

# Forschungsdatenmanagement

Fachhochschule für öffentliche Verwaltung  
und Rechtspflege in Bayern, SS 2014

Michael Franke, Max Planck Digital Library



# Wofür Forschungsdatenmanagement?

- Gute wissenschaftliche Praxis
- Wissenschaftliches Renommee
- Nachnutzung der Daten

# Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

# Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

# Beispiele für Forschungsdaten

# Beispiele für Forschungsdaten

## Bohrkerne

## LfU Bayern

The screenshot shows the website of the Bayerisches Landesamt für Umwelt. The main content area is titled 'Bohrkernarchiv' and contains the following text:

**Bohrkernarchiv**

Das Bohrkerarchiv dient der zentralen Lagerung von Bohrproben, Gesteinsproben sowie Bodenproben aus der rohstoffgeologischen, geologischen und bodenkundlichen Erfassung Bayerns. Zudem sind spezielle Forschungsbohrungen, wie u.a. die Bohrung Altdorf / Landshut, die Thermalbohrung Obersees oder die Bohrung Nördlingen 1 hier eingelagert.

Daneben beherbergt es eine Belegsammlung von Proben aus Bergrechtsverfahren und bodenkundlichen Projekten zum Zweck der Beweissicherung sowie eine Naturwerkstein-Sammlung zur Dokumentation ehemaliger bzw. noch aktiver Gewinnungsteilen von ganz Bayern.

Die Proben werden in einem Hochregallager auf Paletten eingelagert. Insgesamt stehen über 3 000 Palettenstellplätze zur Verfügung. Somit können aneinander gereiht insgesamt ca. 150km Bohrungen archiviert werden.

Aktuell sind ca. 74km Bohrkerstrecke sowie 250 Paletten mit Bodenproben eingelagert. Die Beschickung und Entnahme erfolgt mittels eines Elektro-Kommissionierers und einem Gabelstapler.

Die Lagerverwaltung wird EDV-gestützt abgewickelt, die Bohrungen sind im Bodeninformationssystem Bayern (BIS) für jedermann online recherchierbar.

Below the text are three images: a high-bay warehouse, a forklift loading a pallet, and a close-up of pallets with boxes. A caption reads: 'Beispiele für eingelagertes Archivgut (von links nach rechts: Spitzproben, Bohrkerne, Liner, Bodenproben)'.

Under the heading 'Weiterführende Informationen', there is a link to 'Bodeninformationssystem Bayern (BIS)'.

The footer contains navigation menus for 'Themen', 'Service', 'Wir', and 'Presse'.

# Beispiele für Forschungsdaten

## Gravitationswellen

## BICEP2

### BICEP2 2014 Results Release

1. BICEP2 2014 Release Papers
2. BICEP2 2014 Release Data Products
3. BICEP2 2014 Release Figures from Papers
4. Previous Publications

[Videos for Technical Talk](#)  $\Rightarrow$  and [News Conference](#)  $\Rightarrow$   
[BICEP2 2014 Release Frequently Asked Questions](#)  $\Rightarrow$   
[BICEP2 2014 Release Image Gallery](#)  $\Rightarrow$   
[BICEP2 Public Web Pages and News Releases](#)  $\Rightarrow$

#### 1. BICEP2 2014 Release Papers

**BICEP2 2014 I: Detection of B-mode Polarization at Degree Angular Scales**  
*The BICEP2 Collaboration, 2014*

[PDF](#) / [figures](#) / [arXiv](#) / [ADS](#)

**BICEP2 2014 II: Experiment and Three-year Data Set**

*The BICEP2 Collaboration, 2014*

[PDF](#) / [figures](#) / [arXiv](#) / [ADS](#)

#### 2. BICEP2 2014 Release Data Products

<a href="#">B2_3yr_rlikelihood_20140314.txt</a>	Text file containing the tabulated likelihood for the tensor-to-scalar ratio, $r$ , computed using the "direct likelihood calculation" described in Section 9.3.1 of Barkats <i>et al.</i> and Section 11.1 of BICEP2 2014 I.
<a href="#">B2_3yr_bandpowers_20140314.txt</a>	Text file containing bandpowers and statistical uncertainties, corresponding to Figure 2 of BICEP2 2014 I.
<a href="#">B2_3yr_bandpower_window_functions.tgz</a> <a href="#">B2_3yr_bpwf_bin1_20140314.txt</a> <a href="#">B2_3yr_bpwf_bin2_20140314.txt</a> <a href="#">B2_3yr_bpwf_bin3_20140314.txt</a> <a href="#">B2_3yr_bpwf_bin4_20140314.txt</a>	Bandpower window functions for each $l$ bin.

# Beispiele für Forschungsdaten

## Gensequenzen

The screenshot displays the NCBI GenBank website interface. At the top, there is a search bar with 'Nucleotide' selected and a search button. Below the search bar, the title of the entry is 'Homo sapiens class II AP endonuclease (APE) gene, partial CDS'. The GenBank accession number is M99703.1. The entry includes various fields such as LOCUS, DEFINITION, ACCESSION, VERSION, KEYWORDS, SOURCE, ORGANISM, REFERENCE, JOURNAL, PUBMED, and FEATURES. The FEATURES section shows the location and qualifiers for the source, intron, gene, mRNA, exon, and CDS. On the right side, there are several interactive options: 'Change region shown', 'Customize view', 'Analyze this sequence' (with sub-options like Run BLAST, Pick Primers, Highlight Sequence Features, Find in this Sequence), 'Articles about the APEX1 gene' (with sub-options like Expression and prognostic significance of APE1/Ref-1 and NPM1, Dysregulation of human apurinic/aprimidinic endonuclease 1 (APE1), PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR), and 'Reference sequence information' (with sub-option RefSeq alternative splicing). At the bottom right, there is a section for 'More about the APEX1 gene' with a brief description and a list of also known names.

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Limits Advanced Help

Display Settings: GenBank Send:

**Homo sapiens class II AP endonuclease (APE) gene, partial CDS**  
GenBank: M99703.1  
[FASTA](#) [Graphics](#)

Go to:

LOCUS HUMAPEB 3019 bp DNA linear PRI 31-DEC-1994  
DEFINITION Homo sapiens class II AP endonuclease (APE) gene, partial CDS.  
ACCESSION M99703  
VERSION M99703.1 GI:178748  
KEYWORDS 3' diesterase associated activity; DNA repair protein; DNA-binding;  
class II AP endonuclease; nuclear localized protein.  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Cataarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Harrison,L., Ascione,G., Menninger,J.C., Ward,D.C. and Demple,B.  
TITLE Human apurinic endonuclease gene (APE): structure and genomic  
mapping (chromosome 14q11.2-12)  
JOURNAL Hum. Mol. Genet. 1 (9), 677-680 (1992)  
PUBMED [1284593](#)

FEATURES  
Location/Qualifiers  
source 1..3019  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
intron 596..778  
/gene="APE"  
/number=1  
gene join(779..904,1115..1302,1869..2061)  
/gene="APE"  
mRNA join(779..904,1115..1302,1869..2061)  
/gene="APE"  
exon 779..904  
/gene="APE"  
/number=2  
CDS join(847..904,1115..1302,1869..2061)

Change region shown

Customize view

Analyze this sequence

Run BLAST  
Pick Primers  
Highlight Sequence Features  
Find in this Sequence

Articles about the APEX1 gene

Expression and prognostic significance of APE1/Ref-1 and NPM1 j [Am J Clin Pathol. 2014]  
Dysregulation of human apurinic/aprimidinic endonuclease 1 (APE1) x [Er J Ophthalmol. 2014]  
PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR [Mol Cell. 2014]  
See all...

