

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

„Daten Archivieren“: Lebensabschnitt oder komplementärer Gesamtprozess? Gemeinsamkeiten und Unterschiede in Forschungsdatenmanagement und Langzeitarchivierung im deutschsprachigen Raum

Micky Lindlarⁱ

Natascha Schumannⁱⁱ

2025

Zitiervorschlag

Lindlar, Micky und Schumann, Natascha. 2025. „Daten Archivieren“: Lebensabschnitt oder komplementärer Gesamtprozess? Gemeinsamkeiten und Unterschiede in Forschungsdatenmanagement und Langzeitarchivierung im deutschsprachigen Raum. *Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanager*innen* Nr. 3/2025: S. 1-10. DOI: [10.17192/bfdm.2025.3.8850](https://doi.org/10.17192/bfdm.2025.3.8850).

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

ⁱTIB Leibniz-Informationszentrum Technik und Naturwissenschaften. ORCID: [0000-0003-3709-5608](https://orcid.org/0000-0003-3709-5608)

ⁱⁱGoethe Universität Frankfurt, hebis Verbundzentrale. ORCID: [0000-0001-7897-9616](https://orcid.org/0000-0001-7897-9616)

Abstract

Was ist digitale Langzeitarchivierung? Und wann findet sie statt? Zur Beantwortung dieser Fragen sowie zur Ermittlung von Möglichkeiten einer stärkeren Kooperation untersucht dieser Beitrag die Definition und Verortung von Langzeitarchivierung innerhalb der deutschsprachigen Communities anhand von zwei weit verbreiteten Werkzeugen: dem Datenlebenszyklusmodell sowie dem nestor Siegel Kriterienkatalog. Durch ein Mapping der nestor Siegel Kriterien auf Aufgaben der Lebensabschnitte wird aufgezeigt, dass Archivierung von Beginn an mitgedacht werden muss.

1 Einleitung

Forschungsdatenmanagement (FDM) und digitale Langzeitarchivierung (LZA) - wie passt das zusammen? In Deutschland beschäftigen sich Institutionen seit rund 20 Jahren mit der LZA. nestor hat sich seit den frühen 2000er Jahren als Netzwerk etabliert. Diesem verdanken wir viele, auch heute noch gebrauchte Definitionen, wie der, dass *„Langzeit keine Garantieerklärung über fünf oder fünfzig Jahre“* ist, *„sondern die verantwortliche Entwicklung von Strategien, die den beständigen, vom Informationsmarkt verursachten Wandel bewältigen können“* (Neuroth et al., 2016).

Erste Auseinandersetzungen mit dem FDM in Deutschland fanden ab ca. 2009 im Projekt „WissGrid“ statt, welches die „Langzeitarchivierung von Forschungsdaten“ als eines seiner Hauptaufgaben definierte (SUB Niedersächsische Staats- und Universitätsbibliothek Göttingen, 2009). Im Gegensatz zur nestor Definition beschreibt die FDM-Plattform forschungsdaten.info LZA im Glossar so: *„Das Ziel der Langzeitarchivierung ist es, den Zugang zu archivierten Daten über einen langen Zeitraum zu ermöglichen. Eine Aufbewahrungsfrist der Forschungsdaten für einen Zeitraum von mindestens zehn Jahren hat sich als Standard etabliert“* (forschungsdaten.info, 2025).

Die Gegenüberstellung der beiden Definitionen zeigt die Problematik der unterschiedlichen Nutzung von Begriffen in den Communities. Die Aufbewahrungsfrist für einen Zeitraum von mindestens 10 Jahren stammt aus den Leitlinien der DFG und bezieht sich vor allem darauf, dass Daten reproduzierbar und Aussagen nachprüfbar sind (Deutsche Forschungsgemeinschaft, 2025). LZA bezieht sich dagegen auf die langfristige Verfügbarkeit von Daten, die in der Regel keiner zeitlichen Einschränkung unterliegt. Aber wie wirken sich diese unterschiedlichen Definitionen auf die Praxis aus? Zur Beantwortung dieser Frage werden in diesem Beitrag Gemeinsamkeiten und Unterschiede der FDM und LZA Praxis im deutschsprachigen Raum anhand zweier Standardwerkzeuge betrachtet: des Datenlebenszyklus und des nestor Siegels. Neben dem Aufzeigen von Unterschieden sollen Möglichkeiten einer stärkeren Kooperation zwischen Data Stewards und „Digital Preservationists“ ausgelotet werden.

