

Generelle Tips zur Datenverwaltung (für die Diss und mehr)

Cheatsheet zum Workshop von Elisabeth W. Apicella und Maximilian Krug auf dem 5. DokGF-Arbeitstreffen am 12. Mai 2022

Backups!

- Regelmäßig und möglichst redundant, z. B. Arbeitsstand auf Laptop, Backups auf externem Heimserver und in Nextcloud
- Bei Cloud-Services unbedingt darauf achten, dass die Verarbeitung den Datenschutzanforderungen eurer Daten genügt!
- Wie oft ist oft genug? Das hängt davon ab, mit welchem Fortschrittsverlust man persönlich umgehen könnte. Am besten klappt ein automatisiertes Intervall plus die Möglichkeit zum händischen Start von zusätzlichen Backups bei Bedarf

Verzeichnisstrukturen

- Können sehr hilfreich sein – sofern es nicht zu tief runter geht
- Wie viel „zu tief“ ist, ist sehr persönlich, aber die Übersichtlichkeit hängt stark von der Ansicht im genutzten Navigationstool ab (also: sehe ich immer nur den Ordner, in dem ich bin, oder sehe ich den ganzen Pfad oder gar den kompletten Baum?)
- Je nach Backuplösung darauf achten, dass die Verzeichnisstruktur damit zusammen passt, dass also das Backup verschobene Dateien am Ende nicht zweimal oder gar nicht mehr enthält
- Grundsätzlich: In einem Ordner sollten immer nur entweder weitere Ordner oder Dateien enthalten sein
- Ausnahme: Elisabeth nutzt gerne Ordner mit dem Titel „Archiv“, um alte Versionsstände aufzuheben, ohne den eigentlichen Arbeitsordner zuzumüllen

Benennung von Dateien

- Sprechende Dateinamen sind für Dateien, an denen aktiv gearbeitet wird, ein absolutes Muss. Seid dabei so knapp wie möglich und so spezifisch wie nötig (und bedenkt alle (Nach-)Nutzer*innen mit!)
- Bewährt hat sich z. B. ein Schema wie
Datum(JJJJMMTT)_Titel-Spezifizierung_Version.Dateiendung
20220501_dokgf-workshop-datenverwaltung_v1.0.md
- Bei relativ statischen Dateien wie z. B. Literatur, Papers, reicht oft ein simpleres Schema aus, z. B. NameJahr, sofern zusätzliche Informationen dann anderswo abgelegt werden (siehe unten: Datenbanken)

- Außer, wenn es einen sehr guten Grund gibt¹, solltet ihr mehrfach benutzte generische Benennungen wie "01.mp4" unbedingt vermeiden, auch wenn scheinbar durch das übergeordnete Verzeichnis genauer spezifiziert wird: Sollte der Kontext in irgendeinem Anwendungsfall verloren gehen (z. B. bei Anzeige in Suchergebnissen, versehentliches Löschen oder Verschieben der Datei), so ist eine nachträgliche Zuordnung nicht immer möglich

Datenbanken

- Lagert Information an einen dafür vorgesehenen Ort aus²; z. B. Anmerkungen zu Video- oder Audiodateien, Versionsgeschichten (etwa Änderungen zwischen v1.0 und 1.1 eines Konferenz-Posters), to-dos zur gemeinsamen Bearbeitung etc.
- Nutzt Literaturverwaltungssoftware (siehe das Cheatsheet vom Workshop zu Jabref [[Link 1](#)])

Versionskontrolle

- Bei simultanem Arbeiten an ein und demselben Datenset über lokale Netzwerke oder Clouds: Nutzt nach Möglichkeit Tools, die genau für diesen Anwendungsfall gemacht wurden. Also nicht Dropbox, sondern ggf. git [[Link 2](#)]. Euer Rechenzentrum oder euer lokaler Hackspace geben euch sicher gern eine kurze Einführung. Weniger geeignet für große Dateien wie Videos.

Choose your heroes wisely

- Aus vielen Gründen (auf die wir vielleicht an anderer Stelle mal länger eingehen können) ist es gerade in wissenschaftlichen Projekten höchst ratsam, eure Daten und Informationen mit Freier Software anzulegen, zu verwalten und zu bearbeiten.³
- Generell immer, aber vor allem, wenn Freie Software keine Option ist: speichert so viel Information wie möglich (ggf. zusätzlich) als plaintext (v. a. eigene Schriftstücke).
Denn:
„Software dies, plaintext is forever“ [[Link 3](#)]
- Lernt Markdown [[Link 4](#)], es dauert +/- 5 Minuten und lohnt sich einfach.

¹Ein guter Grund ist z. B., dass ein Programm die Datei erzeugt hat und genau diesen Dateinamen zur Weiterverarbeitung benötigt

²Nennt es „Spreadsheet“, „Exceltabelle“ oder „Informationsdump“, wenn das Wort „Datenbank“ für euch negativ belegt ist – hauptsache, ihr nutzt es!

³Die wichtigsten Argumente könnt ihr bei Interesse auf der Website der Free Software Foundation nachlesen [[Link 5](#)].

Links

Link 1: https://dokgf.files.wordpress.com/2021/07/jabref-workshop_reader-1.pdf

Link 2: <https://de.wikipedia.org/wiki/Git>

Link 3: <https://twitter.com/inderkuerze/status/1433453792712867843>

Link 4: <https://commonmark.org/>

Link 5: <https://fsfe.org/freesoftware/freesoftware.de.htm>