Reference sequence information

RefSeq alternative splicing  
See 6 reference mRNA sequence splice variants for the APEX1 gene.

More about the APEX1 gene

Apurinic/aprimidinic (AP) sites occur frequently in DNA molecules by spontaneous hydrolysis, by DNA damaging agents or by DNA glycosylases  
...  
Also Known As: APE, APE1, APEN, APEX, ...

## Genbank



# Beispiele für Forschungsdaten

## Nahrungsmittelanalysen

**USDA**  
Agricultural Research Service  
United States Department of Agriculture

National Nutrient Database for Standard Reference  
Release 26

[NDL Home](#) [Food Search](#) [Nutrients List](#) [Ground Beef Calculator](#) [Contact Us](#) [SR26 Documentation](#) [Help](#)

**FOOD GROUP**

Baby Foods (1)  
Cereal Grains and Pasta (6)  
Fast Foods (1)  
Meals, Entrees, and Side Dishes (1)  
Vegetables and Vegetable Products (17)

Enter one or more terms and click 'Go' for a new search

26 foods found. Click on a food name to view details

1 2 Next

NDB No.	Description	Food Group
11658	Spinach souffle	Vegetables and Vegetable Products
11457	Spinach, raw	Vegetables and Vegetable Products
11986	Malabar spinach, cooked	Vegetables and Vegetable Products
20126	Spaghetti, spinach, dry	Cereal Grains and Pasta
20127	Spaghetti, spinach, cooked	Cereal Grains and Pasta
11274	Mustard spinach, (tendergreen), raw	Vegetables and Vegetable Products
11276	New Zealand spinach, raw	Vegetables and Vegetable Products
03127	Babyfood, vegetables, spinach, creamed, strained	Baby Foods
11458	Spinach, cooked, boiled, drained, without salt	Vegetables and Vegetable Products
11463	Spinach, frozen, chopped or leaf, unprepared	Vegetables and Vegetable Products
11854	Spinach, cooked, boiled, drained, with salt	Vegetables and Vegetable Products
20095	Pasta, fresh-refrigerated, spinach, as purchased	Cereal Grains and Pasta
20096	Pasta, fresh-refrigerated, spinach, cooked	Cereal Grains and Pasta
20111	Noodles, egg, spinach, dry, enriched	Cereal Grains and Pasta
20112	Noodles, egg, spinach, cooked, enriched	Cereal Grains and Pasta
21500	KASHI Pizza, Mushroom Trio & Spinach	Fast Foods
11275	Mustard spinach, (tendergreen), cooked, boiled, drained, without salt	Vegetables and Vegetable Products
11277	New Zealand spinach, cooked, boiled, drained, without salt	Vegetables and Vegetable Products
11459	Spinach, canned, regular pack, solids and liquids	Vegetables and Vegetable Products
11461	Spinach, canned, regular pack, drained solids	Vegetables and Vegetable Products
11801	Mustard spinach, (tendergreen), cooked, boiled, drained, with salt	Vegetables and Vegetable Products
11802	New zealand spinach, cooked, boiled, drained, with salt	Vegetables and Vegetable Products
11855	Spinach, canned, no salt added, solids and liquids	Vegetables and Vegetable Products
11464	Spinach, frozen, chopped or leaf, cooked, boiled, drained, without salt	Vegetables and Vegetable Products
11856	Spinach, frozen, chopped or leaf, cooked, boiled, drained, with salt	Vegetables and Vegetable Products

USDA

# Beispiele für Forschungsdaten

## Tierbewegungen

Movebank

The screenshot displays the Movebank website interface. At the top, the Movebank logo is visible with the tagline "For Animal Tracking Data". The navigation menu includes Home, Tracking Data Map, Community, Help, Tools, Env-DATA, and Published Data. A search bar is located in the top right corner.

On the left side, there is a "User login" section with fields for Username and Password, and a "Log in" button. Below the login fields are links for "Create new account" and "Request new password".

The main content area is divided into two sections. The left section is titled "Search" and contains an "Advanced Search" form with a search input field containing the word "stork". Below the search input are dropdown menus for "All Sensor Types" and a checkbox for "Only studies where I can see data". A "Search" button is located at the bottom right of the search form.

The right section is titled "Search result" and displays a list of search results. The results are sorted by "Animal Identifier". The list includes various studies such as "TBMU\_Digges\_Glchirat\_2013", "TBMUCOMU GastonMontevicchiNWAatlantic", "Thick-billed Murres: Glchirat, Digges Island, Canada", "Turkey Vulture Acopian Center USA GPS", "Turkey Vulture South Carolina USA", "Turkey vultures in North and South America (data fro", "USFWS California Cander Recovery Program: ARGOT", "USGS Woodcock Migration", "Vultures (Black & Turkey) Pharr Texas", "Wellstorfzug [n=13]", "Amelios, [n=436], Ciconia ciconia", "Amelios I, [n=217], Ciconia ciconia", "Bruno, [n=843], Ciconia ciconia", "Dani, [n=2619], Ciconia ciconia", and "Dusches, [n=92], Ciconia ciconia". Each result has a small icon and a search icon.

On the right side of the search results, there is a map of Europe showing the movement paths of tracked animals. The map includes labels for various countries and regions, such as Deutschland (Germany), France, Italia (Italy), and España (Spain). The map also shows the Mediterranean Sea and the Atlantic Ocean. The map interface includes a search bar, a zoom control, and a "Link" button.

At the bottom of the page, there is a footer with the text "Movebank is coordinated by the Max Planck Institute for Ornithology, the North Carolina Museum of Natural Sciences, and the University of Konstanz. For questions or feedback, contact us at support@movebank.org".

# Beispiele für Forschungsdaten

## Gesichter/Mimik

Max Planck Institute for Human Development

HOME PICTURES ALBUMS

Sort Organize

Sorted by Emotion Person-ID Picture Set Sort

12 of 16 hits


Go to page of 2

### Sad, Happy

Published  
http://faces.mpdl.mpg.de/faces/album/escidoc:102107

Dates  
Date Created: 2009-10-08 at 16:36:05 CEST  
Date published: 2009-10-08 at 16:48:24 CEST

Author  
Noh, SooRim<sup>1</sup>  
Affiliation  
1 Brandeis University  
Description  
Sad, Happy



View View View View

View View View View

View View View View

Info

This album contains 444 private pictures which are only visible for logged in users.  
If you already have an account, please [Login](#).  
If you are interested in an account for special research questions, you can apply for an account by filling in the [application form](#).

