

Forschungsdaten-“Praxis“ an der LMU München

KEM-Sitzung am 5.9.2019

Sonja Kümmer

Open Data LMU (seit 2011)

<https://data.ub.uni-muenchen.de/>

- Für Angehörige aller Fakultäten
- Self-upload
- aktuell 77 Forschungsdaten

| | |
|--------------------------------------|---|
| Item Type: | Data |
| Contact Person: | Schulz, Julian |
| E-Mail of Contact: | julian.schulz at lmu.de |
| URL of Contact: | http://www.hgw.geschichte.uni-muenchen.de/schulz |
| Subjects: | Faculty of History and the Arts |
| Dewey Decimal Classification: | 900 History and geography |
| ID Code: | 110 |
| Deposited By: | Julian Schulz |
| Deposited On: | 29. Mar 2018 09:03 |
| Last Modified: | 29. Mar 2018 09:03 |

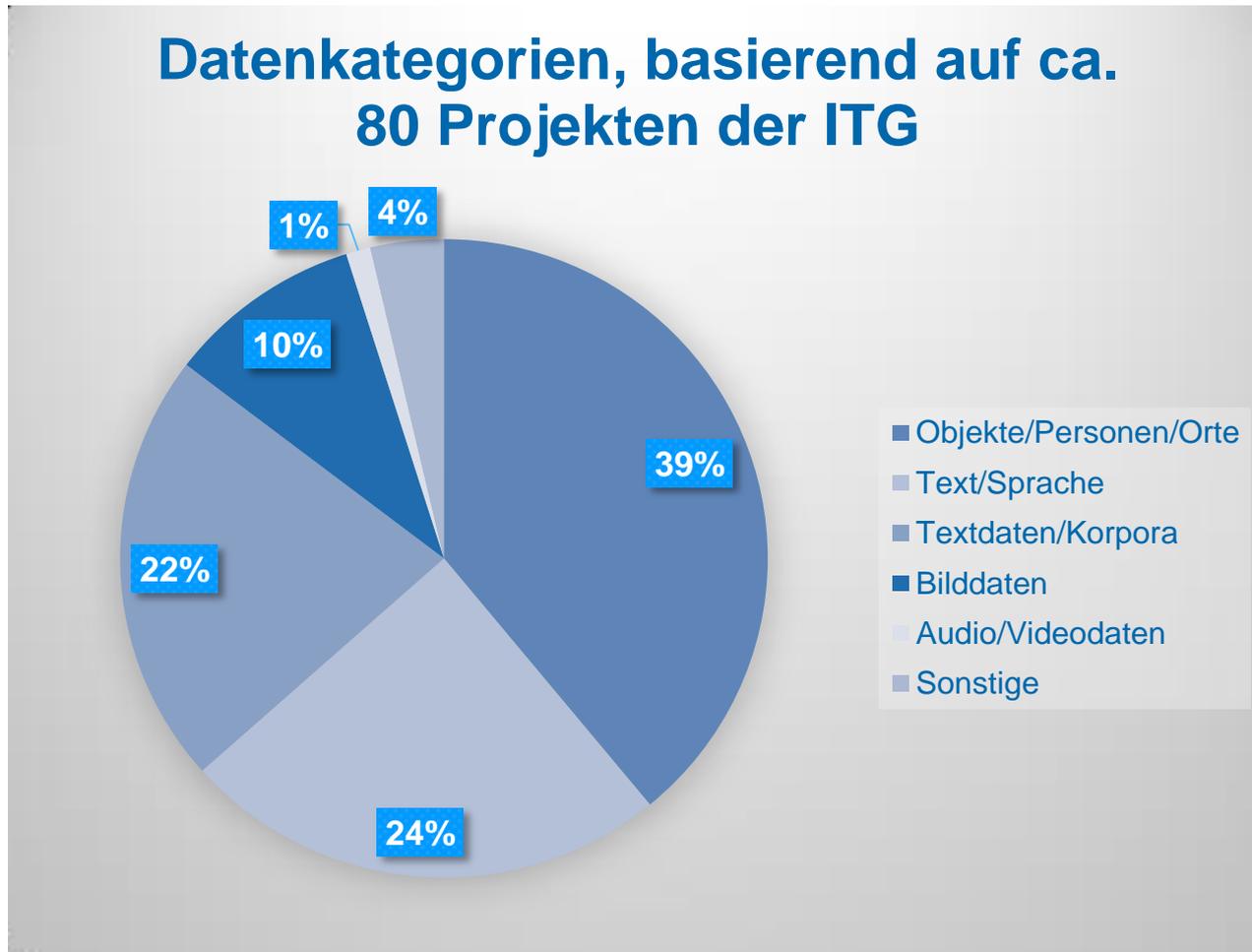
Modellprojekt „eHumanities“



- Projektpartner: UB Erlangen-Nürnberg, UB LMU, IT-Gruppe Geisteswissenschaften an der LMU (ITG)