2 nestor und forschungsdaten.info als Spiegel ihrer Communities

Als aktuelle Spiegel der deutschsprachigen LZA- und FDM Communities können die Netzwerke nestor und forschungsdaten.info angesehen werden.

Seit 20 Jahren besteht mit nestor als Zusammenschluss von Bibliotheken, Archiven, Museen und wissenschaftlichen Einrichtungen ein Kompetenznetzwerk für LZA in Deutschland. Nach Auslauf der Förderung als kooperatives Netzwerk weitergeführt, ist nestor seit 2022 ein eingetragener Verein. In verschiedenen Arbeitsgruppen wird an Standards und Best Practices gearbeitet, deren Output allen Interessierten zur Verfügung steht¹. nestor wird sowohl national als auch international als Vertretung der deutschsprachigen LZA-Community wahrgenommen.

Eine ähnlich lang existierende Interessenvertretung wie nestor besteht für das deutschsprachige FDM (noch) nicht. Perspektivisch kann diese Lücke durch die Nationale Forschungsdaten Infrastruktur (NFDI) e.V. geschlossen werden. Aktuell findet aber ein länderübergreifender Austausch auf einzelnen Plattformen statt, aus welchen sich forschungsdaten.info² als einer der größten Player etabliert hat. Ursprünglich entstanden aus der baden-württembergischen Landesinitiative bwFDM, wurde es über mehrere Projektphasen weiterentwickelt und wird inzwischen von einem organisationsübergreifenden Redaktionsteam betreut. Partner der Plattform repräsentieren die FDM-Community des gesamten D-A-CH Raums.

3 Zyklisch oder linear? FDM und LZA zwischen Lifecycle und vertrauenswürdigen Langzeitarchiv

Die Gegenüberstellung der Definitionen eingangs zeigt, dass „Langzeitarchivierung“ unterschiedlich verstanden wird. Erschwerend kommt hinzu, dass verschiedene Begriffe für den Prozess der langfristigen Archivierung genutzt werden. So finden sich in deutschsprachigen Publikationen neben „Langzeitarchivierung“ auch „Langzeitverfügbarkeit“ oder „Langzeitsicherung digitaler Inhalte“ (Heseler et al., 2024). Auch im Englischen finden sich mehrere Begriffe: „digital preservation“, „long-term preservation“, „long-term archiving“. Während ein genereller Konsens darüber besteht, dass Digitalisierung und Backup nicht LZA bedeuten, sind die Grenzen an anderen Stellen nicht eindeutig. In der englischsprachigen LZA-Community findet im Gegensatz zur Deutschen eine klare Distanzierung vom Begriff des „archivings“ statt. So schreibt Pennock in 2024 „*Digital preservation is not the same as digital archiving*“, und begründet dies unter anderem mit Konnotationen, die der Archivierungsbegriff außerhalb von klassischen Archiven und Bibliotheken mit sich bringt.

