

5.2 Langzeitarchivierung

Daten managen – 5 Datenaufbewahrung

Unsere Versorgungsforscherin beschäftigt sich damit, wie unterschiedliche Behandlungsstrategien zur Verbesserung der Patientenversorgung beitragen können. Während dieses Projekts stößt sie auf alte Datensätze, die nur dank sorgfältiger Langzeitarchivierung verfügbar sind. Diese Daten ermöglichen ihr, Veränderungen in Behandlungsstrategien über viele Jahre hinweg zu erkennen und daraus neue Empfehlungen für die Zukunft abzuleiten. Sie erkennt, wie wichtig Langzeitarchivierung ist und beschließt auch ihre selbst erhobenen Daten so abzulegen, dass auch in der Zukunft Forscher:innen mit ihren Daten und Ergebnissen weiterarbeiten können.

Warum Langzeitarchivierung?

Die [Leitlinie 17](#) zur Sicherung guter wissenschaftlicher Praxis zum Thema Archivierung (DFG-Kodex) formuliert folgendes: Wissenschaftler:innen sollen öffentlich zugänglich gemachte Forschungsdaten und Forschungsergebnisse, inklusive die ihnen zugrunde liegenden Materialien und gegebenenfalls die eingesetzte Forschungssoftware sichern. Die Sicherung erfolgt gemessen an den Standards des betroffenen Fachgebiets in adäquater Weise und sollen für einen angemessenen Zeitraum aufbewahrt werden. Wenn nachvollziehbare Gründe dafür existieren, bestimmte Daten nicht aufzubewahren, sollen die Wissenschaftler:innen dies darstellen. Die Hochschulen oder Forschungseinrichtungen müssen indessen stellen sicher, dass die erforderliche Infrastruktur zur Archivierung vorhanden ist.

Wenn wissenschaftliche Erkenntnisse öffentlich zugänglich gemacht werden, sollten die zugrunde liegenden Forschungsdaten (in der Regel Rohdaten, siehe für mehr Infos **1 Datenstrukturierung**) in der Regel für einen Zeitraum von zehn Jahren (abhängig vom Fachbereich) zugänglich und nachvollziehbar gespeichert werden. Die Speicherung kann in der Einrichtung, wo sie entstanden sind, erfolgen oder, noch besser, in standortübergreifenden Repositorien aufbewahrt werden. Die Aufbewahrungsfrist beginnt mit dem Datum ab dem der öffentliche Zugang möglich ist.

Dateiformate für die Langzeitarchivierung

Dateiformate, die für die Langzeitarchivierung genutzt werden, müssen laut dem IT-Grundschutzkatalog des Bundesamts für Sicherheit in der Informationstechnik (BSI) unter M 4.170 „Auswahl geeigneter Datenformate für die Archivierung von Dokumenten“ folgende Kriterien erfüllen:

- das Datenformat sollte möglichst langfristige Relevanz haben,
- die Dokumentstruktur sollte eindeutig interpretiert werden können,
- der Dokumentinhalt sollte elektronisch weiterverarbeitet werden können, Beachtung gesetzlicher Vorschriften,
- die Grammatik und Semantik des Datenformates muss ausführlich dokumentiert sein, so dass eine spätere Migration problemlos möglich ist,
- Merkmale des Originaldokuments (elektronisch oder in Papierform) sollen später eindeutig nachweisbar sein, auch wenn das Originaldokument nicht mehr vorhanden ist

Für die Langzeitarchivierung geeignet sind die Auszeichnungssprachen SGML (Standard Generalized Markup Language) und XML (Extensible Markup Language), die allerdings in Kombination mit einer Semantikspezifikation (DTD,

Gefördert durch:



Bundesministerium
für Forschung, Technologie
und Raumfahrt



Finanziert von der
Europäischen Union
NextGenerationEU

Document Type Definitions) archiviert werden müssen. PDF-Dokumente (Portable Document Format) können insbesondere in der für die Archivierung optimierte Version PDF/A lange gespeichert werden. TIFF (Tagged Image File Format) ist ein Bilddateiformat, das für gerasterte Bilder verwendet werden kann und in komprimierter Form für die Langzeitarchivierung geeignet ist. Auch JPEG (Joint Photographic Experts Group) können für die Archivierung von Bildern genutzt werden.

Archivierungsmedien

Verschiedene Speichermedien besitzen eine unterschiedliche Lebensdauer. Gerade für die Langzeitarchivierung sollte also darauf geachtet werden ein ausreichend langlebiges Medium gewählt wird.