- (Teil-)Ziel: Etablierung eines **Datenzentrums für die Digital Humanities an der LMU** mit einem Service-Angebot, das den kompletten Forschungsdatenzyklus abdeckt

<https://www.fdm-bayern.org/ehumanities-interdisziplinaer/>

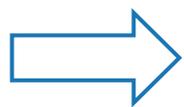
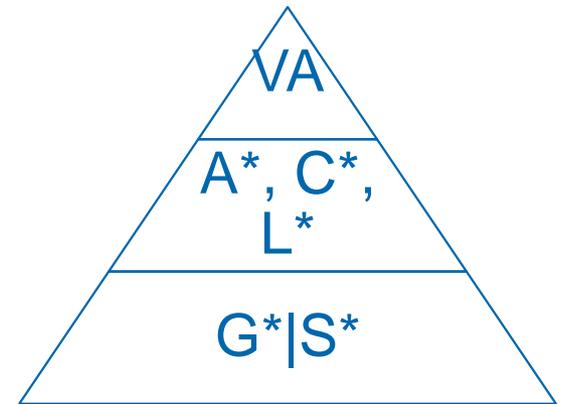
DH-Forschung an der LMU



Pilotprojekt VerbaAlpina

Was ist das Forschungsdatum?

- 5 verschiedene Informationseinheiten:
 - Gesamtprojekt (VA)
 - Gemeinden (A*)
 - Konzepte (C*)
 - Morpho-lexikalische Typen (L*)
 - Einzelbelege (G*|S*)

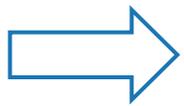


Jede Informationseinheit ist ein Forschungsdatum

Pilotprojekt VerbaAlpina

Was ist das Forschungsdatum?

- VerbaAlpina wird halbjährlich zu einer Version „eingefroren“. Jede dieser Versionen erhält eine eindeutige Versionsnummer, z.B. 18/2.
- Zusätzlich wird jeder einzelne Datensatz versioniert, z.B. C1_v2.

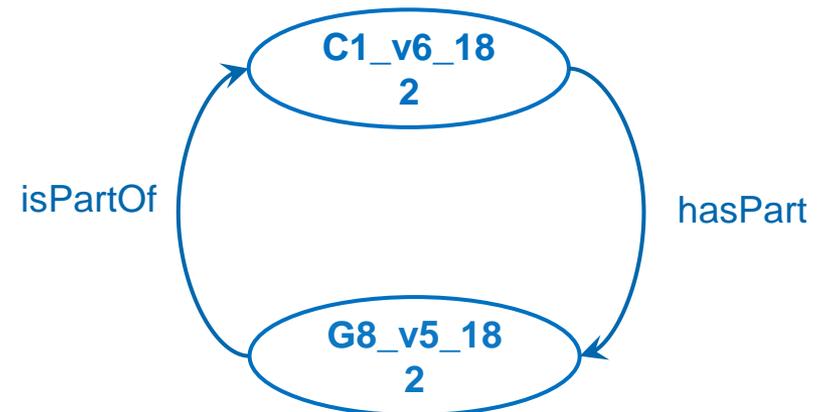
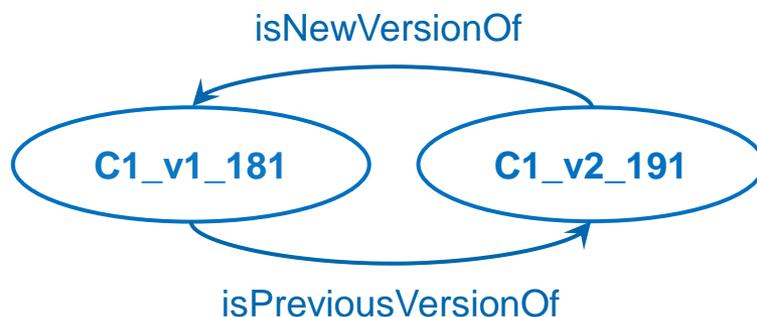


