
Langzeiterhalt von Kulturdaten und wissenschaftlichen Forschungsdaten

Impulsvortrag zur Session im BarCamp „Digitale Kulturvermittlung“ 2023

Eckhart Arnold <arnold@badw.de>

Referat für Digital Humanities Forschung & Entwicklung, Bayerische Akademie der Wissenschaften <www.badw.de>

Worüber wir reden...

Kulturdaten:

- Daten, die Kulturwerke von besonderem Wert abbilden
- **und** sich nicht jederzeit in derselben Qualität neu erzeugen lassen,
- **und** deren Erhalt daher über Jahrzehnte und Jahrhunderte hin geboten erscheint.

Forschungsdaten:

- Daten, die als Ergebnis wissenschaftlicher Forschung entstehen
- **oder** für die Rechtfertigung von Forschungsergebnissen erforderlich sind
- **oder** für die künftige wissenschaftliche Forschung relevant sein können.



Quelle: <https://deckenmalerei.eu> (CC-BY-NC-ND 4.0)

Forschungsdaten können zugleich Kulturdaten sein. Beispiele:

deckenmalerei.badw.de/ (Barocke Deckenmalerei)

bdo.badw.de/ (Bayerische Dialekte Online),

publikationen.badw.de/de/thesaurus (Thesaurus Linguae Latinae)

Kulturdaten sind fast immer auch Forschungsdaten!

Das Problem der digitalen Amnesie

- ein Ungeheuer mit vielen Köpfen...

Bit-Erhalt

- Hardware Schaden
- Unbeabsichtigtes Löschen
- Böses Löschen

Lesbarkeit

- Schlecht-dokumentierte / proprietäre Formate
- Wandel von Software und Datenformaten
- Flüchtigkeit der Präsentationsschicht



Das Problem der digitalen Amnesie

- ein Ungeheuer mit vielen Köpfen...

Bit-Erhalt

- Hardware Schaden
- Unbeabsichtigtes Löschen
- Böses Löschen

Backup und Kaltspeichersysteme

} Backup (Pull-Prinzip), Sicherheitschecks

Lesbarkeit

- Schlecht-dokumentierte / proprietäre Formate
- Wandel von Software und Datenformaten
- Flüchtigkeit der Präsentationsschicht

Datenexport vor dem Kauf prüfen!

{ Überblick über die *verwendeten Datenformate* und die *technische Entwicklung* behalten, ggf. Daten konvertieren

Funktion wichtiger als Optik!,
Moden vermeiden, ggf. Erneuern



Langzeitarchivierung

Technisch: Kaltspeichersysteme (billig)

Administrativ (wiederkehrende Aufgaben):

- Wird das System weiterbetrieben? Existiert der Anbieter noch?
- Müssen Datenformate gewandelt werden? (Daten-Renovierung)
- Stresstest: Gelingt es Daten zurückzuholen und zu „lesen“?

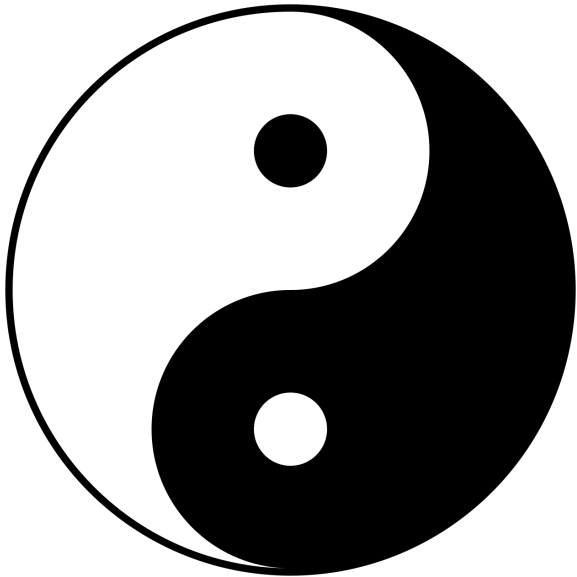
Langzeitbereitstellung (aktive Bereitstellung)

Technisch: Web-Anwendungen (laufende Kosten!)

