

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

**Bewertung von (Forschungs-)Daten für die digitale
Langzeitarchivierung.
Ergebnisse aus der gemeinsamen Arbeitsgruppe
„Bewertungsmodelle“ für die Langzeitarchivierung an Hochschulen in
Hessen (LaVaH)ⁱ**

Elena Hamidy Christoph Marutschke Michael Höfel Dion Wagner
Judith Dähne Natascha Schumann Annegret Holtmann-Mares

2025

Zitiervorschlag

Hamidy, Elena, Marutschke, Christoph, Höfel, Michael, Wagner, Dion, Dähne, Judith, Schumann, Natascha und Annegret Holtmann-Mares. 2025. Bewertung von (Forschungs-)Daten für die digitale Langzeitarchivierung. Ergebnisse aus der gemeinsamen Arbeitsgruppe „Bewertungsmodelle“ für die Langzeitarchivierung an Hochschulen in Hessen (LaVaH). *Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanager*innen* Nr. 3/2025: S. 1-13. DOI: [10.17192/bfdm.2025.3.8849](https://doi.org/10.17192/bfdm.2025.3.8849).

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

ⁱElena Hamidy (ORCID: [0000-0001-6969-5742](https://orcid.org/0000-0001-6969-5742)), Christoph Marutschke (ORCID: [0000-0002-4810-6918](https://orcid.org/0000-0002-4810-6918)), Michael Höfel (ORCID: [0009-0009-2056-0914](https://orcid.org/0009-0009-2056-0914)), Dion Wagner (ORCID: [0009-0000-6958-0881](https://orcid.org/0009-0000-6958-0881)), Judith Dähne (ORCID: [0000-0003-3076-5074](https://orcid.org/0000-0003-3076-5074)), Natascha Schumann (ORCID: [0000-0001-7897-9616](https://orcid.org/0000-0001-7897-9616)), Annegret Holtmann-Mares (ORCID: [0009-0007-9116-405X](https://orcid.org/0009-0007-9116-405X))

Abstract

Der vorliegende Praxisbericht stellt einen Musterworkflow zur Bewertung von im wissenschaftlichen Kontext erzeugten Daten (z. B. Forschungsdaten, elektronische Pflichtablieferungen, (Retro-)Digitalisate u. v. m.) für die Langzeitarchivierung vor. Er gliedert sich in drei Abschnitte: (1) eine rechtlich-organisatorische Prüfung, (2) eine technische Prüfung sowie (3) eine Kontext- und Strukturprüfung. Praktische Ablaufdiagramme und eine kompakte Checkliste können mit der Langzeitarchivierung (LZA) befassten Personen in Kultureinrichtungen den mitunter ausgesprochen komplexen Bewertungs- und Übernahmeprozess erleichtern. Die Arbeitsgruppe „Bewertungsmodelle“, die sich aus Vertreter*innen der Bereiche Hochschulbibliothek, Hochschularchiv und Forschungsdatenmanagement verschiedener hessischer Hochschulen zusammensetzt, möchte damit einen Impuls zur spartenübergreifenden Diskussion zum komplexen Thema Bewertung leisten.

1 Einleitung

Digitale Langzeitarchivierung meint die zuverlässige Verfügbarhaltung von Daten – mitsamt allen Kontextinformationen und Metadaten – über einen nicht befristeten Zeitraum mit dem Ziel ihrer dauerhaften Nutzungsmöglichkeit. Sie geht über die für die gute wissenschaftliche Praxis oftmals genannte Aufbewahrungsfrist von zehn Jahren (z. B. Deutsche Forschungsgemeinschaft, 2025) und die bloße Speicherung in einem Repositorium hinaus. Auch ist sie nicht mit einer redundanten Datensicherung (umgangssprachlich „Backup“) gleichzusetzen. Langzeitarchivierung ist ein aufwändiger, kontinuierlicher Prozess mit laufenden hohen Kosten, sodass eine Bewertung von im wissenschaftlichen Kontext erzeugten Daten unumgänglich ist. Daten werden in diesem Beitrag weiter gefasst als Forschungsdaten, sie können beinhalten Ressourcen, die im weitesten Sinne im wissenschaftlichen Kontext anfallen. Neben Forschungsdaten sind auch Artikel, Monografien, digitalisierte Bibliotheksbestände und weitere gemeint. Wird auf Daten speziell im Archivgewesen eingegangen, dann wird der Begriff „Unterlagen“ erwähnt, die ebenfalls eine Teilmenge von Daten gesehen werden. Die mit der Bewertung einhergehende Auswahl hängt von vielen Faktoren ab, die sich gegenseitig beeinflussen können und auch nicht für jedes Langzeitarchiv gleich sind. Insbesondere bei Forschungsdaten, die sich inhaltlich sehr heterogen und fachspezifisch darstellen, ist es schwierig zu entscheiden, welche für eine Langzeitarchivierung in Frage kommen – und welche nicht.

