/zenodo.3457212>
- Group, n. C. W. (2025). Explanatory Notes on the nestor Seal for Trustworthy Digital Archives. Zugriff 28. Dezember 2025 unter <https://nbn-resolving.org/urn:nbn:de:0008-2501311511155.033954116333>
- Heseler, J., Kusch-Brandt, S., Neumann, J., & Rex, J. (2024). *Forschungsdatenmanagement: Eine Praxiseinführung*. Springer Vieweg.
- Lin, D., Crabtree, J., et al. (2020). The TRUST Principles for Digital Repositories. *Scientific Data*, 7, 144. <https://doi.org/10.1038/s41597-020-0486-7>
- Lindlar, M., Rudnik, P., Neuroth, H., & Strathmann, S. (2020). Begriffsbestimmungen: Langzeitverfügbarkeit, Langzeitarchivierung, Langzeitspeicherung. *Bausteine Forschungsdatenmanagement*, 1, 15–26. <https://doi.org/10.17192/bfdm.2020.1.8268>
- Markus, K., Leinen, P., Naumann, K., & Valena, P. (2025). Long Term Data Issues. <https://doi.org/10.5281/zenodo.15106332>
- Markus, K., Leinen, P., & Stäcker, T. (2025). Concept for Setting up an LTA Working Group. <https://doi.org/10.5281/zenodo.14939132>
- Markus, K., Naumann, K., Schmalzl, M., Watson, J., & Triebel, D. (2024). Langzeitarchivierung in der NFDI. <https://doi.org/10.5281/zenodo.10822613>
- Nationale Forschungsdateninfrastruktur. (2025). Zugriff 28. Dezember 2025 unter <https://www.dfg.de/de/foerderung/foerderinitiativen/nfdi>
- Neuroth, H., Oßwald, A., Scheffel, R., Strathmann, S., & Huth, K. (2010). Nestor-Handbuch. Zugriff 28. Dezember 2025 unter https://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf
- NFDI. (2025a). Gemeinsam zum Ziel: Zusammenarbeit in Sektionen. Zugriff 28. Dezember 2025 unter <https://www.nfdi.de/sektionen/>
- NFDI. (2025b). Konsortien. Verschiedene Disziplinen – ein Ziel. Zugriff 28. Dezember 2025 unter <https://www.nfdi.de/konsortien/>
- NFDI e.V. (2020). Satzung. Zugriff 28. Dezember 2025 unter <https://www.nfdi.de/wp-content/uploads/2021/05/Satzung-NFDI-eV.pdf>
- RfII. (2016). Leistung aus Vielfalt. Zugriff 28. Dezember 2025 unter <https://rfii.de/?p=1998>
- Wilkinson, M. D., Dumontier, M., et al. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>