Administrativ (Daueraufgaben):

- Kontinuierliche Pflege (Serverbetrieb, Updates, Sicherheit)
- Modernisierung der Software, insb. Präsentationsschicht
- ggf. Datenrenovierung

Möglichkeiten der Kostensenkung: nachhaltiges Softwaredesign, statische oder passive Bereitstellung (z.B. auf github), Archivierung, Kassation



Ptolemäus Arabus et Latinus:

- Edition der arabischen und lateinischen Überlieferung des antiken Astronomen Claudius Ptolemaeus (ca. 100-160 n.Chr.)

Langzeitbereitstellung:

- eher statisches, bewusst nicht zu funktionsreiches Anwendungsdesign
- Geringe Softwareabhängigkeiten (Micro-Framework)
- Vermutet dauerhafte Formate: HTML/CSS
- Voraussetzung: Eine zumindest minimale personelle Kontinuität der IT-Betreuung während der Zeit der Bereitstellung

(unter dieser Voraussetzung im Prinzip unbegrenzte Bereitstellung möglich.)



Work A.2.3

Ptolemy
كتاب أربع مقالات
Kitāb Arbaʿ maqālāt (tr. Ibrāhīm ibn al-Ṣalt/Ḥunayn ibn Iṣḥāq)

The most widespread Arabic translation of the *Tetrabiblos* is probably a revision by Ḥunayn b. Iṣḥāq (d. 260/873) of a lost translation from the Greek by Ibrāhīm b. al-Ṣalt (around 138/800), who does not appear in the manuscripts but is mentioned by Ibn al-Nadīm and Ibn al-Qifṭī. The text was most likely written under the patronship of the Abbasid caliphs in Baghdad, where also the earliest commentaries were produced. All extant Arabic commentaries on the *Tetrabiblos*, and probably also the lost ones by Thābit b. Qurra (d. 288/901) and al-Nayrīzī (d. c. 309/922), are based on this version. The work comprises: Book I, 24 chapters on general astrological principles; Book II, 13 chapters on mundane astrology; Book III, 14 chapters, and Book IV, 9 chapters, both on individual astrology and nativities. **Full title:** *Kitāb al-Arbaʿ maqālāt fī l-qaḍāʾ bi-l-nujūm ʿalā l-ḥawādīth*. **Alternative titles:** *Kitāb al-Arbaʿ, al-Maqālāt al-arbaʿ, Kitāb al-Qaḍāʾ ʿalā l-ḥawādīth, Kitāb li-l-Arbaʿ maqālāt li-Baṭlamyūs fī l-qaḍāʾ bi-l-nujūm ʿalā l-ḥawādīth*.

Note Many manuscripts include additions by Thābit b. Qurra (d. 901) in the main text or in the margins; in some cases this leads to confusion in the incipits concerning the actual revisor of the work. In one instance the treatise is even misattributed to Thābit (London, BL, Or. 9115, which has *Thābit ibn Kurrah. Arbaʿ maqālāt* on the spine). The few witnesses without Thābit b. Qurra's glosses appear to have been extracted from ʿAlī b. Riḍwān's commentary (C.2.1050; cf. the colophon in Tehran, Milli Library, Ar. 747).

Text: [Florence, BML, Or. 352]

كتاب الأربع مقالات في الأحكام لبطلموس القلوزي ترجمة حنين بن إسحاق إصلاح ثابت ابن قرّة الحرّاني
جعل ما في المقالة الأولى من هذا الكتاب في القضاء من النجوم على الحوادث. الباب الأول في صدر الكتاب (...). وإذا لم يكن
على شيء من هذه الأحوال فهي ضعيفة لا قوة لها والله أعلم. تمت المقالة الأولى من كتاب بطلموس في القضاء من النجوم على
الحوادث حمد لله وحسن وتوفيقه.