Dieser Praxisbericht zur Bewertung digitaler Daten im Hinblick auf ihre Eignung für eine Langzeitarchivierung ist in einer gemeinsamen Arbeitsgruppe „Bewertungsmodelle“ der Kooperation „Digitales Archiv der Hessischen Hochschulen“ (DAHH)¹ und des Projekts „Langzeitverfügbarkeit an Hessischen Hochschulen“ (LaVaH)² sowie in Ab-

¹<https://www.dahh.de/> (zuletzt geprüft am 20.05.2025).

²<https://www.lavah.de/> (zuletzt geprüft am 20.05.2025).

stimmung mit den „Hessischen Forschungsdateninfrastrukturen“ (HeFDI)³, der hessischen Landesinitiative für Forschungsdatenmanagement, entstanden. DAHH ist eine Kooperation der hessischen Hochschularchive, die mit einer gemeinsamen Infrastruktur die Langzeitarchivierung digitaler Unterlagen ermöglicht und die Digitalisierung von Verwaltungsprozessen sowie die fachliche und kollegiale Zusammenarbeit stärkt. In LaVaH arbeiten die hessischen Universitäten und Hochschulen für angewandte Wissenschaften gemeinsam an der arbeitsteiligen Langzeitarchivierung digitaler Objekte. Hier sind es vor allem die Bibliotheken, die für die Vorbereitung der Daten und damit auch für die Auswahl zuständig sind. Das bedeutet, dass es Unterschiede geben kann, welche Institution jeweils die abliefernde und welche die langzeitarchivierende ist.

Ausgangspunkt für diesen Bericht war die Idee, mit der digitalen Langzeitarchivierung betrauten Personen in kulturellen und wissenschaftlichen Institutionen (z. B. Archiven, Bibliotheken, Hochschulen und außeruniversitären Forschungseinrichtungen) praktikable Kriterien an die Hand zu geben, um transparent über eine (Nicht-)Archivierung von Daten entscheiden zu können. Die Frage nach einem Informationsgehalt, der eine unbefristete Aufbewahrung von Daten rechtfertigt oder gar notwendig macht – im Archivwesen ist diesbezüglich von der „Archivwürdigkeit“ die Rede, die im Rahmen der sogenannten archivischen „Bewertung“ festgestellt wird (z. B. Menne-Haritz, 2011, S. 48 u. 59) –, ist angesichts des breiten Spektrums an Speicherformaten und Fachgebieten mit zum Teil eigenen Standards in Datenerhebung und -organisation sowie der mitunter sehr großen Datenvolumina nicht einfach zu klären. Folglich will (und kann) der vorliegende Praxisbericht kein detailliert ausdifferenziertes Patentrezept vorlegen, das jeden noch so speziellen Einzelfall abdeckt – ein solches Vorhaben müsste zwangsläufig schon an der oben beschriebenen disziplinären und institutionellen Diversität der Autor*innengruppe des vorliegenden Beitrags scheitern. Vielmehr ist ein Musterworkflow entstanden, der im Hinblick auf die mit Auswahl und LZA betraute(n) Institution(en) wie auch auf die Art der zu archivierenden Daten (Forschungsdaten, digitaler Nach- bzw. Vorlass etc.) neutral formuliert ist. Dies geht nicht ohne Kompromisse: So verzichtet der Workflow sowohl auf organisatorisch-strukturelle Vorgaben, welche Stelle innerhalb der betreffenden Institution welche Prozessschritte wahrnehmen sollte, als auch auf technische Spezifizierungen.

2 Musterworkflow zur Bewertung von Daten

Der im Folgenden beschriebene Bewertungsworkflow gliedert sich in drei Abschnitte: (1) die rechtlich-organisatorische, (2) die technische sowie (3) die Kontext- und Strukturprüfung. Die Erfahrung zeigt, dass, unabhängig von der organisatorischen und technischen Realisierung der Langzeitarchivierung, mehrere der folgenden Kriterien nur im direkten Austausch zwischen Datengeber*innen und Langzeitarchiv zu klären sind. Dieser Dialog kann formlos (z. B. per E-Mail) oder über vorgeschriebene Kanäle

³<https://www.hefdi.de/> (zuletzt geprüft am 20.05.2025).

(z. B. Online-Formulare) erfolgen, sollte aber im Sinne einer Arbeitserleichterung entsprechend zweckmäßig durch das Langzeitarchiv gestaltet werden. Im Rahmen der Übergabe- und Bewertungskommunikation unbedingt zu klären sind auch die Übergabemodalitäten (siehe [Abschnitt 2.1](#)).