¹https://www.langzeitarchivierung.de/Webs/nestor/DE/Publikationen/publikationen_node.html

²<https://forschungsdaten.info/>

Wenn sich ein gemeinsames Verständnis des übergeordneten Begriffs „Langzeitarchivierung“ als so schwierig erweist, wie sieht es dann mit den darunter aufgehängten Konzepten aus? Lindlar et al. verglichen hierzu 2020 weit verbreitete Konzepte der LZA- und FDM-Communities. Daraus resultierte die Erkenntnis, dass LZA ein granulareres Verständnis des digitalen Gegenstands voraussetzt, als FDM dies tut. Während beispielsweise Datenmanagementpläne (DMP) pauschal von „Daten“ sprechen, brechen LZA-Modelle wie PREMIS (Preservation Metadata Implementation Strategies) das digitale Objekt in verschiedene Ebenen herunter³. An diese granularen Ebenen angebunden sind in der LZA Konzepte wie Signifikante Eigenschaften (zu erhaltende Bestandteile und Funktionen des digitalen Objekts) und die Designated Community (Zielnutzungsgruppe) und ihre Erwartungen an das digitale Objekt. Eine Aufnahme dieser Konzepte in DMPs würde eine verbesserte Verzahnung von Datenproduzierenden, Forschungsdatenmanagement, Langzeitarchiven und Nutzenden darstellen (Lindlar et al., 2020).

Aber wann findet LZA im Rahmen des FDM statt? Zur Beantwortung dieser Fragen stellen wir zwei Werkzeuge der Communities gegenüber: Das Lebenszyklusmodell von forschungsdaten.info und das nestor Siegel (DIN 31644 Kriterien für vertrauenswürdige digitale Langzeitarchive).

Im FDM sind Datenlebenszyklusmodelle allgegenwärtig⁴ und dienen der Verortung von Aufgaben und Verantwortlichkeiten. Während sich Bezeichnungen einzelner Prozessschritte leicht unterscheiden, ist die Unterteilung in die Schritte Planung, Erhebung, Analyse, Zugänglichkeit, Archivierung, Nachnutzung allgemein gültig. So stellt „Daten archivieren“ auch im Datenlebenszyklus von forschungsdaten.info einen eigenen Schritt dar (siehe [Abbildung 1](#)).

Mit Ausnahme des gleichermaßen in LZA- und FDM-Communities genutzten DCC Lifecycle Models sind Lebenszyklusmodelle in der LZA-Community wenig verbreitet. Zwar ist Langzeitarchivierung auch insofern zyklisch, als dass sie beständig gegen neue technologische Entwicklungen geprüft und angepasst werden muss, jedoch sind gängige Modelle darauf ausgerichtet, die Prozesse und Funktionen des digitalen Langzeitarchivs zu beschreiben und dessen Vertrauenswürdigkeit zu messen. Das bekannteste Modell ist das OAIS (ISO 14721 Reference Model for an Open Archival Information System), welches das digitale Langzeitarchiv als Zusammenspiel von Menschen und Systemen in einer Organisation beschreibt, die Verantwortung für den langfristigen Erhalt von Information übernommen hat. Mit dem CoreTrustSeal (CTS)⁵, dem nestor Siegel und der ISO 16363⁶ existieren drei unterschiedlich granulare Zertifizierungsinstrumente für vertrauenswürdige Langzeitarchive. Zwar wird das nestor Siegel in der

³Für PREMIS: Intellectual Entity / Representation / File / Bitstream

⁴Beispielhaft seien hier das DCC Curation Lifecycle Model (Higgins, 2008) und der UK Data Service Research Data Lifecycle (UK Data Service, 2019) auf internationaler Ebene, sowie die Lebenszyklusmodelle der RWTH Aachen (Baur, 2021) oder des Leibniz Instituts für Altersforschung genannt (fli Leibniz Institut für Altersforschung, n.d.).

⁵<https://www.coretrustseal.org/>

⁶<http://www.iso16363.org/iso-certification/>

internationalen Literatur als Werkzeug zur Zertifizierung wahrgenommen, allerdings wird es in der FDM-Community wenig genutzt. Dort findet das CTS mehr Berücksichtigung, welches die Anforderungen der LZA in seinen 16 Kriterien im Vergleich zu den 34 Kriterien des nestor Siegels oberflächlich behandelt.