Mediathek der BAdW:

- Videos und Podcasts, in denen sich Wissenschaftler zu aktuellen Themen äußern
- Keine originäre Forschung, eher „Grauware“

Bereitstellung:

- Bereitstellung, solange aktuell (einige Jahre)
- Langfristige Bereitstellung nur solange das keinen zusätzlichen Aufwand verursacht.
- Archivierung, wenn keine Bereitstellung mehr möglich.

» Die Akademie » Presse » Mediathek

Mediathek

„Verschwörungsmythen - Hintergründe und Folgen“ - neue Podcastreihe ab 1. Mai 2023 hier in der Mediathek und bei allen gängigen Podcast-Anbietern!

In unserer Mediathek finden Sie den Audio-Podcast "BAdW-Cast", mit Interviews, Kurzvorträgen und Veranstaltungsmitschnitten. Außerdem Videos zu Projekten, Forscherinnen und Forschern der Akademie. Den Podcast gibt es auch bei allen gängigen Podcast-Anbietern wie **Apple Podcasts**, **Spotify**, **Stitcher** oder **Deezer**.

(Post-)Produktion und Redaktion: **Dr. Laura Räuber**, Fragen und Anregungen gerne an presse@badw.de.



Suchbegriff: Typ:

Selte 1 von 11.

1 | 2 | 3 | ... | 11 | Nächste



Russlands Angriff auf die Ukraine. Wie geht es weiter?

Video · 05.10.2023 · 01:18:41

Veranstaltung Politik Gesellschaft Geschichte

Seit Februar 2022 herrscht wieder Krieg in Europa. Das Leid und die Zerstörung sind groß, der Ausgang ist derzeit offen. Bei unserer Veranstaltung am 4. Oktober diskutierten Carlo Masala und Martin Schulze Wessel aus historischer und militärischer bzw. sicherheitspolitischer Perspektive über die aktuelle Entwicklung des Krieges und richteten dabei den Blick in die Zukunft.



