



## **Daten aus den Ländern –**

Was können Regional- und Landesbibliotheken bei der digitalen Konversion des kulturellen Erbes und der Bereitstellung von Forschungsdaten leisten?

*Prof. Dr. Thomas Stäcker  
(ULB Darmstadt/FH Potsdam)*

- Warum digital?
- Begriff der Forschungsdaten und die Anwendung der FAIR Prinzipien
- Digitale Agenda für Landesbibliotheken

# Warum digital?

- Digitale Medien sind im Prinzip ubiquitär verfügbar, beliebig ohne Verluste kopierbar und kostengünstig aggregierbar (digitale Sammlung).
- Informationsbedürfnisse werden maßgeblich aus dem Internet befriedigt
- Medien, die nicht digital vorliegen, werden nicht rezipiert
- Die Rezeption beschränkt sich nicht mehr aufs Lesen: Der Nutzer wird vom *end user* zum *end maker* (McCarty)
- Digitale Medien erlauben offenere Formen der kulturellen Teilhabe
- Digital ist bereits heute die maßgebliche Medienform
- ...

**Digital ist besser!**

## Definitionen

*Forschungsdaten sind Daten, die im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.*

Schwerpunktinitiative Digitale Information der Allianz der deutschen Wissenschaftsorganisationen:

<http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten.html>



## Definitionen

Aber:

*Sind Forschungsdaten **nur** Daten, die im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen **entstehen**?*

—→ *Daten, die nicht im Zuge wissenschaftlicher Vorhaben entstanden sind, können auch Forschungsdaten sein, z.B. Telefonbuchdaten.*



## Definitionen

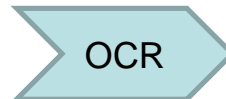
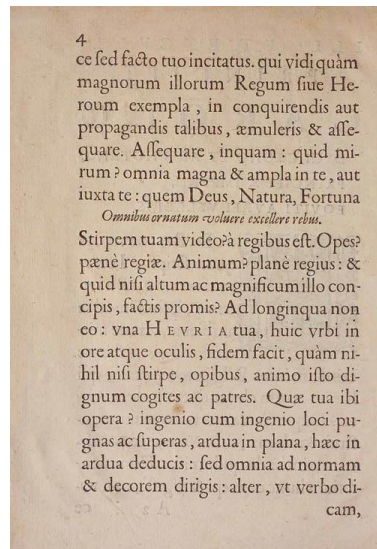
*Unter digitalen geistes- und kulturwissenschaftlicher Forschungsdaten werden innerhalb von DARIAH-DE all jene **Quellen/Materialien** und **Ergebnisse** verstanden, die im Kontext einer **geistes- und kulturwissenschaftlichen Forschungsfrage gesammelt, erzeugt, beschrieben und/oder ausgewertet** werden und **in digitaler Form vorliegen***

DARIAH WP, 2015

- Metadaten (Katalog- und Datenbankeinträge)
- Digitale Reproduktionen (images, AV, 3D)
- Datenzentrierter/dokumentenzentrierter Text (strukturiert/unstrukturiert)
- Annotationen
- ...



## Beispiel: Aufbereiten geisteswissenschaftlicher Daten - *Datafication*



ce sed facto tuo incitatus. qui vidi quàm magnorum illorum  
Regum siue Heroum exempla, in conquirendis aut  
propagandis talibus, æmuleris & assequare. Assequare,  
inquam: quid mirum? omnia magna & ampla in te, aut  
iuxta te: quem Deus, Natura, Fortuna *Omnibus ornatum  
voluere excellere rebus.* Stirpem tuam video? à regibus  
est. Opes? paenè regiae. Animum? planè regius: & quid  
nisi altum ac magnificum illo concipis, factis promissis? Ad  
longinqua non eo: vna HEURIA tua, huic vrbi in ore atque  
oculis, fidem facit, quàm nihil nisi stirpe, opibus, animo  
isto dignum cogites ac patres. Quæ tua ibi opera? ingenio  
cum ingenio loci pugnas ac superas, ardua in plana, hæc  
in ardua deducis: sed omnia ad normam & decorem  
dirigis: alter, vt verbo dicam, BELGICVS LVCVLLVS. Non  
indignare, ILLUSTRISSE PRINCEPS, cum illo componi: verà  
aestimatione, inter raros magnosque fuit: bellis  
victoriisque in florente aetate celebris: & mox inclinante,  
ad quietem & ad se inclinavit, id est animum & studia  
coluit, atque opes in praesentium ac posterorum  
delectationem siue vsum impendit. Ab illo Bibliothecarum  
Romae praecipuum exemplum; ab illo eruditi in sermone,  
mensâ, domo habiti; ab illo denique Praetoria &  
amœnitates structae, quae & Principis rei Romanae  
postea frui vtique aestimarunt. Tu, inquam, noster  
PRINCEPS imitaris. Nam in caligine hac rerum  
temporùmque, patriae & Regi operâ laudabiliter nauatâ,  
copiis triumphisque ductis; vergente iam aevo, portum  
hunc quietis & te respicis, & ab