Tabelle 1: Übersicht der nestor Siegel-Kriterien (nestor-Arbeitsgruppe Zertifizierung, 2024)

nestor Siegel Kriterien	
K1	Auswahl der Informationsobjekte und ihrer Repräsentationen
K2	Verantwortung für den Erhalt
K3	Zielgruppen
K4	Zugang
K5	Interpretierbarkeit
K6	Rechtliche und vertragliche Basis
K7	Rechtskonformität
K8	Finanzierung
K9	Personal
K10	Organisation und Prozesse
K11	Erhaltungsmaßnahmen
K12	Krisen-/Nachfolgeregelung
K13	Signifikante Eigenschaften
K14	Integrität: Aufnahmeschnittstelle
K15	Integrität: Funktionen der Archivablage
K16	Integrität: Nutzerschnittstelle
K17	Authentizität: Aufnahme
K18	Authentizität: Erhaltungsmaßnahmen
K19	Authentizität: Nutzung
K20	Technische Hoheit
K21	Transferpakete
K22	Transformation der Transferpakete in Archivpakete
K23	Archivpakete
K24	Interpretierbarkeit der Archivpakete
K25	Transformation der Archivpakete in Nutzungspakete
K26	Nutzungspakete
K27	Identifizierung
K28	Beschreibende Metadaten
K29	Strukturelle Metadaten
K30	Technische Metadaten
K31	Protokollierung der Langzeiterhaltungsmaßnahmen
K32	Administrative Metadaten
K33	IT-Infrastruktur
K34	Sicherheit

Während also die Archivierung im Lebenszyklusmodell nur einen Lebensabschnitt darstellt, bricht das nestor Siegel den Gesamtprozess der Archivierung in 34 verschiedene zu erfüllende Kriterien herunter (siehe [Tabelle 1](#)). Hieraus resultiert die Frage, ob die

nestor Siegel Kriterien ausschließlich im „Daten archivieren“ Schritt des Lebenszyklusmodells zu verorten sind, oder auch andere Abschnitte berühren. Hilfreich ist in diesem Zusammenhang die Auflistung von einzelnen, den Abschnitten zugehörigen Unteraufgaben auf forschungsdaten.info (siehe Tabelle 2). „Daten archivieren“ beinhaltet demnach die Dateiformat- sowie Speichermedienmigration, die Erstellung von Backups, die Erstellung und Dokumentation von Metadaten sowie das eigentliche, nicht weitere beschriebene „Daten archivieren“. Ein Mapping dieser aufgeführten Schritte auf nestor Siegel Kriterien ist recht einfach möglich. So finden sich Anforderungen an den initialen Transfer von Daten auf geeignete Speicher (5.2) in den erläuternden Hinweisen (nestor-Arbeitsgruppe Zertifizierung, 2025) zu den Kriterien K14 „Integrität: Aufnahmeschnittstelle“, K21 „Transferpakete“ und K33 „IT-Infrastruktur“. Im Gegensatz dazu muss die Unteraufgabe „5.5 Daten archivieren“ zu allen Kriterien gemappt werden. Aber auch Unteraufgaben außerhalb des Lebensabschnitts „Daten archivieren“ enthalten für die LZA relevante Aufgaben. So verweist „1.3 Planung des Datenmanagements“ auf „Formate und Speicherort“ als relevante Punkte. Diese werden respektive in den nestor Kriterien K1-K3, K5-K7, K12 (siehe [Tabelle 1](#)) behandelt.

Gleiches gilt für Unteraufgaben aus allen anderen Lebensabschnitten (siehe [Tabelle 2](#) und [Abbildung 1](#)).