12 of 16 hits

Go to page of 2

## Faces

# Definitionsversuche I

- „Ganz allgemein gesprochen sind Forschungsdaten Daten, die im Forschungsprozess erzeugt, gesammelt oder zusammengestellt werden und auf deren Grundlage wissenschaftliche Hypothesen, Modelle oder Theorien gebildet werden“ (Uni Heidelberg).
- „Forschungsdaten stellen im weitesten Sinne Primärdaten, Sekundäranalysen, Visualisierungen, Modelle, Analysewerkzeuge, Objektsammlungen oder Produkte dar, die während des wissenschaftlichen Arbeitsprozesses erzeugt und benutzt werden“ (Johanna Vompras, Uni Bielefeld).
- *„Forschungsdaten sind Daten, die im Zuge wissenschaftlicher Vorhaben z.B. durch Digitalisierung, Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.“ (Allianz AG „Forschungsdaten“)*
- *„(...) researchers know within their projects what research data is“ (Leibniz-Gemeinschaft)*

# Definitionsversuche II

- *Forschungsdaten sind jene Fakten, auf denen wissenschaftliche Publikationen gründen.*

# Klassifikation von Forschungsdaten

## Nach dem Charakter der Daten

Qualitative Daten

Bilder (Scans, Fotos, Mikroskopdaten, Teleskopdaten, Satellitendaten)

Multimedia (Audiodaten, Videodaten, 3D-Daten, 4D-Daten)

Zahlenreihen (Umfragedaten, Experimentaldaten, Sensormessreihen, Gensequenzen, Geodaten)

Born-digitals (Simulationsdaten, Algorithmen, Websites)

Quantitative Daten

# Klassifikation von Forschungsdaten

## Nach dem Format der Daten

Bilder (TIFF, JPEG, ...)

Multimedia (MP4, Dicom, ...)

Messdaten (TXT, XLSX, ...)

Born-digitals (netCDF, grib, ...)

# Klassifikation von Forschungsdaten

Nach der Quelle der Daten  
bzw. der Art der Forschung

**Observational:** data captured in real time that is usually unique and irreplaceable. For example, remote sensing data, survey data, field recordings, sample data.

**Experimental:** data captured from lab equipment that is often reproducible. For example, gene sequences, chromatograms, magnetic field data.

**Models or simulation:** data generated from test models where model and metadata may be more important than output data from the model. For example, climate models, economic models.

**Derived or compiled:** resulting from processing or combining 'raw' data. For example, text and data mining, compiled databases, 3D models.

**Reference or canonical:** a static or organic conglomeration or collection of datasets, probably published and curated. For example, gene sequence databanks, collection of letters.

(University of Bristol, Boston University)



# Klassifikation von Forschungsdaten

## Nach dem Grad der Aggregation/Verdichtung

**Rohdaten:** Als Rohdaten soll die Gesamtmenge aller Daten bezeichnet werden, die empirischer Wissenschaft als Forschungsgrundlage zur Verfügung stehen.

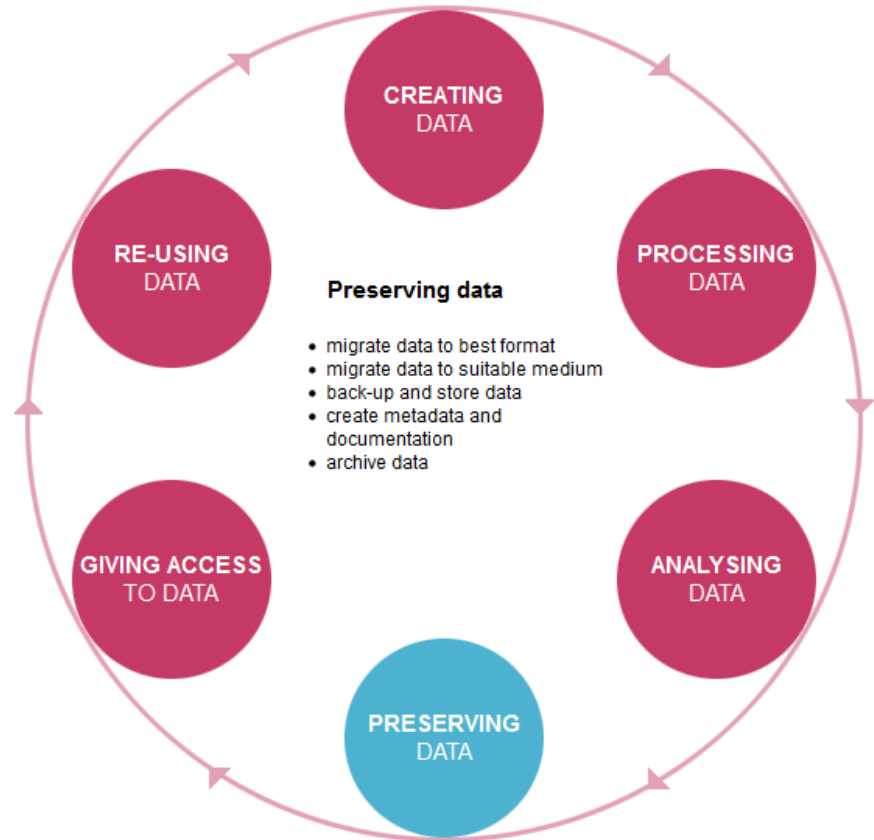
**Primärdaten:** Primärdaten seien diejenigen Daten, die als Teilmenge der Rohdaten tatsächlich zur Forschung herangezogen werden.

**Sekundärdaten:** Sekundärdaten seien Daten, die in Prozessschritten aus Primärdaten gewonnen wurden.

**Tertiärdaten:** Als Tertiärdaten sollen Informationen bezeichnet werden, die nicht aus den Primärdaten herleitbar sind, aber im wissenschaftlichen Prozess anfallen.

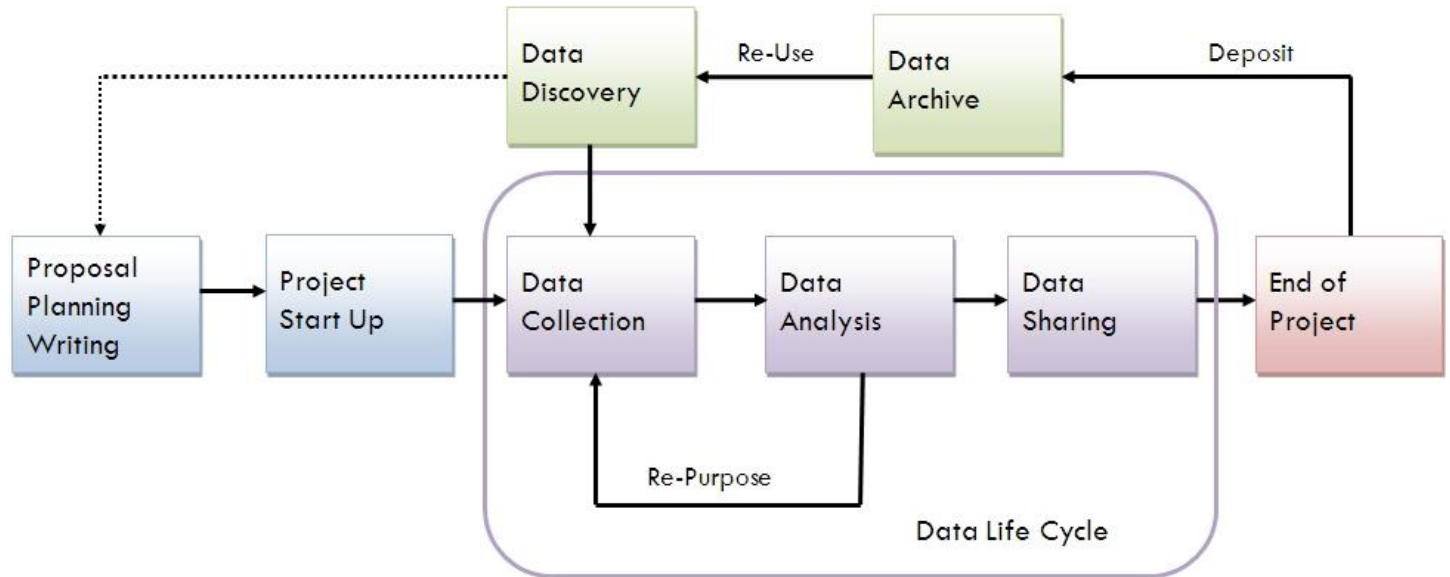
(Franke)

