



Handreichung: Datenbewertung und -löschung

Aus ökonomischen und technischen Gründen können und müssen nicht alle Daten, die im Laufe eines Forschungsprojektes erhoben wurden, dauerhaft aufbewahrt werden. Es ist daher notwendig, mit Abschluss des Projektes eine Datenbewertung vorzunehmen, die als Entscheidungsgrundlage dient, welche Daten archiviert werden sollen und welche Daten gelöscht werden können.

Kriterien für die Datenbewertung:

Bedarf Es gibt potenzielle Interessenten für die Daten.	Verifizierbarkeit Andere können mithilfe der Daten die Forschungsergebnisse verifizieren.	Dokumentation Die Forschungsdaten sind ausreichend mit Metadaten versehen.	Qualität Die Daten sind vollständig, umfangreich, genau, gültig, zuverlässig und repräsentativ.
Einzigartigkeit Es gibt nur eine einzige und vollständige Kopie der Daten.	Technische Erhaltung Die Daten liegen in einem für Langzeitarchivierung geeignetem Format vor.	Kosten Die Kosten für die Vorbereitung und Aufbewahrung wurden abgewogen.	Rechte Das Urheberrecht und das Datenschutzrecht wurden berücksichtigt.

Leitfragen zur Datenaufbewahrung:

Die Entscheidung, welche Daten beibehalten werden sollten, hängt davon ab, wie wertvoll und geeignet die Daten für eine potenzielle Nachnutzung sind.

Mögliche Leitfragen, die Sie sich bei der Datenevaluierung stellen können, sind:

1. **Welche Daten müssen aufbewahrt werden?** Orientieren Sie sich hier an den Vorgaben der Forschungsdaten-Policy, der Journal-Policy (im Falle einer Veröffentlichung der Daten bzw. zu den Daten) und den Förderrichtlinien. Berücksichtigen Sie außerdem, dass personenbezogene Daten (Datennutzung erfordert Einverständniserklärung) ggf. gelöscht werden müssen. Vertragsbindende und patentbezogene Aspekte (Daten haben kommerziellen Wert oder sollen als Patent angemeldet werden), sprechen hingegen für eine langfristige Aufbewahrung.

2. **Welche Zwecke können die Daten über den eigentlichen Forschungskontext hinaus erfüllen? Wie relevant sind die Forschungsdaten für eine eventuelle Wiederverwendung?** Evtl. sind die Daten relevant für die Verifizierung der Forschungsergebnisse. Mithilfe der Daten kann anderen ermöglicht werden, dem Prozess zu folgen, der zu den veröffentlichten Ergebnissen geführt hat und diese zu reproduzieren. Außerdem können die Daten die Möglichkeit für weiterführende Analysen erhöhen.
3. **Welche Daten sollen behalten werden?** z.B. aufgrund der Datenqualität. Qualitätskriterien sind: Vollständigkeit, Stichprobengröße, Genauigkeit, Gültigkeit, Zuverlässigkeit und Repräsentativität. Außerdem sollte das Nachnutzungspotenzial (Wie wahrscheinlich ist eine Nachfrage?) und die Einzigartigkeit der Daten berücksichtigt werden (Gibt es nur eine vollständige Kopie der Daten?).
4. **Welche Daten und Informationen werden für eine Nachnutzung gebraucht?** z.B. für weitere Publikationen (referenzierte Daten mit zusätzlicher Dokumentation) oder zu Lehrzwecken (Proben von Originaldaten und zusammengestellten Daten inklusive Analyseschritten.)
5. **Wurden die Kosten abgewogen?** Die Aufbewahrung von Daten über lange Zeiträume umfasst mehr als nur die Speicherung. Die Daten müssen vorbereitet werden, zugänglich sein, Sicherungskopien erstellt und der Zugriff gewährleistet werden. All dies erhöht die Kosten der Datenerhaltung und muss bei der Finanzierung berücksichtigt werden.