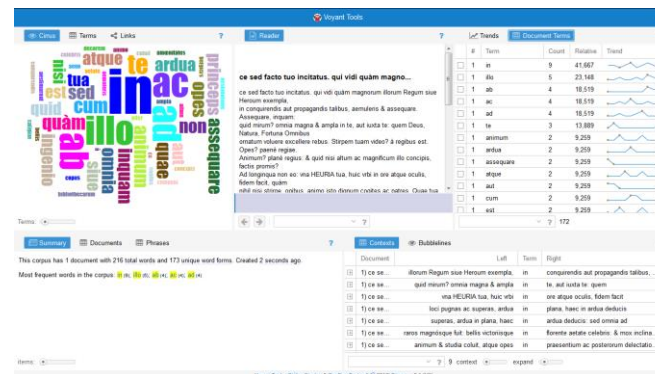




## Aufbereiten (geisteswissenschaftlicher Daten) - Datafication

ce sed factu tuo incitatus. qui vidi quàm magnorum illorum Regum siue Heroum exempla, in conquirendis aut propagandis talibus, aemuleris & assequare. Assequare, inquam: quid mirum? omnia magna & ampla in te, aut iuxta te: quem Deus, Natura, Fortuna *Omnibus ornatum voluere excellere rebus*. Stirpem tuam video? à regibus est. Opes? paenè regiae. Animum? planè regius: & quid nisi altum ac magnificum illo concipis, factis promissis? Ad longinqua non eo: vna HEURIA tua, huic vrbi in ore atque oculis, fidem facit, quàm nihil nisi stripse, opibus, animo isto dignum cogites ac patres. Quae tua ibi opera? ingenio cum ingenio loci pugnas ac superas, ardua in p... in ardua deducis: sed omnia ad normam & decore dirigis: alter, vt verbo dicam, BELGICVS LVCVLLVS. N... indignare, ILLUSTRISSIME PRINCEPS, cum illo con... aestimatione, inter raros magnòsque fuit: bellis victoriosque in florente aetate celebris: & mox inclinante, ad quietem & ad se inclinauit, id est animum & studia coluit, atque opes in praesentium ac posterorum delectationem siue vsum impendit. Ab illo Bibliothecarum Romae praecipuum exemplum; ab illo eruditi in sermone, mensà, domo habiti; ab illo denique Praetoria & amœnitates structae, quae & Principis rei Romanae postea frui vtiq; aestimarunt. Tu, inquam, noster PRINCEPS imitaris. Nam in caligine hac rerum temporumque, patriae & Regi operà laudabiliter nauatà, copiis triumphisque ductis; vergente iam aeuo, portum hunc quietis & te respicis, & ab