2.1 Rechtlich-organisatorische Prüfung

Zuerst ist abzuklären, ob eine Langzeitarchivierung rechtlich möglich und zulässig ist (Schritt 1, siehe [Abbildung 1 im Anhang](#)). Bei personenbezogenen Daten muss gemäß den geltenden datenschutzrechtlichen Bestimmungen abgewogen werden, ob eine Archivierung nicht anonymisierter Daten zweckmäßig und alternativlos ist. Im Falle urheberrechtlich geschützter Daten(sammlungen) muss die abgebende Stelle verfassungsberechtigt sein, um der langzeitarchivierenden Einrichtung die notwendigen Nutzungs- und Verwertungsrechte einräumen zu können. Fällt die Prüfung negativ aus, ist dies entsprechend zu dokumentieren.

Ein Auftrag zur Archivierung (Schritt 2) wird durch entsprechende Vereinbarungen mit Personen und/oder Organisationen begründet. Diese sollten stets schriftlich erfolgen und können sich an entsprechenden Musterverträgen orientieren. Im Sinne der Bewertung ist ferner darauf zu achten, Art und Umfang der durchzuführenden Erhaltungsmaßnahmen wie z. B. Migrationen, Validierungen etc. zu vereinbaren.

Für die inhaltliche Ausrichtung ist die Bedeutung einer möglichst klar umrissenen Sammlungspolitik (Schritt 3), beispielsweise in Form eines Sammlungsprofils, zu unterstreichen. Für den konsistenten Aufbau eines Langzeitarchivs sind diese Vorgaben und ihre Veränderung über die Zeit zu dokumentieren.

Die besondere Bedeutung und/oder Aktualität von Daten (Schritt 4) ist ein zugegeben unscharfes Kriterium, das den Entstehungskontext der betreffenden Daten berücksichtigt, aus dem sich zukünftig ein bleibender Wert ergeben kann, z. B. Daten zur Migrationspolitik, die erst durch bestimmte Ereignisse in den Fokus rücken können.

Ein Langzeitarchiv stellt kein Zwischenarchiv oder gar eine Arbeitsablage dar. Folglich sind keine Daten zu übernehmen, für die lediglich eine temporäre Aufbewahrung angestrebt bzw. gewünscht wird (Schritt 5) oder die zum Zeitpunkt der Übergabe ans Langzeitarchiv noch bearbeitet und sich potentiell oder definitiv verändern werden (z. B. weil das betreffende Forschungsprojekt noch nicht abgeschlossen ist) (Schritt 6, siehe [Abbildung 2 im Anhang](#)). In solchen Fällen kann der*die Datengeber*in an ein institutionelles oder fachliches Repositorium verwiesen werden, das die Aufbewahrung für einen befristeten Zeitraum vorsieht.

Können die Daten nach Ermessen der langzeitarchivierenden Einrichtung von der Designated Community voraussichtlich nachgenutzt werden (Schritt 7)? Diese Frage lässt sich – abhängig von der Art der betreffenden Daten – beliebig kompliziert beantworten und wird in der Literatur unter dem Stichpunkt „Signifikante Eigenschaften“ diskutiert

(z. B. nestor-Arbeitsgruppe Digitale Bestandserhaltung, 2012; SLUB, 2022). Einstiegsfragen könnten sein: Sind potentielle Nachnutzungsszenarien denk- und formulierbar? Ist eine Nachnutzung der Daten zu erwarten? Nachnutzbarkeit kann nicht zuletzt von einem vorausschauenden (Forschungs-)Datenmanagement, das bereits mit der Planung der Datenerhebung einsetzt, erheblich profitieren⁴; umgekehrt kann nachlässiges Datenmanagement die spätere Bewertung, Übernahme, Langzeitarchivierung und Bereitstellung erschweren. Im Falle personenbezogener Daten ist zu klären, ob eine Übernahme in anonymisierter bzw. pseudonymisierter Form ausreichend ist.

Schließlich muss geklärt werden, ob die Daten ohne Weiteres reproduziert werden können oder nicht (Schritt 8). Damit ist nicht die Erstellung physischer Kopien der Daten gemeint (z. B. im Rahmen von Backups), sondern die Frage, ob eine Datenerhebung nach identischen Erhebungskriterien zu einem späteren Zeitpunkt identische Ergebnisse brächte (z. B. Klimamessungen).

Kommt die Bewertung zum Ergebnis, dass eine dauerhafte Sicherung der Daten anzustreben ist, sind die Übergabemodalitäten abzuklären. Diese schließen den Übernahmekanal (z. B. Cloud, externer Datenträger), die Dateiformate, in denen die Daten bereitzustellen sind, die mitzuliefernden Metadaten und Kontextinformationen, (z.B. Anwendungsprogramme, Betriebssysteme, Codierungen) eine Übersicht der Daten sowie Hashwerte zur Prüfung der Datenintegrität nach der Übernahme mit ein.

2.2 Technische Prüfung

Im Anschluss an die Übernahme sollen folgende Prüfschritte sicherstellen, dass eine Langzeitarchivierung der Daten aus technischer Perspektive möglich ist.⁵

Zum Schutz des Langzeitarchivs werden die Daten auf eine mögliche Infektion mit Malware (z. B. Viren) geprüft (Schritt 9, siehe [Abbildung 3 im Anhang](#)). Fällt die Prüfung positiv aus, müssen die bereitgestellten Daten abgelehnt und gelöscht werden.