# Forschungsdatenzyklus



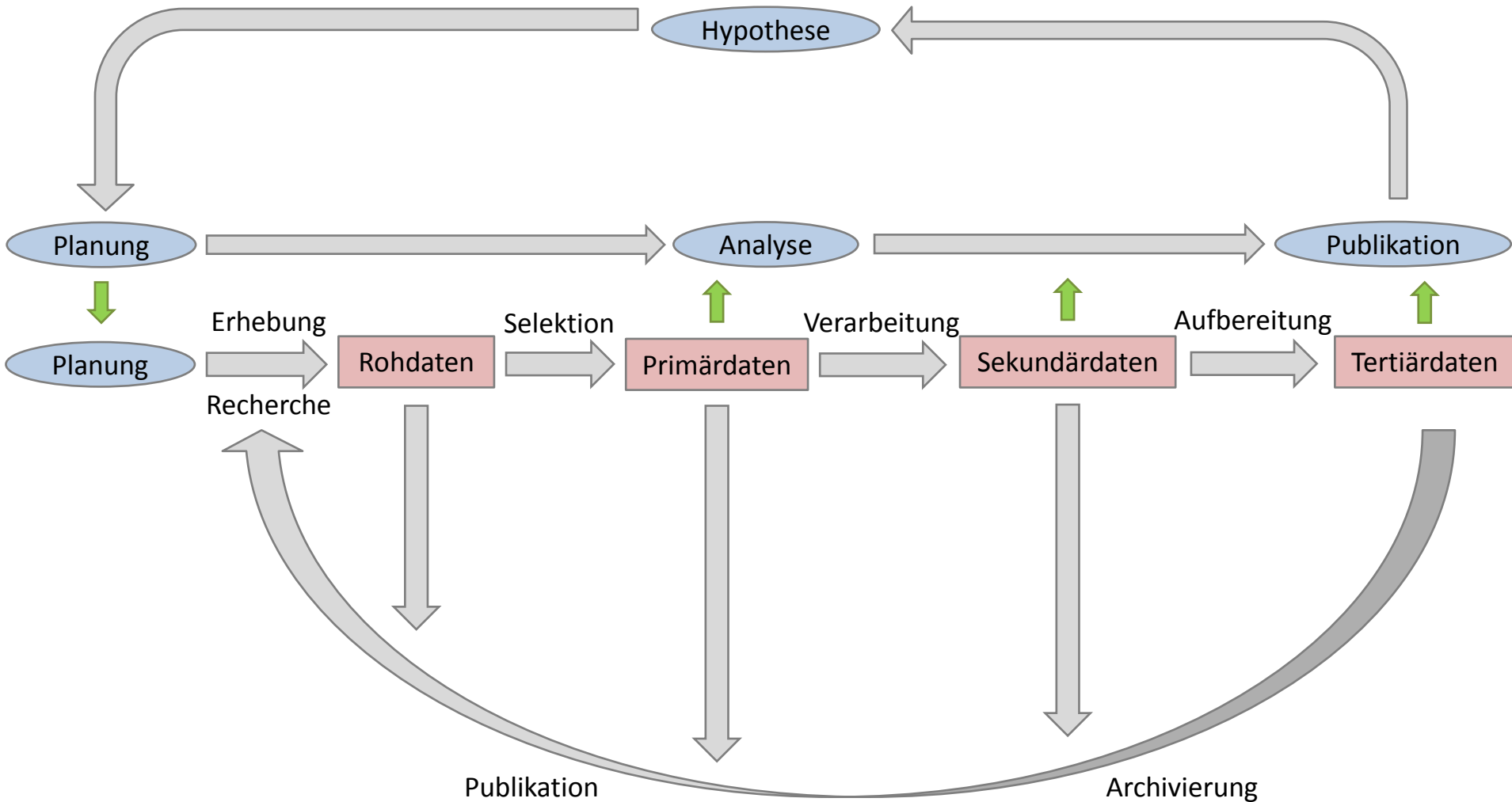
UK Data Archive

# Forschungsdatenzyklus



University of Virginia

# Forschungsdatenzyklus



# Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

# Infrastruktur?

- z.B. Verkehrsinfrastruktur
  - Straßen
  - Schienen, Wasserwege
  - Tankstellennetz, Verkehrsfunk, GPS
  - Straßenverkehrsordnung, Straßenwacht, ADAC
  - Räumdienst, Mautsystem, KFZ-Steuer
  - Autoindustrie, Verkehrsministerium, ÖPNV

# Forschungsdateninfrastruktur

- Wiss. Großanlagen
- Metadaten, Standards, Identifikatoren
- Repositorien, Registries, Datenjournale
- Rechtsgrundlagen
- Wissenschaftler, Universitäten, Wissenschaftsorganisationen, Förderer, Regierung, Gremien, Allianzen, Rechenzentren, Verlage, Archive, Sammlungen, Museen, Bibliotheken

# Wissenschaftliche Großanlagen

- z.B. [LHC](#), [Polarstern](#), [Square Kilometre Array](#)
- [E-ELT](#)
  
- Erheben/Erzeugen Daten
- Verteilen Daten
- Zwingen zur Kooperation



# Metadaten

- [DC](#) – Dublin Core Metadata Initiative
- [RDF](#) – Resource Description Framework
- [SKOS](#) – Simple Knowledge Organization System
- [Metadaten-Registries](#)

# Standards

- Metadaten
  - Daten-/Dateiformate
  - Protokolle/Schnittstellen
  - Lizenzen
- Pronom  
IANA  
Isocat
- HTTP  
REST  
OAI-PMH  
Z39.50
- Creative Commons  
Open Data Commons  
Public Domain

# Persistente Identifikation

- URL/URI <http://zuse.zib.de/collection/wl3XoEDHO8v0ImCa/item/VqNgKUacPrIhqPKu>
- URN <urn:nbn:de:bsz:25-opus-14124>
- Handle-System <hdl:11858/00-001M-0000-0019-D20F-6>
- DOI [10.1088/0004-637X/715/2/1453](https://doi.org/10.1088/0004-637X/715/2/1453)
- Researcher-ID [G-2512-2011](https://orcid.org/0000-0002-2661-8242)
- ORCID [0000-0002-2661-8242](https://orcid.org/0000-0002-2661-8242)