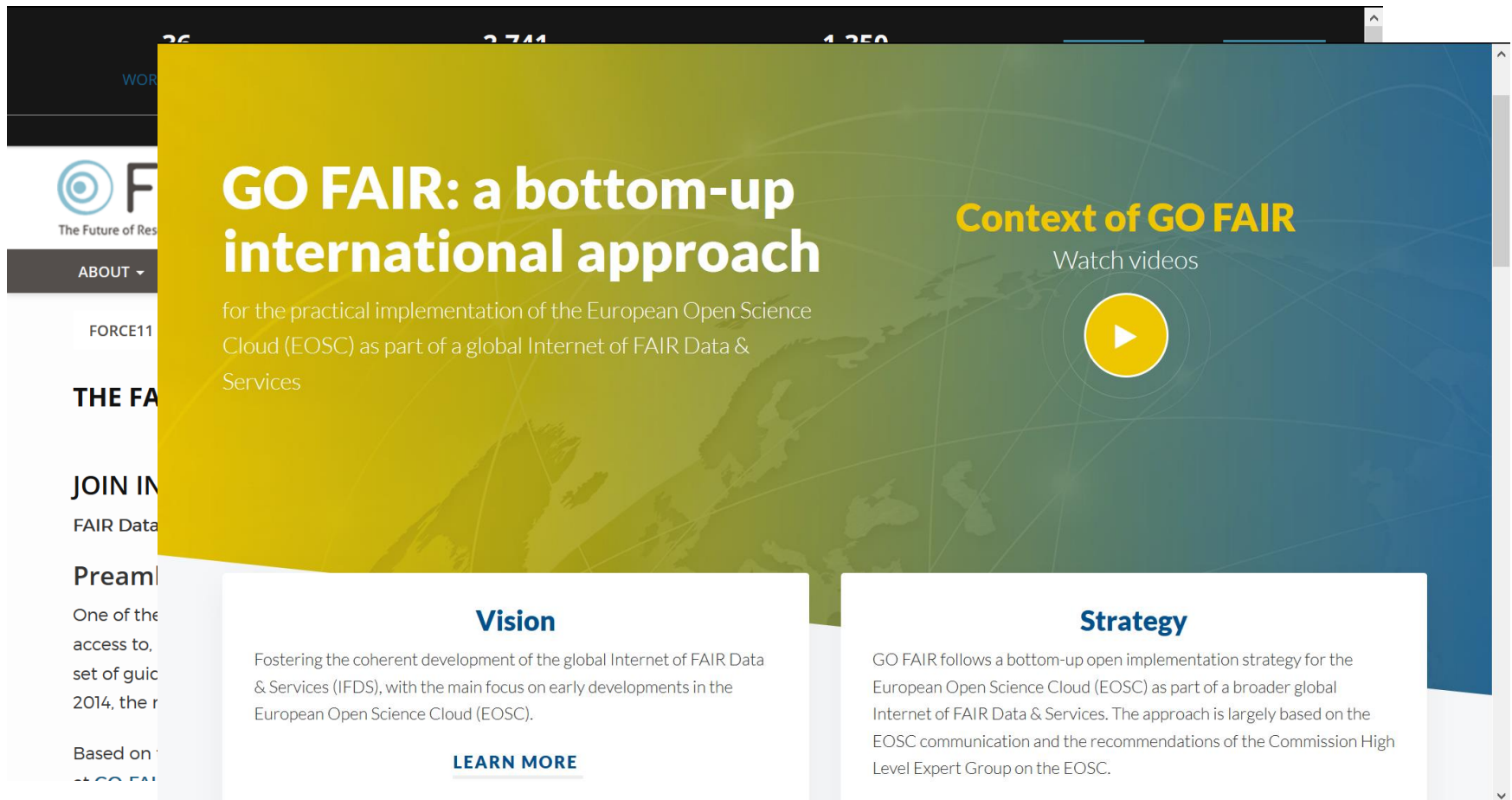
Analyse



Voyant-Tools



# Forschungsdaten FAIR



The screenshot shows the GO FAIR website with a navigation menu on the left and a main content area. The main content area features a large yellow and green graphic with the text 'GO FAIR: a bottom-up international approach' and 'Context of GO FAIR'. Below this, there are two white boxes with the headings 'Vision' and 'Strategy', each containing a paragraph of text and a 'LEARN MORE' link.

**GO FAIR: a bottom-up international approach**

for the practical implementation of the European Open Science Cloud (EOSC) as part of a global Internet of FAIR Data & Services

**Context of GO FAIR**

Watch videos

**Vision**

Fostering the coherent development of the global Internet of FAIR Data & Services (IFDS), with the main focus on early developments in the European Open Science Cloud (EOSC).