Anhand der vorab vereinbarten Übergabemodalitäten (Hashwerte, Liste, Form des Transferpakets) wird geprüft, ob die Daten und Metadaten vollständig und korrekt sind (Schritt 10).

Um die Daten einer sinnvollen technischen und inhaltlichen Prüfung unterziehen sowie ihre künftige Bearbeitung gewährleisten zu können, müssen sie unverschlüsselt vorliegen (Schritt 11).

Das Langzeitarchiv hat in der Regel Vorgaben bezüglich Dateigrößen und akzeptierter Datenformate. Systembedingt können häufig nur Dateien bis zu einer bestimmten

⁴Für den Bereich der Hochschulen für angewandte Wissenschaften siehe (Blümm et al., 2023).

⁵Einen detailliert ausformulierten Bewertungsprozess für digitale Dateisammlungen im archivischen Kontext bieten Birn, M. & Naumann, K. (2019). Erfahrungsberichte zu kostenpflichtigen und -freien Softwares zur archivischen Bewertung schwach strukturierter Dateisammlungen sind in Puchta & Naumann (2017) enthalten.

Größe fehlerfrei verarbeitet werden, ferner können kostenbedingt durch das Langzeitarchiv festgelegte Maximalgrößen eingereicherter Daten existieren (Schritt 12, siehe [Abbildung 4 im Anhang](#)). Wenn die Dateigröße überschritten wird, ist zu prüfen, ob in der Lieferung ggf. vorhandene Dubletten entfernt werden können und/oder ob durch eine Konvertierung eine Datenintegrität und -authentizität nicht gefährdende Reduktion der Größe erreicht werden kann; auch eine Aufteilung der Daten in mehrere Pakete ist denkbar.

Wenn nicht empfohlene Formate vorliegen, kann eine Formatmigration notwendig sein, damit langzeitinstabile Formate durch langzeitarchivfähige ersetzt werden.⁶ In besonderen Fällen kann das Archiv nach interner Beratung auch andere als die geforderten Formate akzeptieren (Schritt 13).

Zuletzt werden die Daten und Metadaten auf ihre Validität, z. B. hinsichtlich der Dateiformatspezifikation, geprüft (Schritt 14).

Werden bei den Schritten 9 bis 11 Probleme festgestellt, ist zu klären, ob die Daten erneut korrekt, vollständig, malwarefrei und/oder unverschlüsselt übermittelt werden können, andernfalls sind die bereitgestellten Daten zu löschen. Sind in den Schritten 12 bis 14 Eingriffe in die Datenintegrität notwendig, sind diese durch das Langzeitarchiv zu dokumentieren.

2.3 Kontext- und Strukturprüfung

Damit archivierte Daten sinnvoll (nach)genutzt werden können, ist eine möglichst präzise Beschreibung durch Metadaten unerlässlich (Schritt 15, siehe [Abbildung 5 im Anhang](#)). Je nach Art der Daten kann dies z. B. durch ein Code-Book, ein ReadMe und/oder ein verbindliches Kernset an Metadaten (z. B. LaVaH, 2025) gewährleistet werden.

Der Begriff der Kontextinformationen ist noch etwas weiter gefasst als jener der Metadaten (Schritt 16). Hierzu zählen etwa Hilfsmittel, die für Verständnis, Ordnung, Nutzung und Erhaltung der Daten notwendig sind (z. B. Informationen zu Codierungen, Softwareanforderungen, Datenbankstrukturen, Versionen).

Schließlich wird noch geprüft, ob die Daten eine übersichtliche und dokumentierte Struktur aufweisen, da verschachtelte Ordnungen die erneute Nutzung erschweren

⁶Exemplarisch genannt werden sollen an dieser Stelle der *Katalog archivischer Dateiformate* der Koordinationsstelle für die dauerhafte Archivierung elektronischer Unterlagen (KOST) (https://kost-ceco.ch/cms/kad_main_de.html), das *Recommended Formats Statement* der Library of Congress (<https://www.loc.gov/preservation/resources/rfs/>) sowie die *International Comparison of Recommended File Formats* der Open Preservation Foundation (OPF) (<https://openpreservation.org/resources/member-groups/international-comparison-of-recommended-file-formats/>). Für HeFDI siehe Cordes et al., 2024, S. 26–28. Alle Hyperlinks zuletzt geprüft am 20.05.2025.

können (Schritt 17). Für diese drei Schritte gilt jeweils die Frage, ob der*die Datengeber*in die gewünschten Informationen nachliefern beziehungsweise die Struktur vereinfachen kann. Dabei geht es um Rücksprache bei leicht zu behebenden Fehlern.