# Forschungsdatenrepositorien I

Institutionell

Fachlich

Allgemein

[Open Data LMU](#)

[PsychData](#)

[DANS Easy](#)

[Dryad](#)

[Edinburgh DataShare](#)

[Pangaea](#)

[Zenodo](#)

# Forschungsdatenrepositorien II

- Daten aufnehmen
- Langzeitarchivierung
- Daten auffindbar machen
- Daten exportieren

# Datenjournale

- Lebenswissenschaften
  - [Gigascience](#)
  - [F1000](#)
  - [Open Health Data](#)
- Geowissenschaften
  - [Earth System Science Data](#)
- Naturwissenschaften
  - [Scientific Data](#)

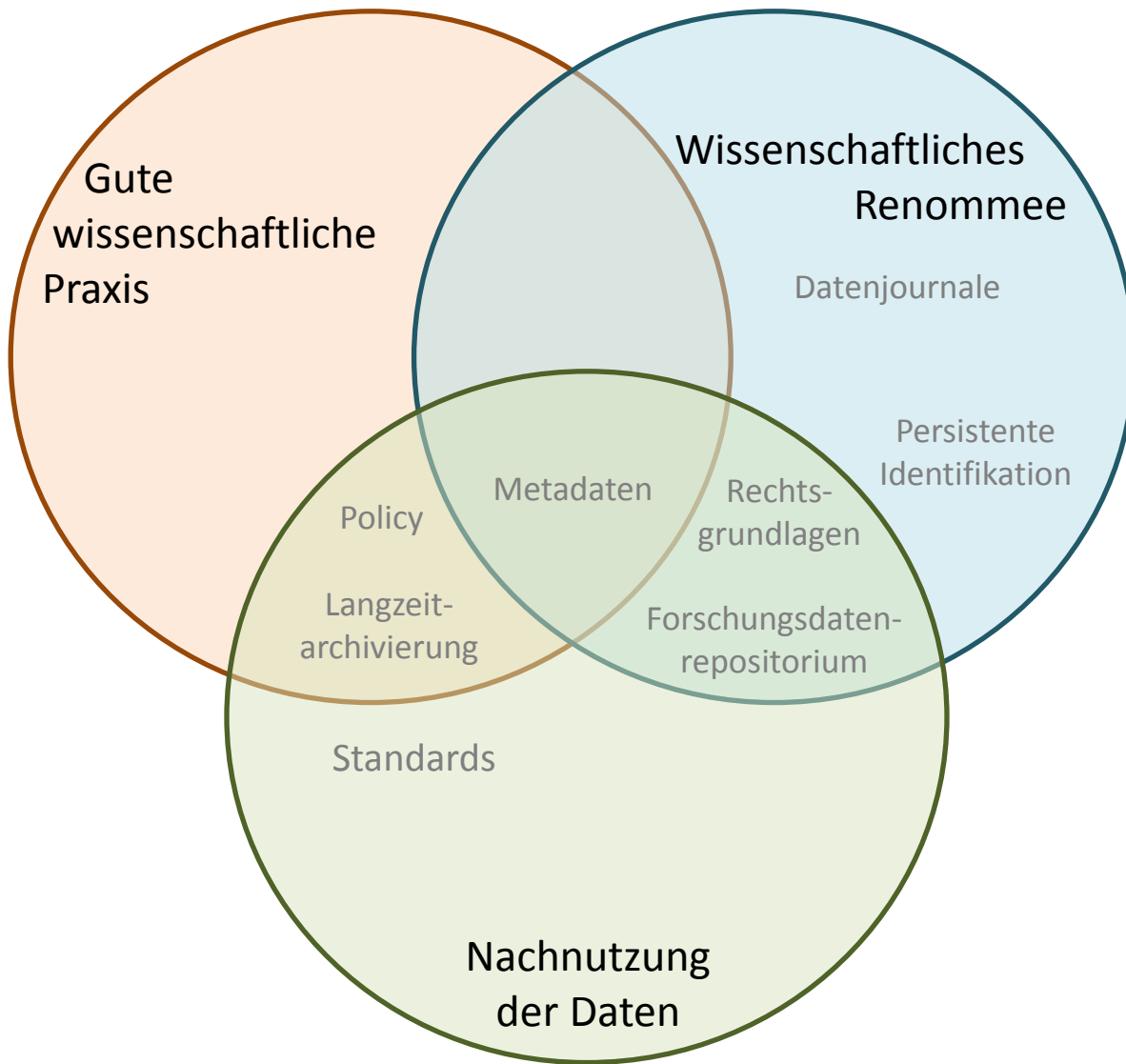
# Rechtsgrundlagen

- Urheberrecht
- Datenbankrecht
- Archivrecht
- Persönlichkeitsrechte
- Strafrecht
- Kriegswaffenkontrollgesetz
- Jugendschutzgesetz
- Kooperationsverbot
  
- Lizenzen für Forschungsdaten
- Einwilligungserklärungen

# Wofür Forschungsdatenmanagement?

- Gute wissenschaftliche Praxis
- Wissenschaftliches Renommee
- Nachnutzung der Daten





# Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

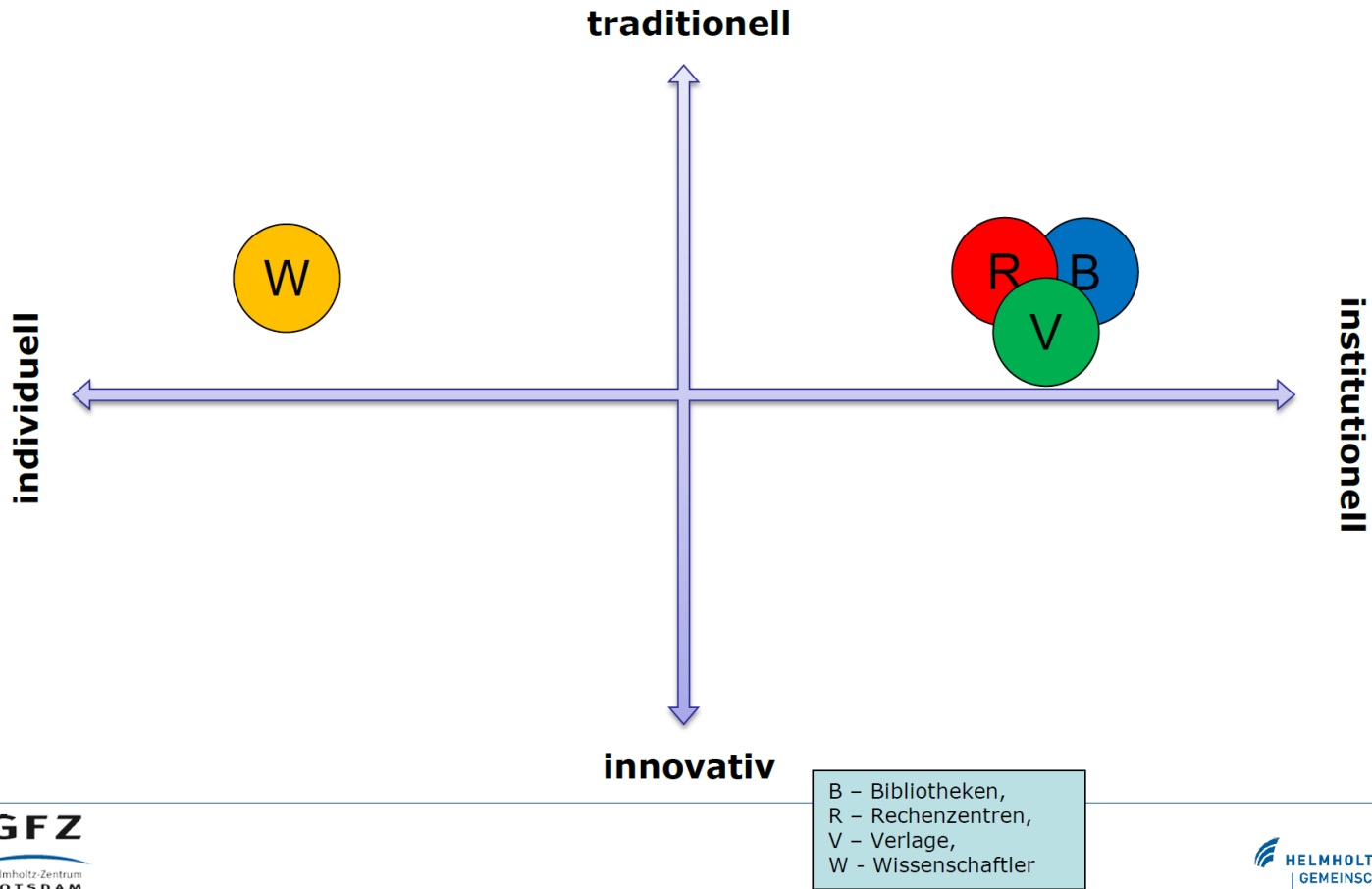
Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