Jede Version eines Datensatzes ist ein eigenständiges Forschungsdatum

Anforderungen an die Metadaten

Projektspezifisch (VerbaAlpina):

- eindeutige Referenzierbarkeit jeder einzelnen Informationseinheit
- Zitierfähigkeit bereits während des Projekts
- Repräsentation der hatTeil- und hatVersion-Beziehungen



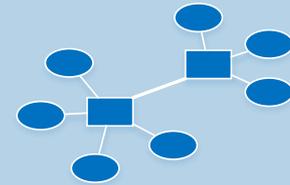
Anforderungen an die Metadaten

Projektübergreifend:

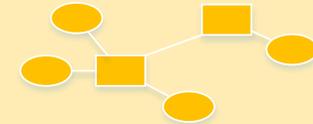
- Datenaustausch (u.a. Harvesting durch GeRDI), DOI-Registrierung » **DataCite**
- Langzeitverfügbarkeit » **PREMIS, technische Metadaten**
- maschinelle Interpretierbarkeit der (Meta)Daten » **RDF, DCAT**
- inhaltliche Erschließung » **CIDOC CRM (?)**

ToDo: Inhaltliche Erschließung

VerbaAlpina



**[Domänenspezifisches
Metadatenmodell]**



DataCite



(Automatische) Erschließung

- Definition eines **VerbaAlpina-Exportformats** samt XML Schema
- Konzeption und Implementierung einer **VerbaAlpina API** (https://www.verba-alpina.gwi.uni-muenchen.de/?page_id=8844&db=191)
- **Konzeptionelles Mapping** von VerbaAlpina nach DataCite (auf Grundlage der „DataCite Best Practices“)
- **Technisches Mapping** (XSLT-Skript)

(Meta)Datenmanagement

- Repository basierend auf Fedora 5.x
- Repräsentation der Metadaten sowohl als **xml-Dateien** als auch als **RDF triples**.
- Anreicherung der Metadaten beim Ingest um **strukturelle Metadaten** (isPartOf, isNewVersionOf, isPreviousVersionOf) und **technische Metadaten** (z.B. Checksum, Dateigröße, ...)

Normdatenanreicherung

- VerbaAlpina
 - Einige der *Konzepte* sind bereits mit Normdaten versorgt (**Wikidata**).
 - Die *Gemeinden* werden voraussichtlich eine **GeoNames-ID** erhalten.
 - Für die *Morpho-lexikalischen Typen* gibt es aktuell keine Normdatenversorgung.
- Allgemein
 - Erarbeitung von Workflows, um möglichst viele Forschungsdaten-Entitäten(typen) (global) eindeutig zu identifizieren (GND u/o Wikidata).

Persistente Identifikatoren

- DOI
 - Registrierungsformat = DataCite
 - UB LMU ist seit August 2019 **DataCite-Mitglied**
- LMU ID
 - *Idee*: einheitlicher, persistenter Identifikator für alle elektronischen Ressourcen (inklusive Forschungsdaten) der UB der LMU
 - *Ziel*: größere Unabhängigkeit (von Vergabep Praxis / Businessmodell)

Nachweissysteme

- Open Data LMU (Relaunch)
- GeRDI
- DataCite Search (+ DataCite Event Data)
- Google Dataset Search
- BASE
- Bibliothekskatalog?
- Verbundkatalog?

Velen Dank für Ihre Aufmerksamkeit!

Kontakt:

sonja.kuemmet@ub.uni-muenchen.de

forschungsdaten@ub.uni-muenchen.de