3 Fazit und Ausblick

Die Arbeitsgruppe möchte mit diesem Bericht zur spartenübergreifenden Diskussion zum komplexen Thema Bewertung – insbesondere im deutschsprachigen Raum – beitragen.⁷ Die Checkliste ist als Arbeitsmittel für die Nachnutzung gedacht und wird begleitend in bearbeitbarer Form veröffentlicht (Wagner, 2025). Das soll interessierten Einrichtungen die Möglichkeit geben, diese für ihre Zwecke und speziellen Bedürfnisse anzupassen.

Literaturverzeichnis

- Birn, M. & Naumann, K. (2019). Bewertung schwach strukturierter Unterlagen. *Brandenburgische Archive*, 36, 8–14.
- Blümm, M., Burkart, C., Chlastak, M., Fritsch, K., Neuroth, H., Schieberle, A., Schmunk, S., & Werner, S. (2023). *Datenmanagementpläne an Fachhochschulen / Hochschulen für Angewandte Wissenschaften: eine Bestandsaufnahme und Empfehlungen*.
<https://doi.org/10.5281/zenodo.8275739>
- Cordes, B., Pietsch, A., Dähne, J., Geissner, A., Jordan, S., Richter, T., & Trebels, B. (2024). Kurationsempfehlungen für HeFDI-Repositorien: Handreichung und Erläuterung zur Kuration eingereicher Datensätze in den HeFDI-Repositorien.
<https://doi.org/10.5281/ZENODO.14442045>
- Deutsche Forschungsgemeinschaft. (2025). Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Zugriff 22. Mai 2025 unter
<https://doi.org/10.5281/zenodo.14281892>
- LaVaH. (2025). LaVaH Metadaten Kernset.
<https://doi.org/10.5281/zenodo.15064874>.
- Menne-Haritz, A. (2011). *Schlüsselbegriffe der Archivterminologie: Lehrmaterialien für das Fach Archivwissenschaft*. (Nachdruck der 3., durchgesehenen Auflage). Archivschule Marburg.
- nestor-Arbeitsgruppe Digitale Bestandserhaltung. (2012). *nestor-materialien 15: Leitfaden zur digitalen Bestandserhaltung. Vorgehensmodell und Umsetzung, Version 2.0*. <https://nbn-resolving.de/urn:nbn:de:0008-2012092400>

⁷In diesem Kontext sei auf die nestor-Entwurfsgruppe zur Bewertung von Forschungsdaten hingewiesen, die sich zum Zeitpunkt der Einreichung des vorliegenden Beitrags in Gründung befindet: https://www.langzeitarchivierung.de/Webs/nestor/DE/Arbeitsgruppen/AG_Archivstandard/2025ausschreibungArchivstandard.html (zuletzt geprüft am 20.05.2025).

- Puchta, M., & Naumann, K. (Hrsg.). (2017). *Kreative digitale Ablagen und die Archive: Ergebnisse eines Workshops des KLA-Ausschusses Digitale Archive am 22./23.11.2016 in der Generaldirektion der Staatlichen Archive Bayerns*. Generaldirektion der staatlichen Archive Bayerns.
- SLUB. (2022). *Signifikante Eigenschaften im SLUB Langzeitarchiv. Diskussionspapier*. https://slubarchiv.slub-dresden.de/fileadmin/groups/slubsite/slubarchiv/SLUBArchiv_Diskussionspapier_Signifikante_Eigenschaften_v1.2.pdf
- Wagner, D. (2025). LaVaH - Checklisten zur Bewertung von (Forschungs-)Daten für die LZA. <https://doi.org/10.48718/4wct-1x42>