Tabelle 2: Mapping des forschungsdaten.info Lebenszyklus auf nestor Siegel Kriterien

Lebensabschnitt	Lebenszyklus-Unteraufgabe	nestor Siegel Kriterien
Forschungsvorhaben planen	1.1 Untersuchungsdesign unter dem Aspekt der Datenverarbeitung erstellen	NA
	1.2 Lokalisieren bereits vorhandener Daten	NA
	1.3 Planung des Datenmanagements (Formate, Speicherort etc.)	K1–K3, K5–K7, K12
	1.4 Zustimmungsprozedere für das Teilen von Daten planen und vorbereiten	NA
Daten erheben	2.1 Experimente, Beobachtungen, Messungen, Simulationen etc. durchführen	NA
	2.2 Erlaubnis zur Datennutzung ggf. einholen	K6, K20
Daten aufbereiten und analysieren	3.1 Daten eingeben, digitalisieren, transkribieren, übersetzen	K20, K21, K30
	3.2 Daten sichern und verwalten	K14, K17
	3.3 Daten prüfen, validieren, bereinigen	K5, K10
	3.4 Daten, wenn nötig, anonymisieren	K7
	3.5 Daten beschreiben	K5, K17, K29–K30, K32
	3.6 Daten interpretieren	NA
	3.7 Forschungsergebnisse ausgeben	NA
	3.8 Datenerhaltung vorbereiten	K1–K3, K13, K14, K17
Daten teilen und publizieren	4.1 Urheberrechte festlegen	K6, K7
	4.2 Zugang kontrollieren	K4, K33, K34
	4.3 Teilen der Daten	K4, K16, K19, K25, K26
	4.4 Daten bekannt machen	NA
Daten archivieren	5.1 Daten in geeignete Formate migrieren	K5, K13, K17, K18, K22, K24
	5.2 Daten auf geeignete Medien migrieren	K14, K21, K33
	5.3 Backups erstellen und Daten sichern	K12, K14, K15
	5.4 Metadaten erstellen und dokumentieren	K28–K32
	5.5 Daten archivieren	K1–K34
Daten nachnutzen	6.1 Prüfen der bisherigen Ergebnisse	K1, K10
	6.2 Weitere Untersuchungen unter Verwendung der Daten durchführen	NA
	6.3 Bisherige Forschung rezensieren	NA
	6.4 Lehren und Lernen	NA

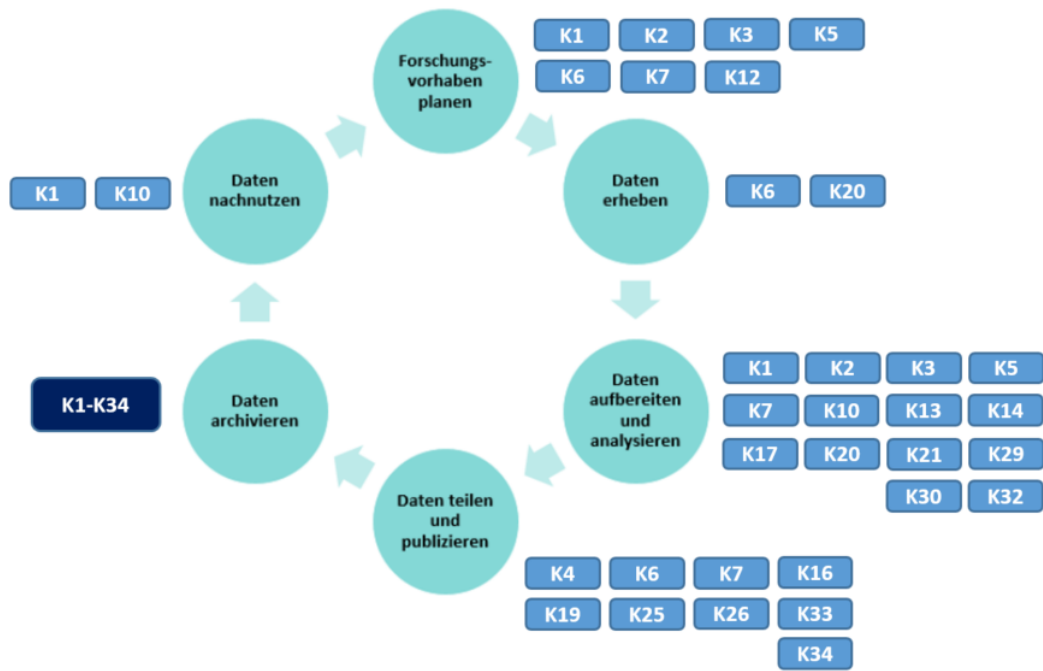


Abbildung 1: Der forschungsdaten.info Datenlebenszyklus 2024 mit darauf gemappten nestor Kriterien

