Checkliste zur Datendokumentation[[1]](#footnote-1)

Prof. Dr. Mirjam Blümm, Katharina Fritsch, Sina Bock, Prof. Dr. Ursula Arning, Prof. Dr. Konrad Förstner

Entstanden im Rahmen des Kooperationsvorhabens FDM@Studium.nrw, gefördert vom Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen in der Zeit vom 01.10.2022 bis zum 31.08.2024.

Lizenz: CC BY 4.0 // OER

Zitiervorschlag: Blümm, M., Fritsch, K., Bock, S., Arning, U., & Förstner, K. U. (2024). Checkliste zur Datendokumentation. FDM@Studium.nrw Blended-Learning-Basiskurs „Forschungsdatenmanagement“ der TH Köln (Version 1.0).

|  |  |
| --- | --- |
| Dozent\*in: |  |
| Kurs (Semester): |  |
| Datum: |  |
| Name: |  |
| Titel: |  |
| Studienleistung / Prüfungsleistung: |  |

# 1. Planen

|  |  |  |
| --- | --- | --- |
| Themenfindung und Forschungsfrage | Wie lautet die Forschungsfrage / Aufgabe? |  |
| In welchem Zeitraum beschäftigen Sie sich damit? |  |
| 3 Schlagwörter zur Frage / zum Thema |  |
| Verantwortlichkeiten | Verantwortliche Person(en)  1. für die Datendokumentation:  2. für die Datenerhebung:  3. für die Datenspeicherung während des Bearbeitungszeitraums:  4. für die Datenspeicherung nach Ende des Bearbeitungszeitraums:  5. für den Schutz der Daten: | 1.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  2.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  3.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  4.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  5.\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| Literaturrecherche | Wie soll die verwendete Literatur organisiert werden? |  |
| Wird ein Literaturverwaltungsprogramm verwendet (z. B. Citavi, Zotero)? |  |
| Zeit-/Arbeitsplan | Wurde ein grober Zeit- bzw. Arbeitsplan für die Bearbeitung der Forschungsfrage / Aufgabe erstellt? | ☐ ja / ☐ nein |

# 2. Inhaltliche Einordnung

|  |  |  |
| --- | --- | --- |
| Herkunft der Daten | Werden die Daten nachgenutzt (Fremddaten) oder selbst erzeugt (Eigendaten)? | ☐ Eigendaten / ☐ Fremddaten / ☐ Mischform |
| Wann erfolgt(e) die Erhebung bzw. Erstellung der Daten? |  |
| Bei Fremddaten | Wer hat die Daten erzeugt? |  |
| Wo wurden die Daten erzeugt? |  |
| Wo wurden sie publiziert? |  |
| Wurde sichergestellt, dass die Daten verwendet werden dürfen und mit einer entsprechenden Lizenz versehen sind? | ☐ ja, CC\_\_\_\_\_\_  ☐ nein |
| Sind die Daten permanent verfügbar und zitierbar? Das heißt, verfügen sie über einen Persistent Identifier (PI), wie z.B. eine DOI? | ☐ ja / ☐ nein |
| Datensatz | Um welche Arten von Daten handelt es sich? |  |
| Sind ethische und rechtliche Vorgaben zu berücksichtigen? | ☐ ja (s. Punkt 5) / ☐ nein |
| Sind die Daten reproduzierbar, d. h. ließen sie sich, wenn sie verloren gingen, erneut erstellen oder erheben? | ☐ ja / ☐ nein  Begründung: |
| Welche Instrumente, Software, Technologien oder Verfahren werden für die Erzeugung, Erfassung, Bearbeitung, Analyse, Visualisierung etc. der Daten genutzt? |  |
| Sind die Daten auf Fehlerhaftigkeit und / oder Unvollständigkeit geprüft worden? | Fehlerprüfung:  ☐ ja / ☐ nein  Vollständigkeitsprüfung:  ☐ ja / ☐ nein |
| Wurden die Daten bereinigt und in eine geeignete Form gebracht (z. B. Tabellen in statistische Analyse-Tools, Textanalyse Software usw.)? | ☐ ja / ☐ nein |
| Methodik | Werden bestimmten Regeln / Richtlinien / Methoden im Umgang mit den Daten gefolgt (bspw. qualitative / quantitative Methoden)? |  |
| Warum wurden diese Regeln / Richtlinien / Methoden gewählt? |  |
| Auf welchen Theorien / Prinzipien / Vorgaben / Vorlagen beruht die angewandte Methode? Wird diese ausreichend dokumentiert? |  |
| Metadaten | Können die Daten auch ohne eine schriftliche Arbeit verstanden werden oder müssen sie aufgearbeitet werden? |  |
| Welche Informationen sind für Außenstehende (z. B. Lehrende, Studierende) notwendig, um die Daten sowie die auf ihrer Basis gewonnenen Forschungsergebnisse nachvollziehen und nachnutzen zu können? |  |