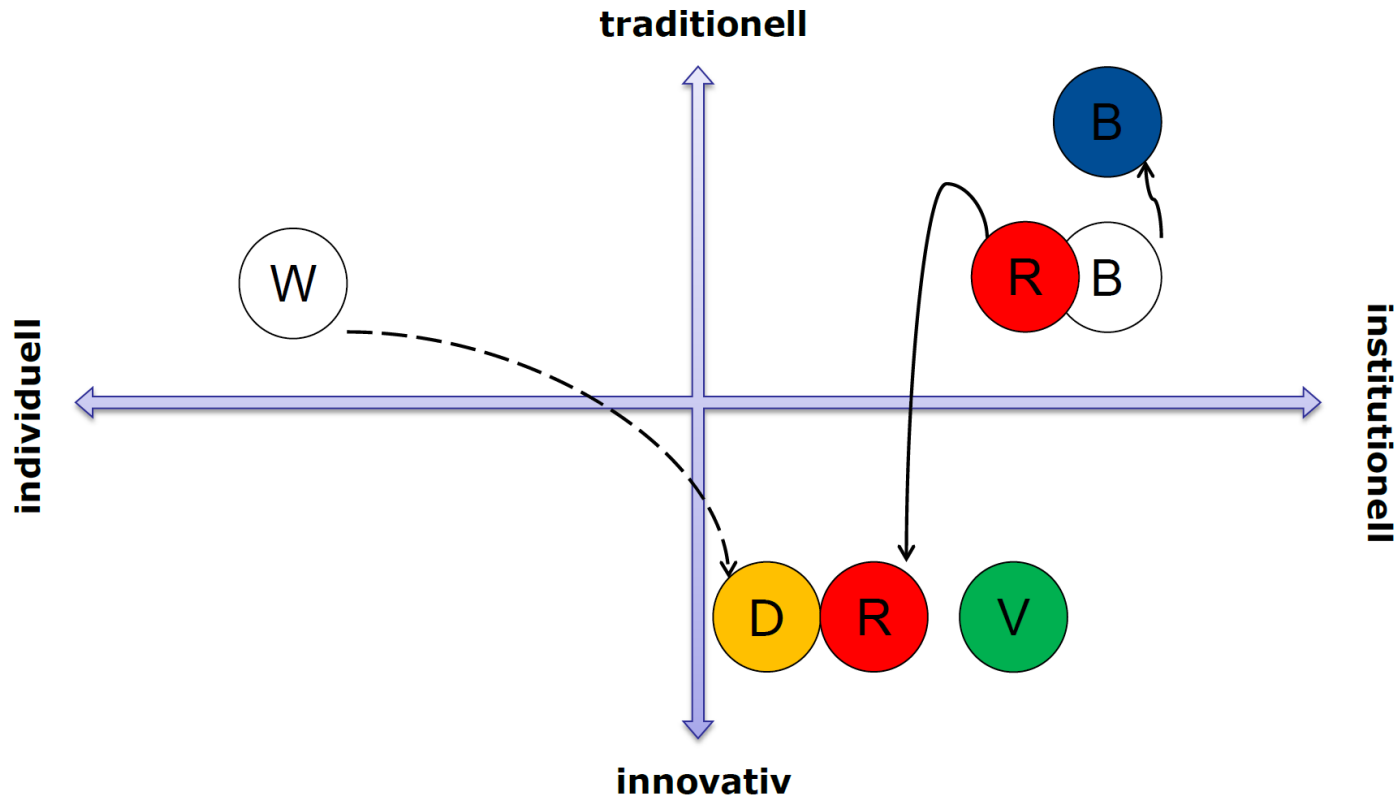
# Warum?

- Neue Technologien
- Neue Anforderungen
- Neue kommerzielle Dienste
  
- → Krise der Bibliotheken?