4 Fazit

Der Abgleich zwischen dem forschungsdaten.info Lebenszyklusmodell und dem nestor Siegel als Werkzeuge der Communities FDM und LZA zeigt, dass sich die Aufgaben der LZA durch alle Schritte des FDMs ziehen. Archivierung ist somit nichts, was später passiert, sondern muss von Anfang an mitgedacht werden. Die von Pennock empfohlene Bezeichnung von „Digital Preservation“ in Abgrenzung zu „Digital Archiving“ ist im Deutschen aber insofern schwierig, als dass beide Begriffe mit digitaler (Langzeit-)Archivierung übersetzt werden können. Unterscheiden wir digitale Erhaltung von digitaler Archivierung, dann liegt der Unterschied augenscheinlich nur in der Dauer der Archivierung und nicht in einem aktiveren, früher einsetzenden Prozess. Steht der Prozess „Daten archivieren“ dann noch an der richtigen Stelle und hat er die korrekte Bezeichnung für den Lebensabschnitt? Zwar stimmt es, dass die Daten in diesem Lebensabschnitt an ein Langzeitarchiv übergeben werden können, allerdings sollten auch die weiter gefassten Anforderungen der LZA zumindest in den Unteraufgaben der anderen Lebensabschnitte ergänzt werden. Hier zeigt sich eine generelle Schwäche des Lebenszyklusmodells, welches die Anforderungen des wissenschaftlichen Arbeitens (z. B. „Forschungsvorhaben planen“, „Daten erheben“, „Daten aufbereiten und analysieren“) mit den Anforderungen einer Informationsinfrastruktur („Daten teilen und publizieren“, „Daten archivieren“) vermengt. Die Erwartung an Data Stewards ist es, beiden Anforderungen gleichermaßen gerecht zu werden. Hierbei

führt das Segmentieren in an den Lebensabschnitten orientierten Aufgaben allerdings zu fehlender Betrachtung der Einflussnahme der einzelnen Prozesse aufeinander, wie am Beispiel der LZA gezeigt. Das DCC Lifecycle Modell verfolgt hier schon seit vielen Jahren eine andere Darstellungsform, in welcher LZA („Curate“, „Preserve“, „Preservation Planning“) Kernaufgaben darstellen, die den kompletten Lebenszyklus umspannen. Die Aufgabe der Data Stewards ist ein sehr weit gefasstes Gebiet und am Beispiel LZA wird deutlich, dass es in der Umsetzung an der adäquaten Einbindung von Anforderungen aus verwandten informationswissenschaftlichen Fachgebieten mangelt. Daran wird deutlich, dass die Aufgaben von Data Stewards eher kooperativ zu lösen sind.

Neben der bereits angemerkten Ergänzung von LZA in DMPs erlaubt das im Paper vorgestellte Mapping auch die gezielte Identifikation von weiteren Kooperationsmöglichkeiten. Dies sei kurz am Beispiel von Punkt 2.1 „Erlaubnis zur Datennutzung einholen“ erläutert, der mit K6 und K20 (siehe [Tabelle 2](#)) verknüpft ist. K6 besagt, dass neben den Nutzungsbedingungen auch die Rechte, Pflichten und Kosten im Rahmen der LZA geregelt sein sollen. Im Kontext von Forschungsdatenarchivierung kann dies beispielsweise Pflichten zum Datenschutz und Schutzstufenkonzepte beinhalten, wenn personenbezogene Daten erhoben werden. Diese Dokumentation sollte als Provenienzinformation mit archiviert werden, was im Umkehrschluss zu einer verbesserten Datendokumentation und Interoperabilität führt, die wiederum in einer Qualitätssteigerung des archivierten Objekts resultiert.