# 3. Datenspeicherung

|  |  |  |
| --- | --- | --- |
| Datengröße | Was ist die (zu erwartende) Datengröße? |  |
| Datenformate | In welchen Formaten liegen die Daten vor? |  |
| Sind die Formate offen und standardisiert? | ☐ ja / ☐ nein |
| Sind die Formate weit verbreitet? | ☐ ja / ☐ nein |
| Speicherort | Wo werden die entstehenden Daten in der Bearbeitungszeit gespeichert (Zwischenspeicher)? | ☐ lokal, eigener Computer  ☐ extern, Festplatte / USB-Stick  ☐ extern, Cloud-Speicher  ☐ anderes: |
| Wo werden die Daten nach dem Ende des Vorhabens gespeichert (finaler Speicherort)? | ☐ lokal, eigener Computer  ☐ extern, Festplatte / USB-Stick  ☐ extern, Cloud-Speicher  ☐ anderes: |
| Welche Ressourcen benötige ich zur Datenspeicherung?  *z. B. Software oder Speicherressourcen, wie sciebo als Cloudspeicher oder USB-Stick/externe Festplatte als externes Speichermedium* |  |
| Backup | Wie werden/wurden Backups erstellt? | ☐ automatisch / ☐ manuell |
| Wie oft werden/wurden Backups erstellt? | ☐ täglich / ☐ wöchentlich / andere: |
| Mein Backupschema folgt dem 3-2-1 Prinzip. Zwei der Kopien sind auf unterschiedlichen Arten von Speichermedien gespeichert (z. B. Cloud und Lokal). Und eine der Kopien ist an einem anderen Ort gelagert (z. B. USB-Stick). | ☐ ja / ☐ nein |

# 4. Datenorganisation

|  |  |  |
| --- | --- | --- |
| Dateibenennung | Wie wurden die Daten benannt? Gibt es eine einheitliches Schema? Wenn ja, welches?  *Bspw.:*   * *Die Dateibenennungen beinhalten keine Leerzeichen, Umlaute oder Sonderzeichen.* * *Die Dateibenennungen beinhalten eine Datumsangabe (JJJJ-MM-TT / JJ-MM-TT oder JJJJMMTT / JJMMTT).* | ☐ ja,\_\_\_\_\_\_  ☐ nein |
| Ordnerstruktur | Wurde eine übersichtliche, einheitliche Ordnerstruktur angelegt? Wenn ja, welche?  *Bspw.:*   * *Die Ordner sind nach Thema kategorisiert.* * *Die Ordnernamen verweisen auf die Ordnerinhalte.* * *Die Ordnerstruktur beinhaltet übergeordnete (allgemeinere) und untergeordnete (konkretere) Elemente.* | ☐ ja,\_\_\_\_\_\_  ☐ nein |
|  | Können die Dateibenennungen und Ordnerstrukturen auch von Dritten nachvollzogen werden? | ☐ ja / ☐ nein  Falls nicht, sind Erläuterungen zum Benennungsschema in einer separaten Datei abgespeichert?  ☐ ja / ☐ nein |
| Versionierung | Werden verschiedene Versionen der Daten erzeugt (z. B. durch verschiedene Weiterbearbeitungsprozesse)? | ☐ ja / ☐ nein  Wenn ja, Versionskontrolle:  ☐ automatisch / ☐ manuell |
| Weitergabe / Nachnutzung  Wenn ja,… | Ist es geplant, die Daten nach der Aufgabe zu teilen? Z. B. mit Lehrenden / Studierenden. | ☐ ja / ☐ nein |
| Sind bestimmte Software, Verfahren oder Technologien notwendig, um die Daten nachzunutzen? | ☐ ja,\_\_\_\_\_\_  ☐ nein |
| Können die Daten auch ohne eine schriftliche Arbeit verstanden werden oder müssen sie aufgearbeitet werden? | ☐ ja / ☐ nein |
| Welche Informationen (wie Metadaten) sind für Außenstehende (z. B. Lehrende, Studierende) notwendig, um die Daten sowie die auf ihrer Basis gewonnenen Forschungsergebnisse nachvollziehen und nachnutzen zu können? |  |

# 5. Rechtliche und ethische Fragen

|  |  |  |
| --- | --- | --- |
| Personenbezogene Daten | Enthalten die Daten personenbezogene Informationen? | ☐ ja / ☐ nein |
| Sensible Daten | Enthalten die Daten besondere Kategorien personenbezogener Daten nach Artikel 9 der DSGVO *(“Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit, Gesundheit oder Sexualleben”)*? | ☐ ja / ☐ nein |
| Urheber- oder verwandte Schutzrechte | Werden Daten genutzt und / oder erstellt, die durch Urheber- oder verwandte Schutzrechte geschützt sind? | ☐ ja / ☐ nein |
| Maßnahmen zum Schutz von personenbezogenen / sensiblen Daten | Wie werden diese Daten geschützt? |  |
| Könnte Ihre Testpersonen durch den Kontext identifiziert werden? | ☐ ja / ☐ nein |
| Werden die Daten anonymisiert oder pseudonymisiert? | ☐ anonymisiert /  ☐ pseudonymisiert |
| Haben Sie eine "informierte Einwilligung" der Betroffenen eingeholt? Wo und wie sind die "informierten Einwilligungen" abgelegt? Wenn keine "informierte Einwilligung" eingeholt wird, begründen Sie dies bitte. | ☐ ja, abgelegt\_\_\_\_\_\_  ☐ nein, \_\_\_\_\_ |
| Bis wann werden die (un-anonymisierten bzw. un-pseudonymisierten) Originaldaten spätestens sicher vernichtet? |  |

1. Abgewandelt von: Institut für Bibliotheks- und Informationswissenschaft. (2021, 8. Dezember). Datenmanagementplan für die Abschlussarbeit – Template. Humboldt-Universtität zu Berlin. Zugriff am 29.08.2024, von https://www.ibi.hu-berlin.de/de/studium/rundumdasstudium/fdm-fuer-studierende/dmp\_template\_finale\_version\_dez\_21.pdf. [↑](#footnote-ref-1)