[LEARN MORE](#)

**Strategy**

GO FAIR follows a bottom-up open implementation strategy for the European Open Science Cloud (EOSC) as part of a broader global Internet of FAIR Data & Services. The approach is largely based on the EOSC communication and the recommendations of the Commission High Level Expert Group on the EOSC.

# Findable, Accessible, Interoperable, Reusable in der Praxis von RB/LB

## Findable

The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the FAIRification process.

### F1. (Meta)data are assigned a globally unique and persistent identifier

- Metadata: Bibliographische ID in URI-Form (VD16,17, 18, ISBN, persistente Record-ID ... )
- Data: DOI,URN, PURL auf definierten Granularitätsleveln

### F2. Data are described with rich metadata (defined by R1 below)

- vgl. R1

### F3. Metadata clearly and explicitly include the identifier of the data they describe

### F4. (Meta)data are registered or indexed in a searchable resource

- Verbundkatalog
- Linked Open Data



## Accessible

Once the user finds the required data, she/he needs to know how can they be accessed, possibly including authentication and authorisation.

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol

A1.1 The protocol is open, free, and universally implementable

A1.2 The protocol allows for an authentication and authorisation procedure, where necessary

A2. Metadata are accessible, even when the data are no longer available

- Der typische Zugang dürfte HTTP sein, bevorzugt REST, LOD-dumps oder klassisch OAI



# API DOKUMENTATION

Home > Hilfe > A070012

← Zurück

## Inhaltsverzeichnis

- [http/https](#)
- [Ausgabeformate und Content Negotiation](#)
- [TEI-Ausgabeformate](#)
- [Sprachauswahl](#)
- [Abfragen von GND IDs](#)
- [Correspondence Metadata Interchange-Format \(CMIF\)](#)
- [International Image Interoperability Framework \(IIIF\)](#)

## Dokument

Text
Rückverweise
XML-Vorschau
Download

[https://weber-gesamtausgabe.de/de/Hilfe/API\\_Dokumentation.html](https://weber-gesamtausgabe.de/de/Hilfe/API_Dokumentation.html)



## Interoperable

The data usually used on the web need to be interoperable for analysis, storage, and retrieval.

### 1. (Meta)data in an applicable language

➤ RDF, XML, JSON

### 2. (Meta)data in a standard format

➤ Ontologieentwurf

(bibliothekarische Sammlungen)

Sammlung: DC

### 3. (Meta)data in a standard format

➤ Hypertextualität

<https://www.w3.org/>

*Tim Berners-Lee*

*Date: 2006-07-27, last change: \$Date: 2009/06/18 18:24:33 \$*

*Status: personal view only. Editing status: imperfect but published.*

[Up to Design Issues](#)

### Linked Data

The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data.

Like the web of hypertext, the web of data is constructed with documents on the web. However, unlike the web of hypertext, where links are relationships anchors in hypertext documents written in HTML, for data they links between arbitrary things described by RDF. The URIs identify any kind of object or concept. But for HTML or RDF, the same expectations apply to make the web grow:

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF\*, SPARQL)
4. Include links to other URIs. so that they can discover more things.

Simple. In fact, though, a surprising amount of data isn't linked in 2006, because of problems with one or more of the steps. This article discusses solutions to these problems, details of implementation, and factors affecting choices about how you publish your data.

### The four rules

I'll refer to the steps above as rules, but they are expectations of behavior. Breaking them does not destroy anything, but misses an opportunity to make data







## Data C

### W3C Re

# DCAT-AP.de als formaler Metadatenstandard für offene Verwaltungsdaten bestätigt

#### This versio

<http://>

#### Latest pub

<http://>

#### Implement

<http://>

#### Previous v

<http://>

#### Editors:

[Fadi I](#)

[John](#)

#### Contributo

[Phil A](#)

#### Please refe

#### This docum

#### The English

#### Copyright © 2

#### Abstract

Die Schaffung einer einheitlichen „Metadatenstruktur für offene Verwaltungsdaten“ ist bereits seit 2013 ein Vorhaben auf der Standardisierungsagenda ([PDF](#) [↗](#)) des [IT-Planungsrates](#) [↗](#).

