

Handout "Datenspeicherung und -archivierung"

Worum geht es?



Um Maßnahmen zur Datenspeicherung und zur Datenarchivierung. Ersteres meint Maßnahmen, die eher die kurzfristige Speicherung von Daten betreffen, d.h. während mit diesen noch aktiv und meist täglich gearbeitet wird und ein schneller Zugriff notwendig ist. Dies bezieht sowohl die softwareseitige Speicherung ein, aber beispielsweise auch Überlegungen über die Sicherheit der Hardware.

Letzteres bezieht sich auf Maßnahmen zur langfristigen Speicherung. Die Datenarchivierung beschäftigt sich im Vergleich zur Datenspeicherung mit der Lagerung und Verfügbarmachung von Forschungsdaten nach Beendigung der Forschungsarbeit und stellt dahingehend Anforderungen an Datenformate und eine ausführliche Dokumentation über Metadaten. Man unterscheidet zwischen der mittelfristigen Verfügbarmachung und der sogenannten Langzeitarchivierung.

Warum ist das wichtig?



Maßnahmen zur Datenspeicherung helfen beispielsweise, um dem Risiko eines Datenverlusts vorzubeugen. Es ist deshalb wichtig, (automatische) Backupstrategien zu kennen und diese auf die eigenen Projektdaten anzuwenden, um auch der Dokumentations- und Nachweispflicht im Sinne der guten wissenschaftlichen Praxis nachzukommen. Viele Experimente oder Daten können nur einmalig erhoben werden, weshalb durch einen endgültigen Verlust dieser Daten dann der positive Ausgang eines ganzen Projekts gefährdet sein kann.

Datenarchivierung ist notwendig, um langfristig die Verfügbarkeit von Forschungsdaten gemäß den FAIR-Prinzipien zu gewährleisten, um auf diese zugreifen und sie interpretieren, verstehen und nachnutzen zu können. Dazu gehört auch eine kritische Auseinandersetzung mit Dateiformaten und die Kenntnis über die Anforderungen einer Langzeitarchivierung.

Wie setze ich das um?



Maßnahmen zur Datenspeicherung:

- mögliche Speicherorte- und Medien:
 - ➔ institutioneller Speicherort: Hochschul-Clouddienst, Netzlaufwerk
 - ➔ externer Speicherort: Dropbox, Google Drive
 - ➔ mobiles Speichermedium: USB-Stick, externe Festplatte
 - ➔ lokaler Arbeitsplatzrechner oder eigener PC
- geklärter Datenzugang: bei Projekten mindesten zwei Personen mit Passwortkenntnis
- regelmäßige Backups gemäß 3-2-1 -Regel: **3** Kopien auf **2** Speichermedien und **1** davon dezentral

Maßnahmen zur Datenarchivierung:

- informieren über und nutzen von geeigneten Repositorien
- für Langzeitarchivierung frühzeitig geeignete Datenformate beim ausgewählten Archiv oder Repository erfragen
- den gesamten Erhebungsprozess dokumentieren (Software, Hardware, Migrationshistorie)