Abbildung 1: Rechtlich-organisatorische Prüfung Schritte 1-5

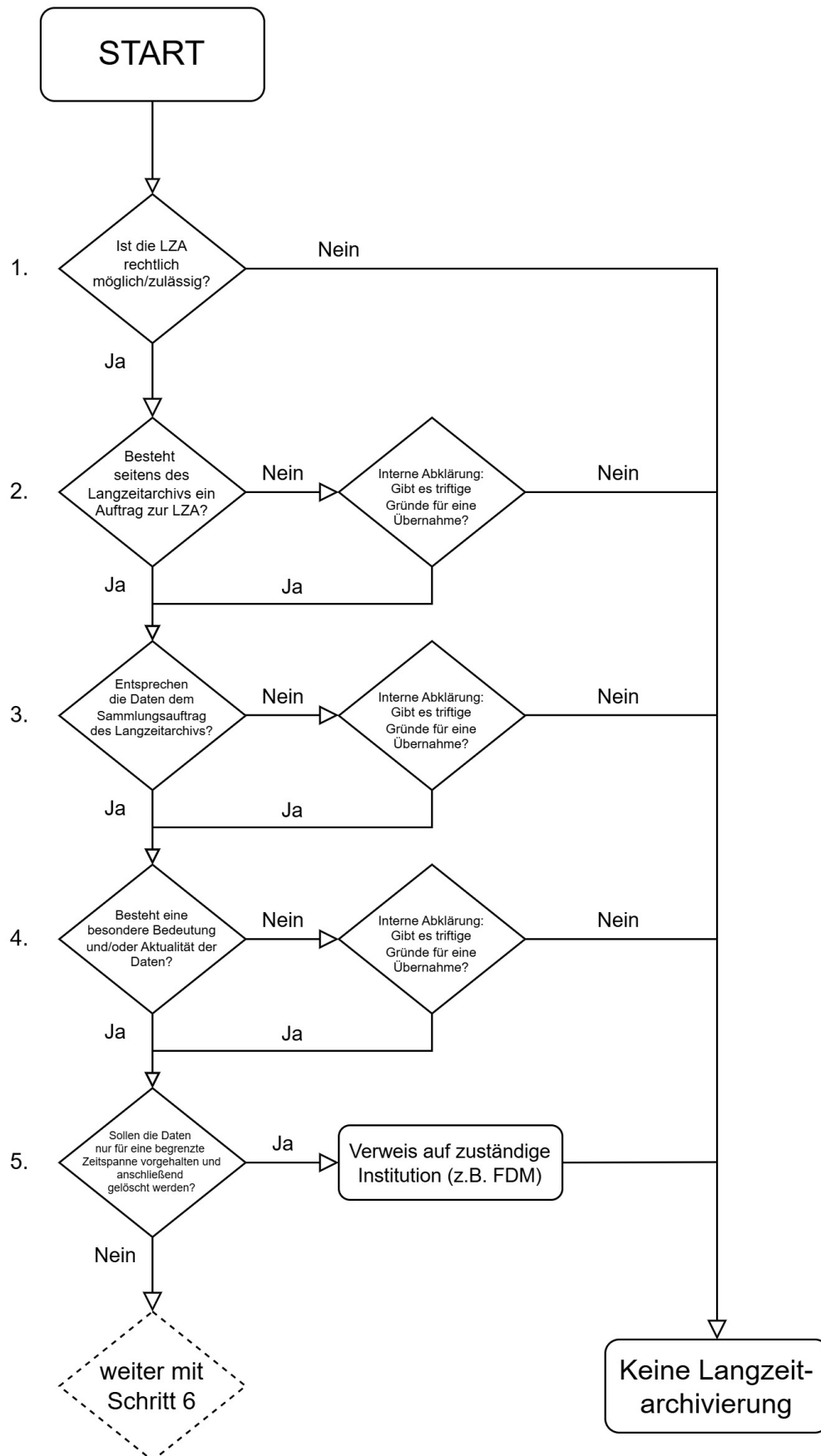


Abbildung 2: Rechtlich-organisatorische Prüfung Schritt 