Ziel der Bearbeitung des Standardisierungsbedarfs ist es, mit einem einheitlichen, durch den IT-Planungsrat festgelegten Standard Metadaten künftig einfacher und umfassender austauschen zu können und so die Mehrwerte aller Datensysteme für ihre Nutzer zu erhöhen. Der Standard muss dabei so beschaffen sein, dass er alle für eine Daten-Recherche erforderlichen Informationen einbezieht und mit anderen fachlich getriebenen oder auch internationalen Formaten kompatibel ist.

Im September 2015 erhielt die Geschäfts- und Koordinierungsstelle GovData (GKSt) den Auftrag, einen Standard für ein einheitliches Metadatenschema für offene Verwaltungsdaten (Open Government Data) zu erarbeiten und eine Beschlussfassung im IT-Planungsrat vorzubereiten.

Auf der Grundlage von Analysen und Befragungen wurde Anfang 2016 die Entscheidung getroffen, ein eigenes Metadatenmodell OGD 2.0 als Standard für Deutschland zu entwickeln. Dieses Modell wurde im Juni 2016 auf [Joinup](#) [↗](#) öffentlich zur Diskussion gestellt.

Das Ergebnis dieser Diskussion, sowie zahlreiche relevante aktuelle Entwicklungen in Bezug auf den von der EU betriebenen Standard DCAT-AP, die auch während der öffentlichen Beteiligungsphase zu dem vom Projekt bisher entwickelten Metadatenmodell zu Bedenken gegeben wurden, haben uns dazu bewogen, den bisher gewählten Standardisierungspfad vor dem Hintergrund der neuen Erkenntnisse kritisch zu prüfen.

Auf dieser Grundlage hat die Fachgruppe GovData am 21. November 2016 beschlossen, dass als deutscher Metadatenstandard eine exakte DCAT-AP konforme deutsche Ableitung von DCAT-AP erstellt wird, die eine direkte Kompatibilität zum EU-Standard sicherstellt.

DCAT-AP.de wurde gemeinsam mit dem Freistaat Sachsen und der ]init[ AG entwickelt und in der Version 1.0 auf [dcat-ap.de](#) [↗](#) veröffentlicht (inzwischen liegt eine Version 1.0.1 vor). Auf dieser Plattform finden Sie die Spezifikationsdokumente in der jeweils aktuellen Version, Beispieldateien, sowie die DCAT-AP.de eigenen Vokabulare.

**Der IT-Planungsrat hat auf seiner 26. Sitzung im Juni 2018 DCAT-AP.de als formales Austauschstandard für offene allgemeine Verwaltungsdaten festgelegt**



## Reusable

The ultimate goal of FAIR is to optimise the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

[R1. Meta\(data\) are richly described with a plurality of accurate and relevant attributes](#)

[R1.1. \(Meta\)data are released with a clear and accessible data usage license](#)

➤ urheberrechtsrelevante Materialien: CC BY

➤ urheberrechtsfreie Materialien: CC0 (Quellen und Datenbanken)

[R1.2. \(Meta\)data are associated with detailed provenance](#)

➤ Vgl. PROV-DM (W3C) , SammlungsID (vgl. ISIL + SammlungsID),

[R1.3. \(Meta\)data meet domain-relevant community standards](#)

➤ Dynamische Angebote auf der Grundlage digitaler Sammlungen (Datenaufbereitung: Aggregation, Selektion, Verknüpfung, Anreicherung ...)

# Digitale Agenda für Landesbibliotheken

- Systematische (koordinierte) Konversion des gedruckten Erbes in eine maschinenlesbare Form nach den FAIR Prinzipien
- Landesbibliotheken übernehmen nach dem Sprengelprinzip die Digitalisierung in der Fläche (vgl. HSS Programm der DFG)
- Konsequente Umsetzung des Open Access Prinzips
- Aufbau digitaler Sammlungen durch „Aggregation“ (Kopieren) und „Anreichern“
- Entwicklung eines Konzeptes der Bibliothek als Schnittstelle
- Die digitale Bibliothek als Ort der Citizen Science und kulturellen Teilhabe
- Umbau der Bibliothek zu einem Ort der *datafication*. Anpassung von Sammlungs-, Erschließungs-, und Benutzungsworkflows.



**<emph>  
Herzlichen Dank  
für Ihre Aufmerksamkeit!  
</emph>**