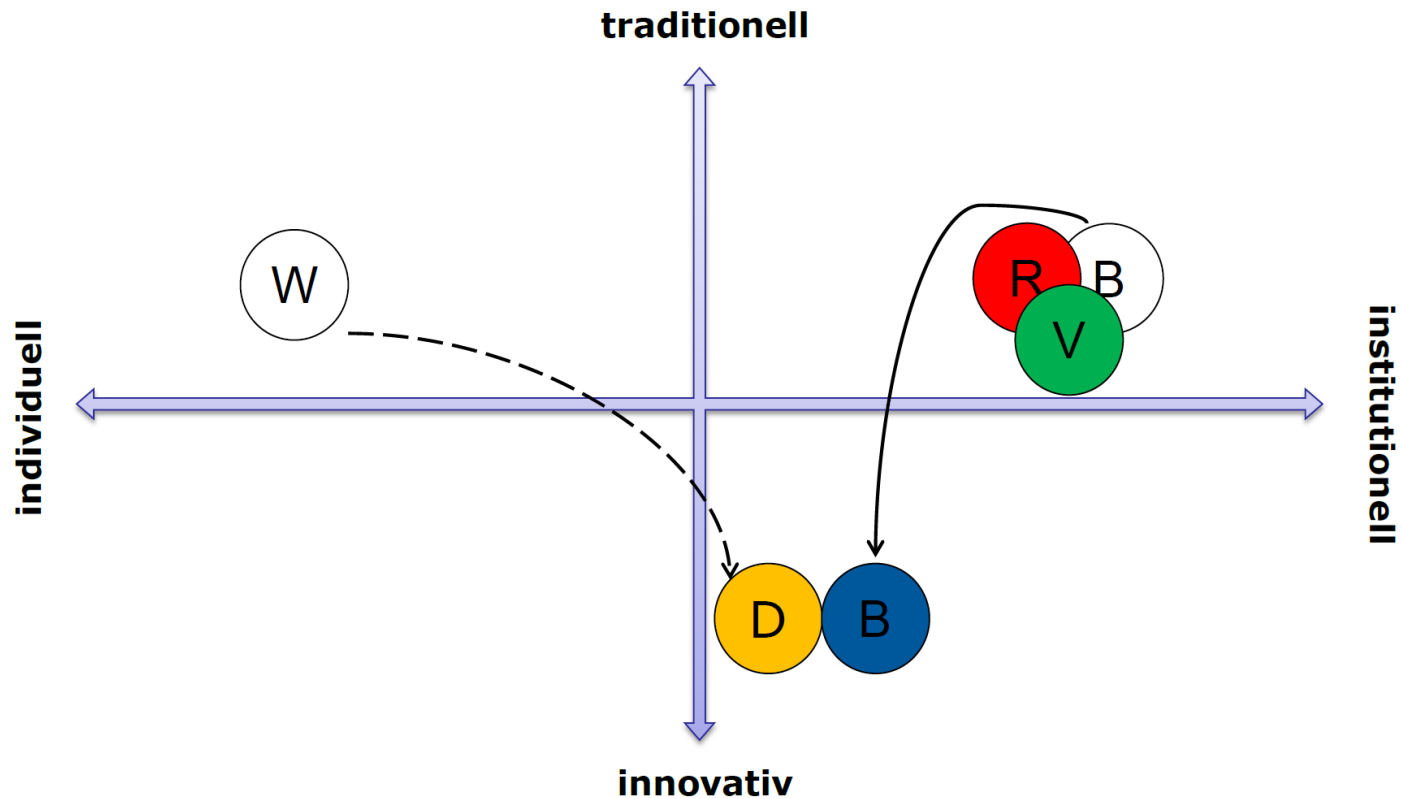
# Szenario ‚Forschungsdaten‘ heute



2025: Datascientists sind bei Rechenzentren/Datenzentren angesiedelt, generische Dienste werden kommerziell angeboten. Bibliotheken fallen zurück auf traditionelle Rollen.



2025: Datascientists sind bei Bibliotheken angesiedelt,  
Bibliotheken sind Träger der Innovation.



# LIBER

- [Ligue des bibliothèques européennes de recherche](#)
- 400 Forschungsbibliotheken
- 48 deutsche Bibliotheken
- München: UB TU, BSB, VDB, UB UniBwM

# Ten recommendations for libraries to get started with research data management

[Final report](#) of the LIBER working group on E-  
Science / Research Data Management, 2012



Offer research data management support, including data management plans for grant applications, intellectual property rights advice and information materials. Assist faculty with data management plans and the **integration of data management into the curriculum.**



# Datenmanagementpläne I

- Welche Daten werden erhoben/verwendet?
- Wie werden die Daten verarbeitet?
- Welche Standards werden benutzt?
- Wie werden die Daten dokumentiert?
- Wo werden die Daten wann (offen) abgelegt?
- Was spricht ggf. gegen eine Veröffentlichung?
- Welche Aufwände entstehen durch das Forschungsdatenmanagement?

# Datenmanagementpläne II

- [DFG](#)
- [EU](#)
- [Data Curation Profiles Directory](#) (Purdue University)

# Lizenzen für Forschungsdaten

- Siehe Teil 2

# Kurse für Forschungsdatenmanagement

- [Uni Bielefeld](#)
- FHVR
- ~~[LMU München](#)~~

Engage in the development of metadata and data standards and provide **metadata services** for research data.

- Beispiel [RDA/FRBR](#)
- Beispiel [DDI](#)
- Metadaten-Services:
  - Transformationen
  - Kontrollierte Vokabulare



Create Data Librarian posts and develop professional staff skills for **data librarianship**.

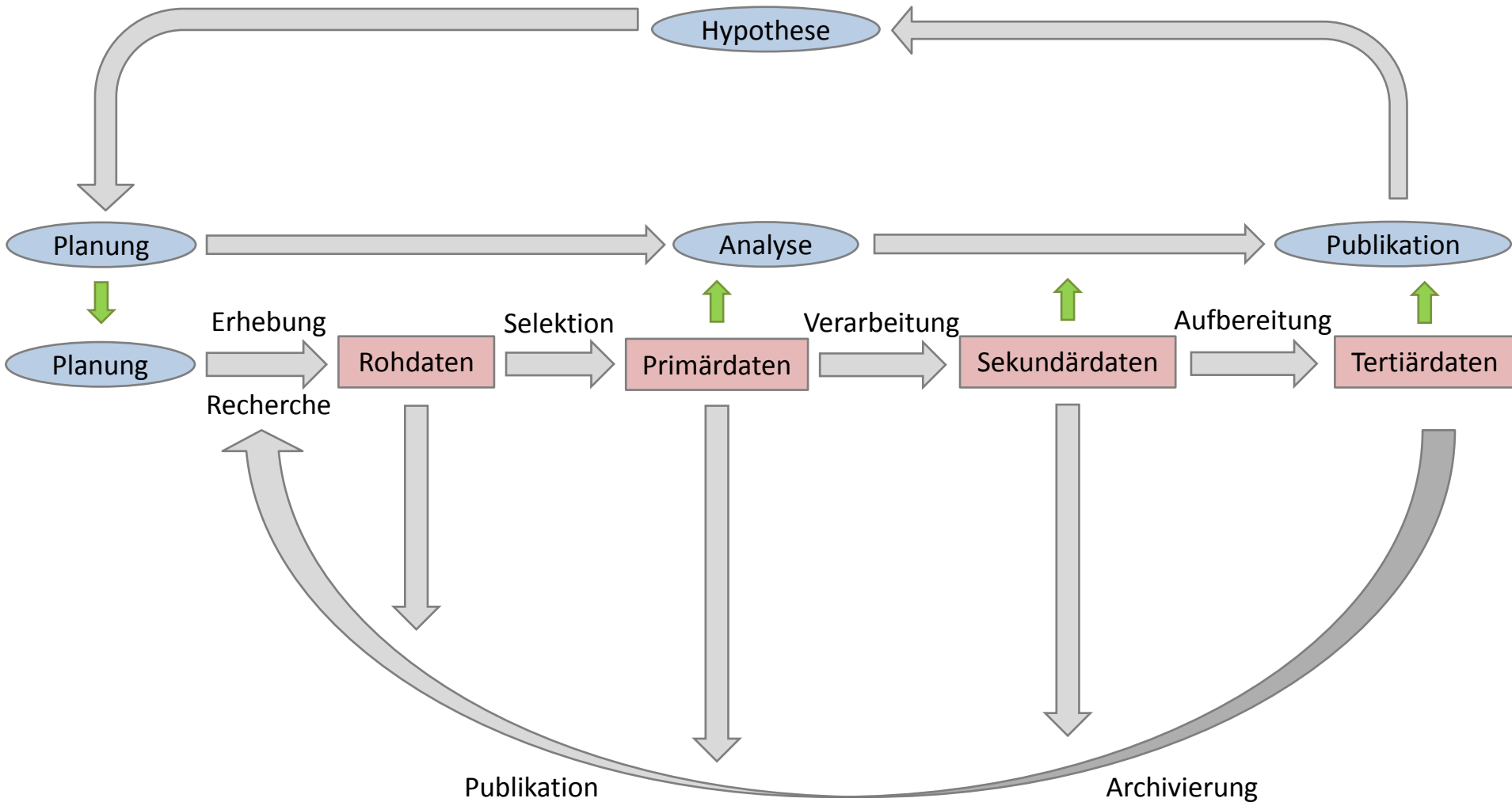
3

Actively participate in institutional research data policy development, including resource plans. Encourage and adopt open data policies where appropriate in the **research data life cycle**.





