

# Ein Best-Practice-Guide für das Metadatenschema DataCite

Ein Beitrag zu mehr Interoperabilität von Metadaten

Workshop „Interoperabilität“ der AG Datenzentren  
Frankfurt/Main, 23./24.01.2020

Julian Schulz M.A.  
IT-Gruppe Geisteswissenschaften (ITG)  
Ludwig-Maximilians-Universität München

# DataCite



- Internationales Konsortium
- Generisches Metadatenschema
  - Aktuell: Version 4.3 (<https://schema.datacite.org/meta/kernel-4.3/>)
  - Umfangreiche Dokumentation (>70 Seiten)
- Warum verwenden wir es?
- Warum ein Best-Practice-Guide?

# Ausgangssituation

- Projekte der ITG weitgehend noch nicht mit Metadaten erschlossen
- Projektverantwortliche zeichneten ihre Projekte mit Hilfe des [DataCite-Generators](#) in DataCite aus
- Evaluierung der DataCite-XML-Files ergab: Trotz Verwendung des gleichen Metadatenschemas uneinheitliche, wenig interoperable Metadaten
- Beispiele für ITG-Projekte:

## DataCite Metadata Generator

**Mandatory Elements**

DOI: [DOI] ?

Title(s): [TITLE] ? +  
[LANG] ? [titleType] ▾

Creator(s): [CREATOR NAME] [nameType] ▾ ? +  
[GIVEN NAME] (optional) [FAMILY NAME] (optional)  
[NAME IDENTIFIER] [nameIdentifierSchem ▾  
[IDENTIFIER SCHEME URI] ? +  
[CREATOR AFFILIATION] [LANG] ? +

Publisher: [PUBLISHER] ?  
[LANG]

Publication Year: [YYYY] ?

Resource Type: [RESOURCE TYPE] [resourceTypeGenera ▾ ?

+ Recommended Elements

+ Other Elements

ERBA  
ALPIN



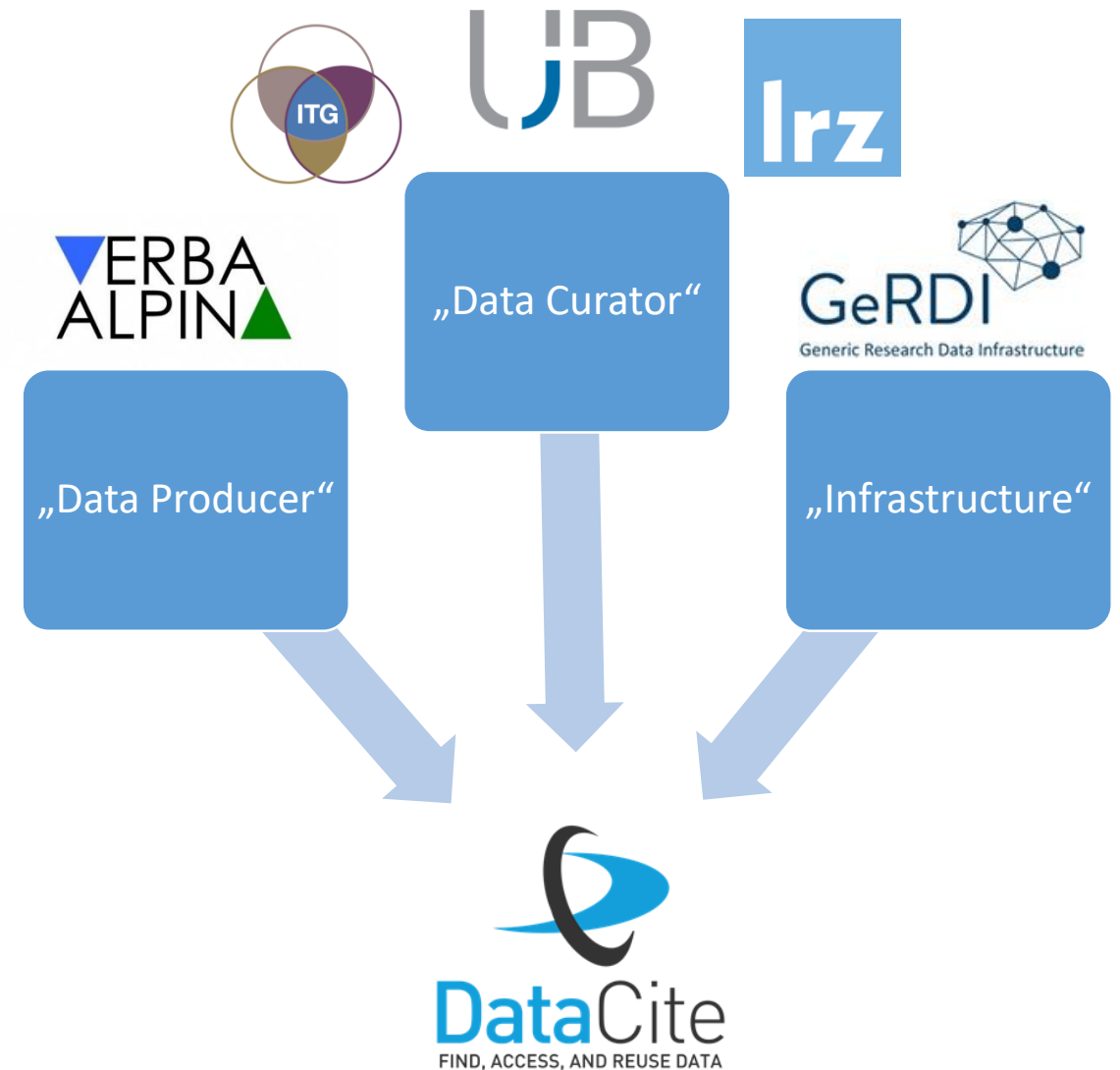
bayerisches  
musiker  
lexikon  
online