- Optische Speichermedien (z.B. CD-ROMs, DVDs, Blu-ray-Disks): Tests zeigen bei wiederbeschreibbaren Medien wie CD und DVD eine Haltbarkeit von weniger als 30 Jahren. Bei Blu-ray sind es etwa 50 Jahre. Wichtig ist eine richtige Lagerung: Bei Temperaturen von weniger als 25 Grad Celsius und unter 80% Luftfeuchtigkeit kann sich die Haltbarkeit verlängern.
- Festplatten (SATA, iSCSI): Festplatten, die nur als Backup Medium dienen können etwa 10 Jahre halten. Bei Platten, die fest in einem Rechner verbaut sind und produktiv bespielt werden, beträgt die Haltbarkeit nur 5 Jahre.
- Flash-Speicher (USB-Sticks, Solid State Drives / SSD): Die Haltbarkeit von Flashspeichern richtet sich stark nach der Anzahl der Lese- und Schreibzyklen. Namenhafte Hersteller versprechen eine Dauer von mindestens 5 Jahren. Bei Preisgünstigen Herstellern können es auch nur 3 Jahre sein. Als Archivmedium (mit wenigen Zugriffen) genutzt, kann ein Flash-Speicher auch 10 Jahre lang halten.

Eigenschaften die im Auswahlprozess eines geeigneten Archivs eine Rolle spielen und abgewogen werden sollten sind:

1. Technische Anforderungen,
2. Siegel für vertrauenswürdige Langzeitarchive (beispielsweise CoreTrustSeal, nestor-Siegel, DIN 31644),
3. Kosten der Nutzung
4. Zugänglichmachung der Daten
5. Langlebigkeit des Dienstleisters

Aufbewahrungsfristen für Arztpraxen

Für den Bereich der ärztlichen Dokumentation gibt grundsätzlich eine 10-jährige Aufbewahrungspflicht nach Abschluss der Behandlung (§ 630f Abs. 3 BGB). Aufgrund von Schadenersatzansprüchen und zivilrechtlichen Verjährungsfristen kann eine längere Aufbewahrung von bis zu 30 Jahren von Aufzeichnungen in manchen Fällen sinnvoll erscheinen.

Eine Übersicht über die Aufbewahrungsfristen für verschiedene Dokumente findet sich im [folgenden](#) Dokument von Bundesärztekammer & Kassenärztliche Bundesvereinigung (2021) aus.

Quellenverzeichnis

Blümm, M., Fritsch, K., Bock, S., Hackenbuchner, J., Arning, U., & Förstner, K. U. (2024). 08_LE_Daten-Speichern. FDM@Studium.nrw Blended-Learning-Basiskurs „Forschungsdatenmanagement“ (Version 1.0).

https://landesinitiativefdmnrw.github.io/FDMatStudium/thk/texte/08_LE_Daten-Speichern.html

Bundesamt für Sicherheit in der Informationstechnik (2016). M 4.170 Auswahl geeigneter Datenformate für die Archivierung von Dokumenten. IT-Grundschutz-Kataloge (S. 3490-3495). Abgerufen am 30.09.2025, von https://download.gsb.bund.de/BSI/ITGSK/IT-Grundschutz-Kataloge_2016_EL15_DE.pdf

Bundesärztekammer & Kassenärztliche Bundesvereinigung (2021). Hinweise und Empfehlungen zur ärztlichen Schweigepflicht, Datenschutz und Datenverarbeitung in der Arztpraxis. Deutsches Ärzteblatt. <https://doi.org/10.3238/arztebl.2021.ds02>. Abgerufen am 30.09.2025, von https://www.kbv.de/documents/infothek/rechtsquellen/weitere-vertraege/praxen/datenschutz-schweigepflicht/Empfehlungen_aerztliche_Schweigepflicht_Datenschutz.pdf

Deutsche Forschungsgemeinschaft (DFG)(2019). Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Abgerufen am 30.09.2025, von <https://wissenschaftliche-integritaet.de/kodex/>

Lowe, D.B., White, D., & Novak, E.R. (2022). Module 7: Repositories, Archiving & Preservation. New England Collaborative Data Management Curriculum. Abgerufen am 30.09.2025, von <https://oercommons.org/courses/new-england-collaborative-data-management-curriculum>

Mertzen, D., Neuroth, H., Schneemann, C., Woywod, K., Haase, C., Jacob, B., Kroehling, M., Mittelbach, J., Straka, J., Szczukowski, A., & Weise, K. (2023). Zertifikatskurs "Forschungsdatenmanagement für Studierende": Spring School 2023 der Landesinitiative für Forschungsdatenmanagement in Brandenburg. Zenodo. <https://doi.org/10.5281/zenodo.7936966>



Das Material ist im Rahmen des Projekts „DIM.RUHR: Datenkompetenzzentrum für die interprofessionelle Nutzung von Gesundheitsdaten in der Metropole Ruhr“ (Förderkennzeichen: 16DKZ2008[A-F]) entstanden. Dementsprechend spiegelt das Material in Version 1.0 den Stand von 2025 wider.

„Daten managen - 5.2 Datenaufbewahrung – Langzeitarchivierung“ von Anne Mainz und Sven Meister.

Dieses Werk und dessen Inhalt sind – sofern nicht anders angegeben – lizenziert unter CC BY 4.0. Ausgenommen aus der Lizenz sind die verwendeten Logos.

