# Best Practice Forschungsdaten

Dario Kampkaspar
Herzog August Bibliothek Wolfenbüttel

DARIAH-DE

# Wozu »Best Practice«?

- Mehrere Mitarbeiter mit verschiedenen Aufgaben
- Mehrere Szenarien der Nachnutzung
- Probleme der Langzeitarchivierung
- Vorgaben der Mittelgeber
- ?

Ein Forschungsprojekt hat viele Aspekte, die berücksichtigt werden müssen, (nicht nur) wenn die Ergebnisse später nützlich sein sollen!

# »Forschungsdatenlebenszyklus« (RDLC)

Verschiedene Modelle entwickelt; grundlegende Inhalte meist:

1. Grundlage sichten und Projekt planen
2. Daten sammeln, aufbereiten und analysieren
3. (Daten und/oder Ergebnisse) publizieren

DARIAH-DE AG RDLC:
http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2015-11.pdf

# Planen

- Forschungsdaten von Anfang an in die Projektplanung einbeziehen!
- Metadaten nicht vergessen!
- Von hinten denken: Was und für wen?
- Lizenzen festlegen!
- Abläufe festlegen!
- Standards beachten!
- Pläne diskutieren und überarbeiten!
- Dokumentieren!

# Planen Projektplanung

- Forschungs- und Metadaten sind zentrale Bestandteile eines Projektes, nicht notwendiges Übel oder Abfallprodukt!
- Zeit (Personal) für Pflege und Verwaltung der Daten und Metadaten einplanen
- Datenhaltung einplanen
- Standards prüfen und festlegen (cave: Arbeits- vs- Archivformat!?)
- Lizenzen für die veröffentlichten Daten prüfen und festlegen
- Dokumentieren!

# Planen Standards

- Welche Daten sollen erhoben und weitergegeben werden?
- Für wen sollen die Daten später bereitgestellt werden?

- Offene/freie und weit verbreitete Formate für Archivierung und (allgemeinen) Austausch
- Arbeitsformat (ggfs. Austauschformat innerhalb der Fachcommunity) nach Verbreitung, *aber möglichst umwandelbar*

- Dokumentieren!

# Planen Abläufe

- Abläufe (»Workflows«) sind nicht nur für Softwareentwickler oder BWLer interessant!
- Umfangreiche Projekte brauchen festgelegte Abläufe, an die sich alle halten, um Chaos zu vermeiden und Ziele zu erreichen
- Abläufe dienen der Nachvollziehbarkeit der Forschung und der Erhebung der Daten
- Datenaustausch festlegen
- Dokumentieren!

# Arbeiten

- Fortwährende Kommunikation im Projekt sicherstellen
- Anfallende Daten regelmäßig sichern, vorzugsweise mit Versionierung
- Pläne diskutieren und überarbeiten
- Metadaten aktuell halten
- Dokumentieren!

# Arbeiten Kommunikation

- über die Arbeitsrealität (Abläufe und Standards)
- Abgleich Zeitplan und tatsächlicher Forstschritt
- Pläne, Ziele, Standards diskutieren und anhand des Projektverlaufs und der Ressourcen aktualisieren
- Dokumentieren!

# Arbeiten Datenhaltung und -austausch

- zunächst alle anfallenden Daten sichern (Speicherplatz ist billig!)
- relevante Daten sortieren (Art, Zugehörigkeit)
- relevante Daten versionieren

- nachvollziehbaren Datenaustausch gewährleisten
- verlustbehaftetes Überschreiben vermeiden
- Verwantwortlichkeiten klären

- Dokumentieren!

# Arbeiten Metadaten

- Metadaten für das gesamte Projekt wie auch für einzelne Teile erheben und dokumentieren
- aktuell halten
- Vergleichbarkeit mit anderen Metadaten sicherstellen (Normdaten!)

- Dokumentieren!

# Publizieren

- Vor endgültiger Publikation aktualisieren und prüfen
- Weite Verfügbarkeit sicherstellen (offene Formate, verzeichnen in Datenbanken, Katalogisierung)
- Feste Adressierung sicherstellen
- Spätere Änderungen dokumentieren und Versionen adressierbar machen

- Dokumentieren!

# Beispiel

- Kommunikation: regelmäßige Treffen, (DARIAH-DE) Wiki
- Standards: TEI-XML, MEI, SVG, TIFF
- Metadaten: METS, MODS
- Versionierung: git
- Verfügbarkeit: PURL, Bibliotheks-OPAC, DARIAH-DE CR

- Dokumentieren!

# Literaturauswahl

- https://de.dariah.eu/working-papers-beitraege
- https://wiki.de.dariah.eu/pages/viewpage.action?pageId=38080370
- http://www.forschungsdaten.org/index.php/Hauptseite
- U.v.a.m.

- Dokumentieren!