Datenlöschung:

Bestimmte Daten müssen spätestens nach Abschluss des Forschungsprojektes / der Forschungsphase gelöscht werden. So muss die Löschung personenbezogener Daten dann erfolgen, wenn eine Person eine zugrunde liegende Einwilligungserklärung widerruft. Auch wenn der Zweck der Datenverarbeitung erfüllt ist, sind die Informationen zu löschen. Verantwortlich für die Löschung ist der Datenproduzent bzw. die Projektleitung.

Wichtig: Sind Sie nach DSGVO angehalten, personenbezogene Daten zu löschen, müssen bei der Löschung Datenschutzstandards eingehalten werden. Daten müssen so vernichtet werden, dass sie nicht oder nicht ohne Weiteres wiederhergestellt werden können (z.B. Schreddern auch bei physischen Datenträgern wie Festplatten).

Grundsätze der Datensparsamkeit:

Die Verwaltung großer Datenmengen erfordert Aufwand und mit Zunahme der Datenmenge kann auch die Unübersichtlichkeit der zu verwaltenden Daten steigen. Im Umgang mit datenschutzrechtlich relevanten Daten wächst mit steigender Datenmenge auch die Gefahr, schätzenswerte Daten zu übersehen, der mit der Reduzierung der Datenmenge begegnet werden kann (vgl. Heber et al., 3).

Checkliste zur Überprüfung der Daten auf Sparpotentiale:

Während der Datenerhebung:

- ✘ Erfolgt die Speicherung der Daten in einer geordneten und nachvollziehbaren Verzeichnisstruktur?
- ✘ Ist eine Beschreibung der Daten mit Metadaten vorhanden?
- ✘ Sollen nur zu Test- oder Übungszwecken erzeugte Daten in einem separaten Ordner gespeichert werden?
- ✘ Werden Daten beispielsweise aus Methodeneinstellungen oder der Kalibrierung von Instrumenten, sofern sie für die spätere Beurteilung der Datenqualität oder für die weitere Arbeit nicht mehr benötigt werden, gelöscht?

Monatlich:

- ✘ Wurden ältere Daten durch neue Daten oder Versionen redundant oder obsolet?
- ✘ Gibt es Zwischenergebnisse, die sich leicht aus anderen Daten reproduzieren lassen?
- ✘ Sind die Daten für eine Publikation geeignet (Datenqualität, Reproduzierbarkeit, Nachvollziehbarkeit)?
- ✘ Können die Daten für andere Forschende/Studierende von Nutzen sein?
- ✘ Sind die Daten nach Abschluss des Projekts noch in irgendeiner Weise von Nutzen oder können es zukünftig sein?

Halbjährlich:

- ✘ Wie viele Kopien bestehen von erhobenen Daten (auf derselben oder anderen Ressourcen)?
- ✘ Befinden sich erhobene Daten in akutem Gebrauch (Möglichkeit zur Archivierung)?
- ✘ Haben sich die Rechte an bestimmten Daten zwischenzeitlich geändert?
- ✘ Wurden die Daten zwischenzeitlich in einem Repositorium veröffentlicht und darüber zugänglich gemacht?

Quellennachweise:

DCC (2014): Five steps to decide what data to keep. A checklist for appraising research data. V.1.
Edinburgh: Digital Curation Centre

DCC (2010): How to Appraise & Select Research Data for Curation:

<https://www.dcc.ac.uk/sites/default/files/documents/How%20to%20Appraise%20and%20Select%20Research%20Data.pdf> [abgerufen 25.07.2024]

Heber, M., Jakob, M., Landwehr, M., Leendertse, J., Müller, M., Schneider, G., von Suchodoletz, D., & Ulrich, R. (2024). Leitfaden zur digitalen Datensparsamkeit. O-Bib. Das Offene Bibliotheksjournal
Herausgeber VDB, 11(2), 1-8. <https://doi.org/10.5282/o-bib/6036>

Forschungsdaten.info: <https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/datvalidierung/> [abgerufen 25.07.2024]

VFR Verlag für Rechtsjournalismus GmbH (2024): <https://www.datenschutz.org/impresum/>
[abgerufen 15.07.2024]