# Forschungsdatenzyklus



# Forschungsdaten-Policies

- [Universität Bielefeld](#) (2012)
- [Universität Edinburgh](#) (2011)
- [Universität Southampton](#) (2012)
- [KIT](#) (2010)

# (Offene) Fragen

- Who owns the data?
- What Requirements are Imposed By Others?
- Which Data Should Be Retained?
- For How Long Should Data Be Maintained?
- How Should Digital Data Be Preserved?
- Are there Ethical Considerations?
- How are Data Accessed?
- How Open Should the Data Be?
- How Will Costs Be Managed?
- What are the Alternatives to Local Data Management?

Liaise and partner with researchers, research groups, data archives and data centers to **foster an interoperable infrastructure for data access, discovery and data sharing.**

5

Support the lifecycle for research data by  
**providing services for storage, discovery and  
permanent access.**

6

# Promote research data citation by **applying persistent identifiers to research data.**

- URNs → [DNB](#) – Deutsche Nationalbibliothek
- Handles → [GWDG](#) – Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen
- DOIs
  - [TIB](#) – Technische Informationsbibliothek Hannover
  - [DaRa](#) – Registrierungsagentur für Sozial- und Wirtschaftsdaten



**Provide an institutional Data Catalogue or Data Repository, depending on available infrastructure.**

8

Get involved in **subject specific data management** practice.

9



Offer or mediate secure storage for dynamic and static research data in **co-operation with institutional IT units** and/or seek exploitation of appropriate cloud services.

10

# **Klaus Tochtermann: 10 Thesen zum zukünftigen Profil von wissenschaftlichen Informations- Infrastruktureinrichtungen mit überregionaler Bedeutung**

- [ZBW, 2013](#)

# Struktur der Veranstaltung

Teil 1: Forschungsdaten

Teil 2: Forschungsdateninfrastruktur

Teil 3: Forschungsdatenmanagement in Bibliotheken

Teil 4: Weiterführende Überlegungen

# Langzeitarchivierung

- ≠ Backup
- Lange Zeit? 5 Jahre? 10 Jahre? 300 Jahre?
- Verwendung von Daten- und Metadatenstandards (alternativ Dokumentation durch Schemata/Application Profiles)
- Ablegen der Metadaten in maschinen- und menschenlesbarer Form
- Einfach verständliche Verbindung zwischen Daten und Metadaten
- Versehen der Daten mit prozessualen Metadaten (ggf. Versionierung der Daten)
- Extraktion technischer Metadaten
- Persistente Identifikatoren
- Verfolgen des Vollständigkeitsprinzips
- Formatkontrolle (Standard, Emulation, Migration)

# Die Spieler

# Wissenschaftler

- Erheben Daten
- Verwenden Daten
- Nutzen Daten nach
- Beschreiben Daten?
- Archivieren Daten?

# Spieltheorie

## Exkurs: Gefangenendilemma

		<b>B</b>	
		C	D
<b>A</b>	C	4, 4	1, 5
	D	5, 1	2, 2

# Wissenschaftler vs. Community

		<b>C</b>	
		O	N
<b>W</b>	O	4, 4	1, 3
	N	5, 2	2, 2



# Möglichkeit 1: Sanktionen

		<b>C</b>	
		O	N
<b>W</b>	O	4, 4	1, 3
	N	3, 2	2, 2

# Möglichkeit 2: Reputation

		<b>C</b>	
		O	N
<b>W</b>	O	6, 4	1, 3
	N	5, 2	2, 2

# Möglichkeit 3: Reputation/Sanktionen

		<b>C</b>	
		O	N
<b>W</b>	O	6, 4	1, 3
	N	3, 2	2, 2

# Möglichkeit 4: Commitment

		<b>C</b>	
		O	N
<b>W</b>	O	6, 4	1, 3
	N	3, 2	2, 2

# Universitäten

- Achten auf wissenschaftliche Standards
- Stellen Policies auf
- Erstellen institutionelle Repositorien
- Setzen neue Studiengänge auf  
(Datenbibliothekar, Datenwissenschaftler)

# Wissenschaftsorganisationen

- Achten auf wissenschaftliche Standards
- Stellen Policies auf
- Planen und fördern langfristige Strategien
- Stellen allgemeine Repositorien bereit?

# Förderorganisationen

- Geben Geld für strategische Projekte
- Verlangen Veröffentlichung der Daten

# Regierungen

- Bauen administrative Hürden ab?
- Schaffen Rechtssicherheit?



# Verlage

- Verlangen Veröffentlichung der Daten
  - [PLOS](#)
- Veröffentlichen Daten (Data Journals)

# Rechenzentren

- Stellen  
Speicherplatz/Netzwerk/Rechnerleistung  
bereit
- Bitstream Preservation

# Archive/Museen/Sammlungen

- Langzeitperspektive
- Kulturelles Erbe
- [DDB](#) – Deutsche Digitale Bibliothek
- [Europeana](#)

# Bibliotheken

- Öffentlichkeitsperspektive

# Gremien und Allianzen

- [Wissenschaftsrat](#)
- [GWK](#) - Gemeinsame Wissenschaftskonferenz
- RII – Rat für Informationsinfrastrukturen
  
- [Allianz](#) der deutschen Wissenschaftsorganisationen
- [LIBER](#) – Ligue des bibliothèques européennes de recherche
- [LERU](#) – League of European Research Universities
- [Science Europe](#)
- [RDA](#) – Research Data Alliance

# Open Science

- den öffentlichen Zugang zu Forschungsergebnissen in Form von Publikationen (**Open Access**)
- die freie Verfügbarkeit von Forschungsdaten/Rohdaten (**Open Research Data**)
- die Zugänglichkeit und Vernetzung von öffentlichen Datenbeständen (**Linked Open Data**)
- die Suche nach Alternativen zu den traditionellen Begutachtungsverfahren der Zeitschriften-Verlage, um eine höhere Transparenz in diesem Bereich zu gewährleisten (**Open Review**)
- die Entwicklung und den Einsatz quelloffener Software für die Wissenschaft (**Open Source**)
- die Einbeziehung nicht-professioneller Wissenschaftler in den wissenschaftlichen Prozess (**Citizen Science**)

# Science 2.0

- Wikis
- Blogs
- Soziale Medien

# Methoden/Werkzeuge