6 bis zum Abschluss

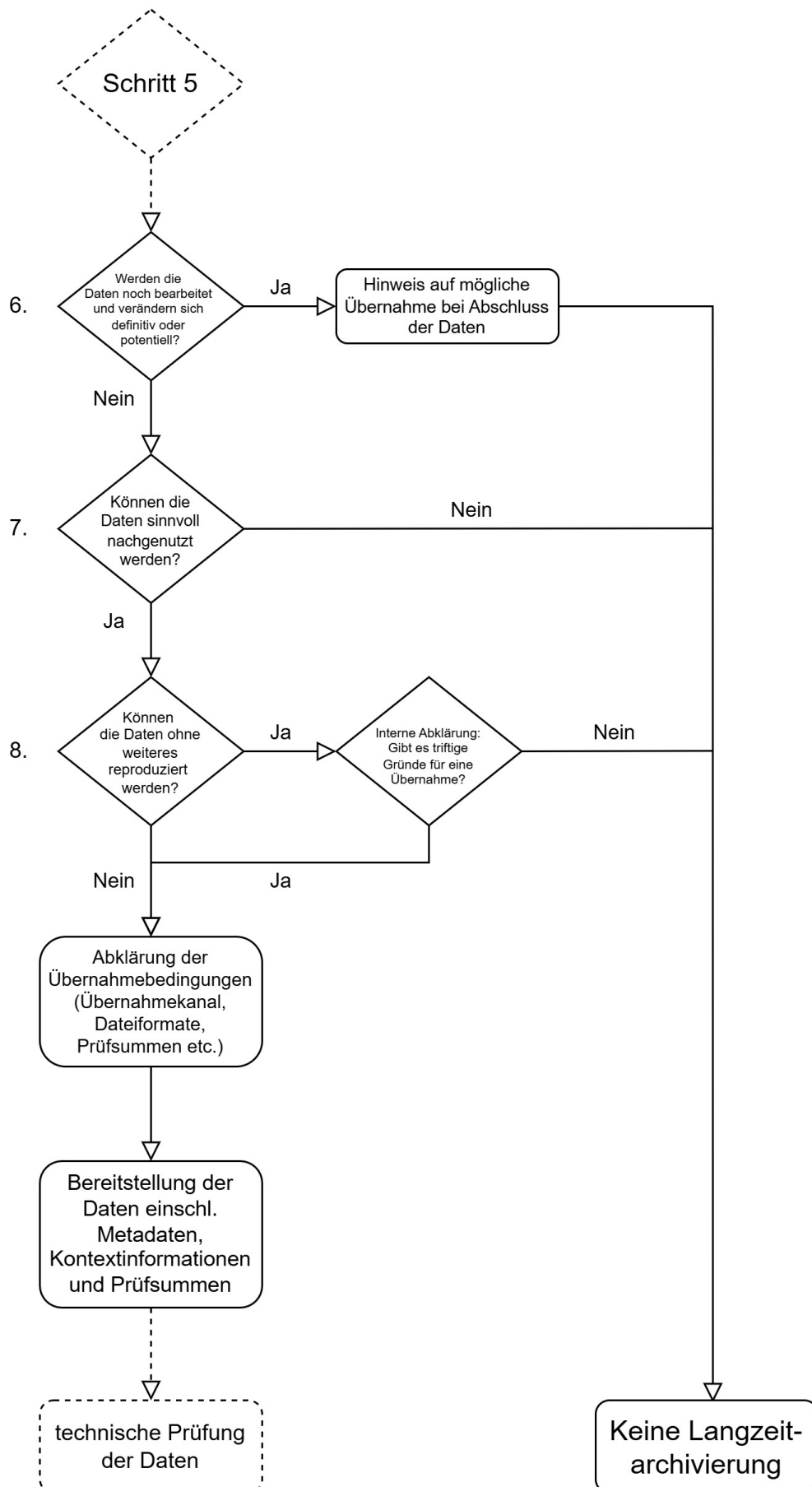


Abbildung 3: Technische Prüfung Schritte 9-11

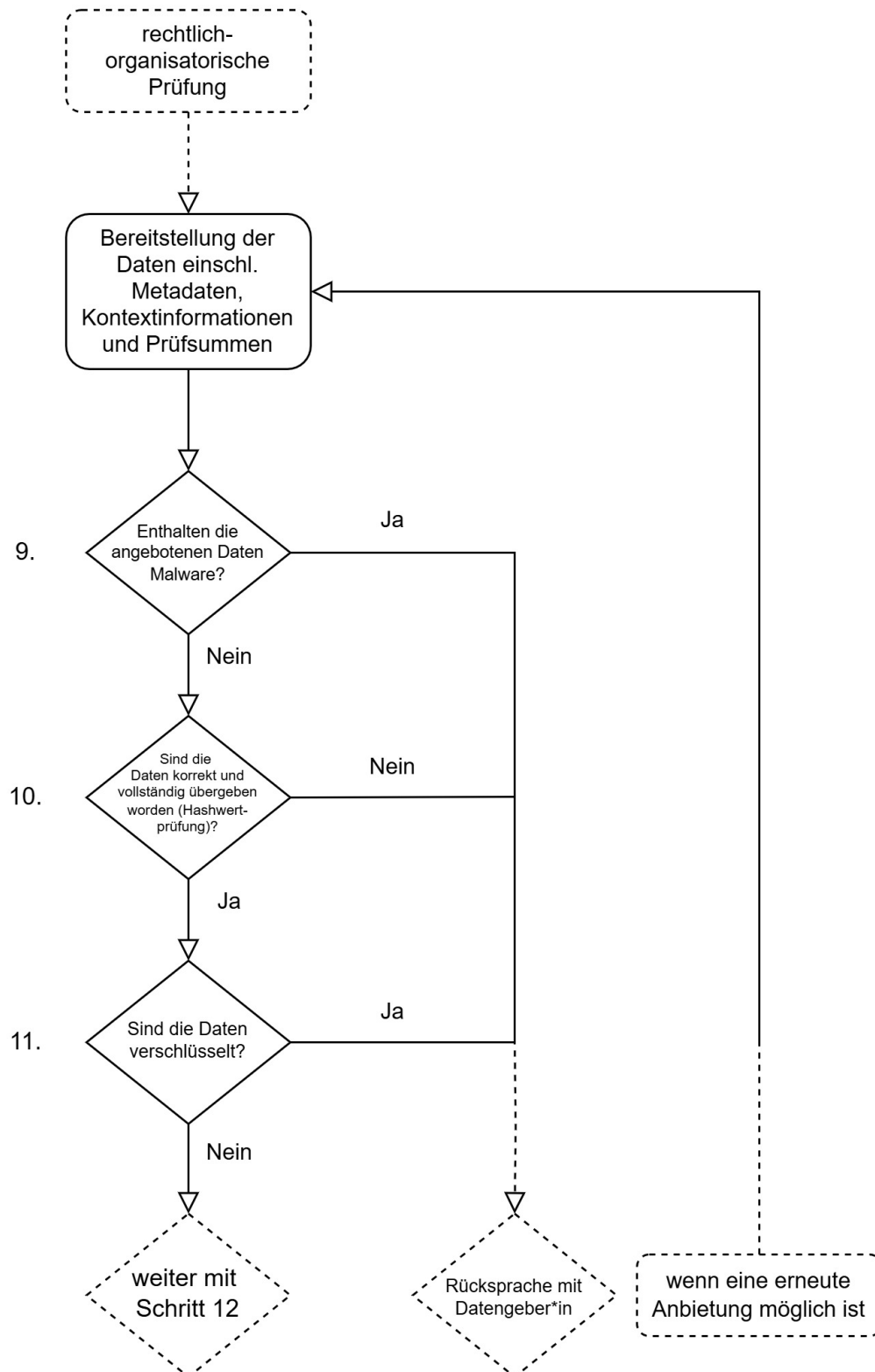


