



Handreichung: Gute Datenorganisation

Ein umfassendes Forschungsdatenmanagement erfordert eine sorgfältige und gut strukturierte Datenorganisation. Dies mag zu Beginn aufwendig erscheinen. Dieser Mehraufwand erleichtert die Arbeit erheblich, wenn für alle Beteiligten nachvollziehbar ist, wo sie welche Daten finden und wo und nach welchem System sie diese ablegen (können). Vor allem in Forschungsk Kooperationen mit mehreren Partnern ist dies unerlässlich.

Eine gute Datenorganisation spart viel Zeit und Kosten, da hierdurch ein Datenverlust, aufwendiges Suchen und Doppelarbeit vermieden werden können und auch Dritte die Daten nachvollziehen und diese (nach-)nutzen können. Daher ist es ratsam, zu Beginn des Projektes ein Konzept zu erstellen und unter allen Projektbeteiligten verbindlich abzustimmen. Dieses Konzept sollte mind. die **1) Ordnerstruktur**, **2) Dateibenennung** und **3) Dateiversionierung** regeln.

Für die Konzepterstellung gibt es einige Empfehlungen. Wichtig ist aber, dass individuelle Lösungen gefunden werden, mit denen alle Projektmitarbeitenden gut arbeiten können.

1) Ordnerstruktur

- Die Ordnerstruktur sollte logisch anhand von Datentypen oder Themenkomplexen aufgebaut sein.
- Es empfiehlt sich, verständliche und kurze Ordnernamen zu verwenden, die den Inhalt des Ordners beschreiben.
- Für eine bessere und standardisierte Ordnung sollten die Ordner durchnummeriert werden.
- Empfehlenswert ist das Sieben-Ordner-Modell: Es sollten maximal sieben Hauptordner angelegt werden, in denen jeweils höchstens sieben Unterordner mit jeweils maximal sieben Unterordnern angelegt werden.
- Mehr als drei Ebenen sollten nicht verwendet werden, da dies die Übersichtlichkeit erschwert.

2) Dateibenennung

- Dateien sollten nach einheitlichem Schema benannt werden.
- Dateinamen sollten aussagekräftig und nicht zu lang sein.
- Es ist zu empfehlen, das Erstellungsdatum im Dateinamen zu integrieren.
- Der Dateiname darf keine Sonder- und Leerzeichen enthalten.

- Beispiele für Konventionen:
„JJJJ-MM-TT_Projekt_Interview_Nr.mp4“
„JJJJ-MM-TT_Sediment_Probe_Instrument.csv.“

3) Dateiversionierung

- Die Dateiversion kann im Dateinamen oder der Datei selbst vermerkt werden.
- Dies kann über eine Versionsnummer und/oder das Datum erfolgen. Das Speicherdatum sollte immer im JJJJ-MM-TT-Format vermerkt werden.
Falls das Datum als Dateiversionierung verwendet wird, ist zu beachten, dass es vorkommen kann, dass neue Versionen einer Datei auch innerhalb eines Tages erstellt werden.
- Es ist zu empfehlen, eine Versionstabelle zu führen, in der Änderungen und Benennungen festgehalten werden, insb. wenn mehrere Mitarbeitende mit denselben Daten arbeiten.
- Es sollten zu festgelegten Zeitpunkten „Meilenstein-Versionen“ erstellt werden, die nicht geändert oder gelöscht werden. Es sollte geklärt und protokolliert sein, ob die entstandenen Zwischenversionen innerhalb der Projektlaufzeit oder in der Abschlussphase gelöscht werden können und wer dafür verantwortlich ist.

Literaturempfehlung zur Vertiefung:

Netscher, S. & Jensen, U. (2019): Forschungsdatenmanagement systematisch planen und umsetzen. In Netscher S., Jensen U., & Weller K. (Hrsg.): Forschungsdatenmanagement sozialwissenschaftlicher Umfragedaten: Grundlagen und praktische Lösungen für den Umgang mit quantitativen Forschungsdaten, S. 37–56.

Quellennachweise:

Netscher, S. (2019): Einführung in das Forschungsdatenmanagement. Die Generierung hochwertig (nach-)nutzbarer Forschungsdaten. Vortrag bei BW-CAR Kolleg. 26.06.2019.

KIM der Universität Konstanz (2024): <https://forschungsdaten.info/themen/organisieren-und-aufbereiten/datenorganisation/> (22.07.2024).