Dialekt zwischen Kulturerbe und Klischee

Video · 13.09.2023 · 01:51:11

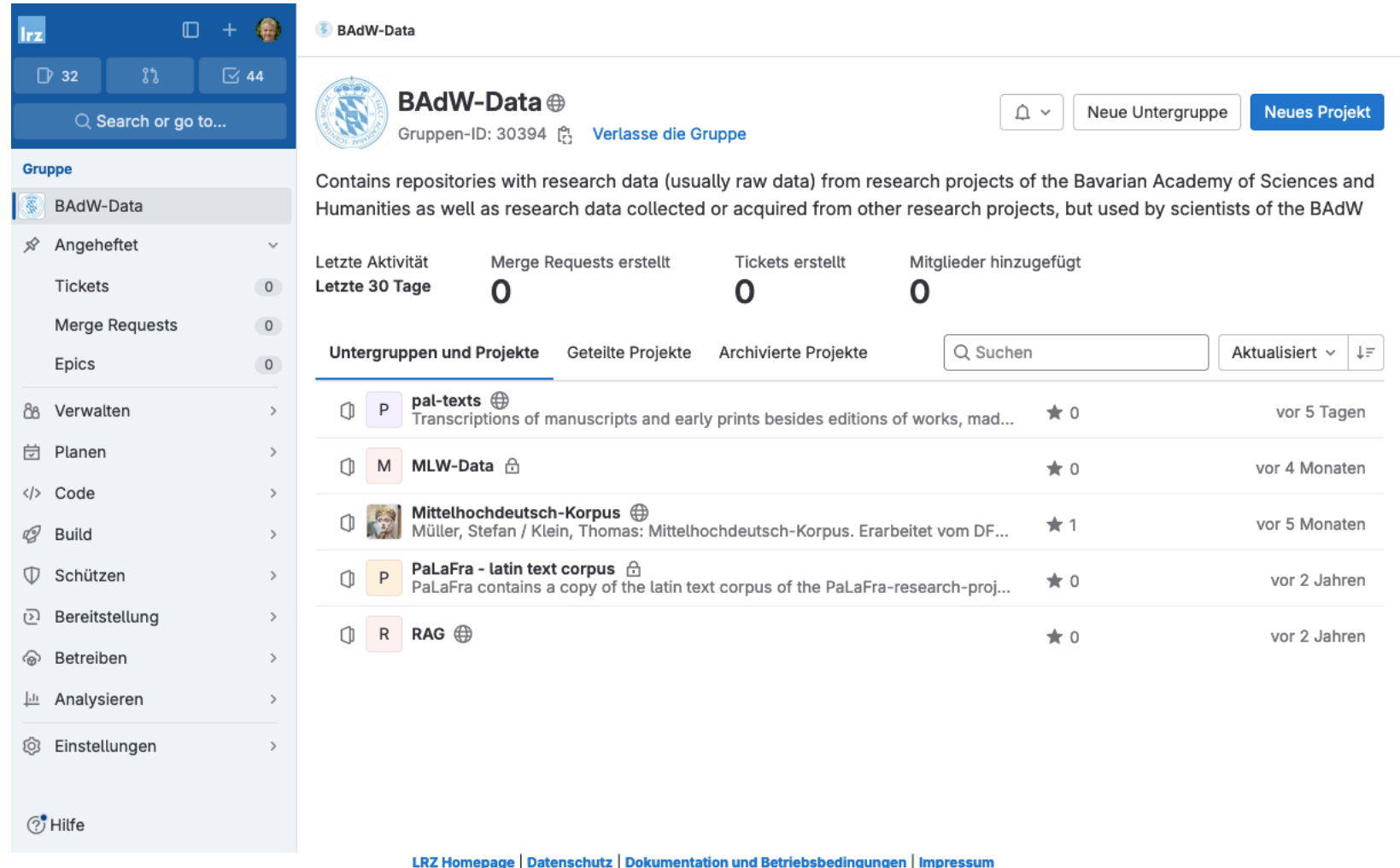
Veranstaltung Kultur

Welche Zukunft hat der Dialekt? Was können Wissenschaft, Medien und Politik zur Bewahrung und Stärkung des Dialekts in der Gesellschaft des 21. Jahrhunderts beitragen? Und warum ist es für eine Gesellschaft wichtig, dass sprachliche Vielfalt erhalten bleibt? Diese und weitere Fragen beleuchtete eine Podiumsdiskussion am 13.9. in der Akademie, die Sie hier im Video sehen können.

Beispiel 3: Passive Bereitstellung

gitlab-System für:

- Daten, die nur maschinenlesbar bereit gestellt werden müssen, bzw. für maschinenlesbare Form anderswo gehaltener Daten.
- Daten von Fremdprojekten, zu deren aktiver Bereitstellung keine Mittel vorhanden sind.
- Daten von Projekten, die nicht mehr aktiv bereit gestellt werden können (bisher und auch in absehbarer Zukunft noch nicht vorgekommen.)



The screenshot shows the GitLab interface for the 'BAdW-Data' group. The group description states: 'Contains repositories with research data (usually raw data) from research projects of the Bavarian Academy of Sciences and Humanities as well as research data collected or acquired from other research projects, but used by scientists of the BAdW'. The interface displays statistics for the last 30 days: 0 Merge Requests created, 0 Tickets created, and 0 Members added. A list of projects is shown, including 'pal-texts', 'MLW-Data', 'Mittelhochdeutsch-Korpus', 'PaLaFra - latin text corpus', and 'RAG'. The bottom of the page features navigation links for 'LRZ Homepage', 'Datenschutz', 'Dokumentation und Betriebsbedingungen', and 'Impressum'.

Offene Fragen und Probleme:

- Erhalt der mindestens notwendigen personellen Kontinuität zur Datenpflege
(in der BAAdW durch DH-Referat gewährleistet)
- Höhere Kosten der Bereitstellung von digitalen Texten im Vergleich zu Büchern
- Einschätzung der Dauerhaftigkeit von Technologien und Datenformaten gleicht einer Wette
- (Geo-)Redundante Speicherung (in Analogie zu Bibliotheken) bisher nicht Teil des Konzepts
 - Digitale Daten überleben vermutlich keinen Zivilisationsbruch