KAISER  
UND  
HÖFE

# Beteiligte:

- Sonja Kümmer (UB LMU) 
- Stephan Lücke (ITG)  
- Julian Schulz (ITG)  
- Martin Spenger (UB LMU) 
- Tobias Weber (LRZ)  



# Best-Practice-Guide für DataCite

- **Zielgruppe**

- Datenproduzenten (= Forschende)
- Datenkuratoren
- Infrastrukturanbieter

- **Aufbau**

- „General Best Practice“
- „Best Practice for specific fields“ (19 Bereiche) mit Beispielen
- Exemplarische DataCite-Datensätze

# Merkmale des Best-Practice-Guides

- **Präzisierung** des DataCite-Standards: Vorgegebener Rahmen wird enger gefasst und konkretisiert
- **Spezifikation** von Konventionen bei mehrdeutigen Eingabemöglichkeiten
- **Normierung** der Eingabemöglichkeiten: Normvokabular anstelle tendenziell heterogener Stichwörter
- **Konformität** zum DataCite-Schema wird gewahrt

# Mehrwert hinsichtlich Interoperabilität

- Homogene Daten erleichtern Aggregation und Nachnutzung:
  - Konvertierung in maschineninterpretierbare Formate (z.B. RDF)
  - Erleichterter Austausch zwischen Forschungsdateninfrastrukturen
  - Erleichtertes Mapping in andere Metadatenschemata
  - Perspektivisch: Berücksichtigung bei anderen FDM-Angeboten (z.B. RDMO) führt zur Vermeidung redundanter Arbeitsschritte
- Standardisierung der Metadaten führt zu besserer Sichtbarkeit von Forschungsdaten:
  - Beitrag zu besserer Auffindbarkeit und Nachnutzbarkeit im Sinne der FAIR-Prinzipien
  - Gesteigerter Impact für Datenproduzenten

# Vielen Dank für Ihre Aufmerksamkeit!

Best-Practice-Guide:

<https://zenodo.org/record/3559800>

Austauschplattform für Weiterentwicklung:

[https://github.com/UB-LMU/DataCite\\_BestPracticeGuide](https://github.com/UB-LMU/DataCite_BestPracticeGuide)

Metadaten-Generator:

<https://dhvlab.gwi.uni-muenchen.de/datacite-generator/>

Begleitpublikation:

<http://www.kit.gwi.uni-muenchen.de/?p=42800&v=1>