Abbildung 4: Technische Prüfung Schritt 12 bis zum Abschluss

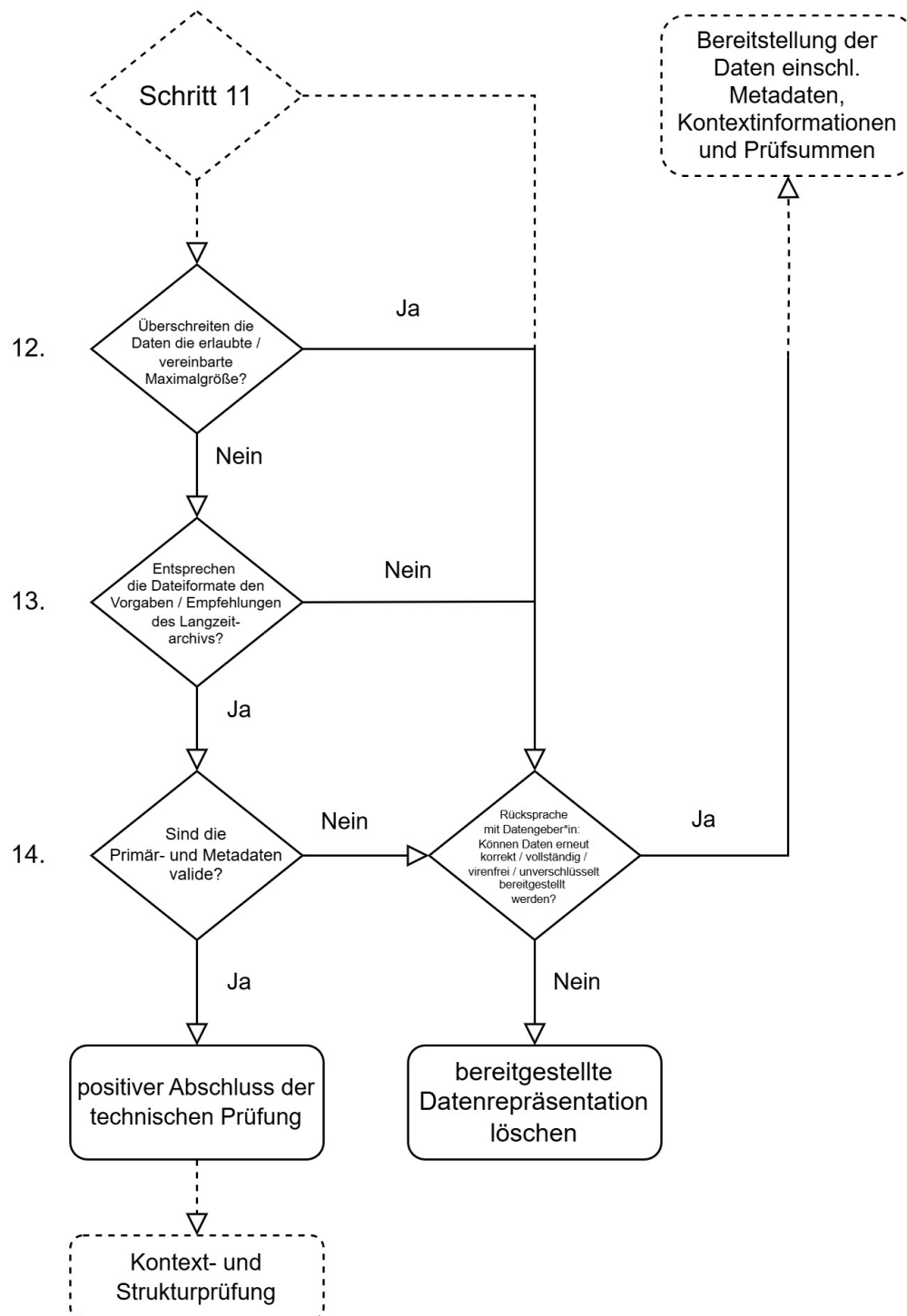


Abbildung 5: Ablauf der Kontext- und Strukturprüfung