Auf gleiche Weise lassen sich auch andere Anforderungen des nestor Kriterienkatalogs auf den Lebensdatenzyklus anwenden und sich vermeintliche Differenzen als Chancen zur stärkeren Kooperation nutzen.

Literaturverzeichnis

- Baur, J. (2021, 16. April). *FDM erklärt – Was ist eigentlich ein Daten-Lebenszyklus?* *Forschungsdaten – Aktuelles und Wissenswertes*. Zugriff 25. Mai 2025 unter <https://blog.rwth-aachen.de/forschungsdaten/2021/04/16/fdm-erklaert-was-ist-eigentlich-ein-daten-lebenszyklus/>
- Deutsche Forschungsgemeinschaft. (2025). *Guidelines for Safeguarding Good Research Practice. Code of Conduct*. <https://doi.org/10.5281/zenodo.3923601>
- fli Leibniz Institut für Altersforschung. (n.d.). *Forschungsdatenmanagement: Gute wissenschaftliche Praxis angewendet auf Daten*. Zugriff 25. Mai 2025 unter <https://www.leibniz-fli.de/de/forschung/gute-wissenschaftliche-praxis/fdm-am-fli>
- forschungsdaten.info. (2024, 9. September). *Der Datenlebenszyklus: Stationen des Forschungsdatenmanagements*. Zugriff 9. September 2025 unter <https://forschungsdaten.info/themen/informieren-und-planen/datenlebenszyklus/>

- forschungsdaten.info. (2025, 29. April). *Glossar: Langzeitarchivierung*. Zugriff 29. April 2025 unter <https://forschungsdaten.info/praxis-kompakt/glossar/#c269838>
- Heseler, J., Büttner, A., & Arnold, M. (2024). *Grundlagen der digitalen Langzeitarchivierung: Eine Handreichung zur digitalen Langzeitarchivierung aus Perspektive der NFDI4Culture Community*. NFDI4Culture. <https://doi.org/10.5281/zenodo.10281971>
- Higgins, S. (2008). The DCC Curation Lifecycle Model. *International Journal of Digital Curation*, 3(1), 134–140. <https://doi.org/10.2218/ijdc.v3i1.48>
- Lindlar, M., Rudnik, P., Horton, L., & Jones, S. (2020). "You say potato, I say potato Mapping Digital Preservation and Research Data Management Concepts towards Collective Curation and Preservation Strategies. <https://doi.org/10.5281/zenodo.3672773>
- nestor-Arbeitsgruppe Zertifizierung. (2024). *Erläuterungen zum nestor-Siegel für vertrauenswürdige digitale Langzeitarchive* (Nr. v2.2). <https://d-nb.info/1334021635/34>
- Neuroth, H., Oßwald, A., Scheffel, R., Strathmann, S., & Jehn, M. (Hrsg.). (2016). *nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung*. Hülsbusch. https://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf
- Pennock, M. (2024). Foreword: A Critical Reflection [Pre-Print]. In *Digital Preservation: A Critical Vocabulary*. <https://digital-preservation-a-critical-vocabulary.pubpub.org/pub/6imgiy3u>
- SUB Niedersächsische Staats- und Universitätsbibliothek Göttingen. (2009). *WissGrid—Projektdetails (Grid für die Wissenschaft—WissGrid)*. <https://www.sub.uni-goettingen.de/projekte-forschung/projektdetails/projekt/wissgrid/>
- UK Data Service. (2019, 13. August). *Research Data Lifecycle* [Video recording]. <https://www.youtube.com/watch?v=-wjFMMQD3UA>