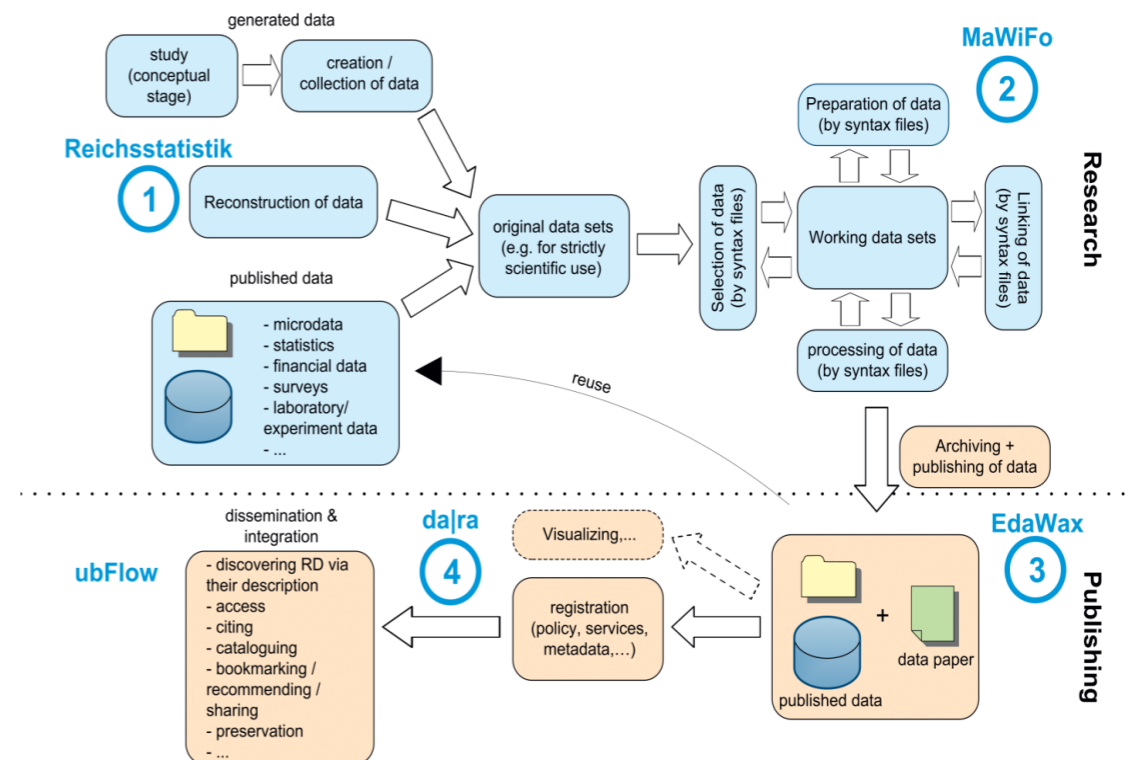


# Zitieren – Recherchieren – Reproduzieren

## Forschungsdaten in den Wirtschaftswissenschaften



*Kirsten Jeude*

Metadaten-Management

DINI AG KIM Workshop Forschungsdaten

Universität Mannheim

25.03.2013

# Zitieren – Recherchieren - Reproduzieren

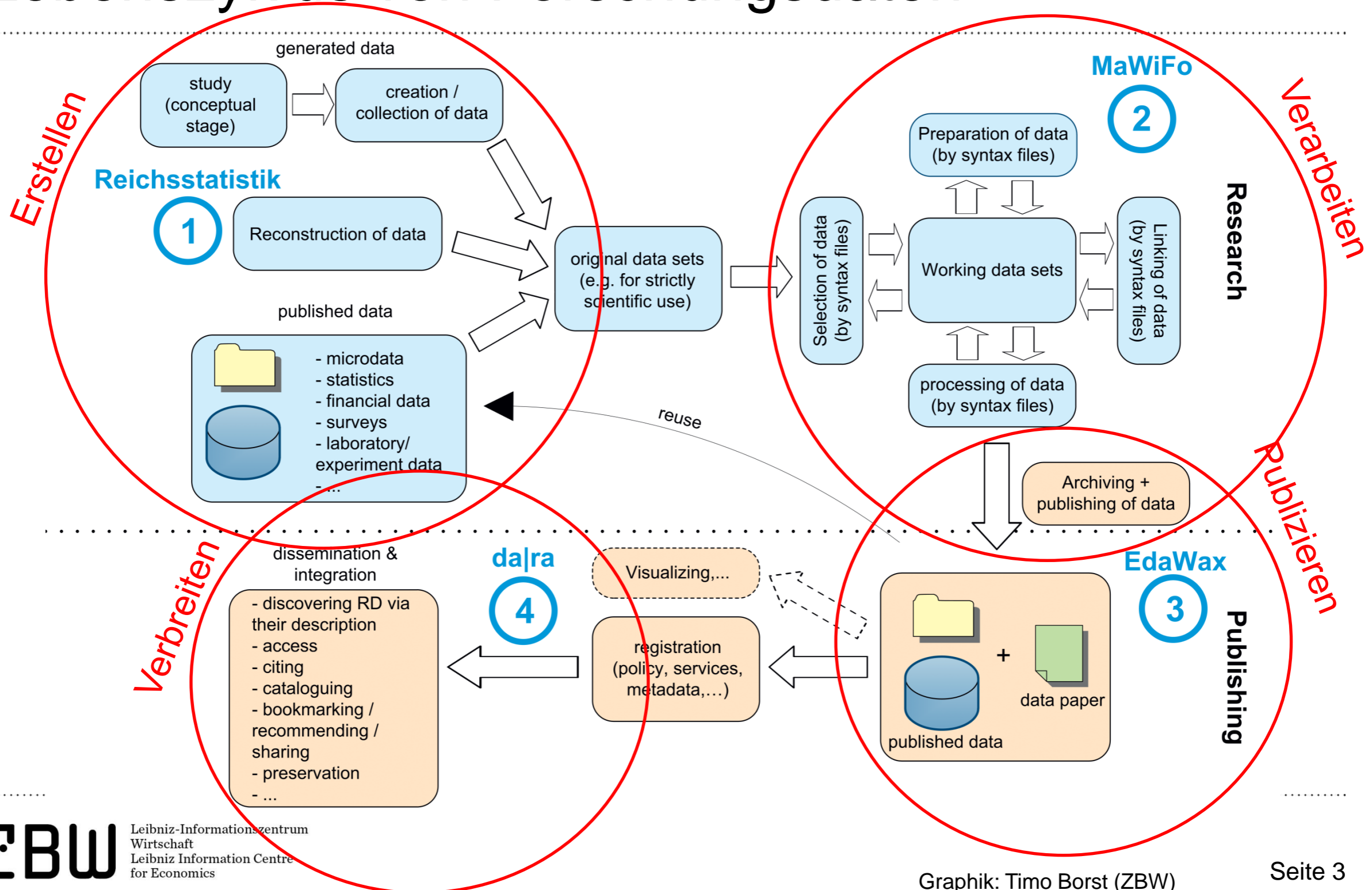
---

## Gliederung

---

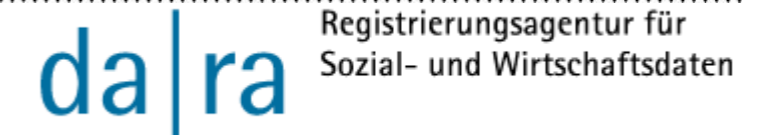
1. Kurzvorstellung der Forschungsdatenprojekte an der ZBW
2. Situation in den Wirtschaftswissenschaften: Ergebnisse aus den Projekten
3. ZBW-Arbeitsmodell: Metadaten für Forschungsdaten
4. Forschungsdaten im Semantic Web
5. Fazit

# Lebenszyklus von Forschungsdaten



# da|ra – DOI-Registrierungsagentur für sozial- und wirtschaftswissenschaftliche Forschungsdaten

## Forschungsdaten zitieren



**Ziel:** dauerhafte Identifizierung, Sicherung, Lokalisierung und verlässliche Zitierbarkeit von Forschungsdaten

- 07/2010: Service online (GESIS), ab 10/2010 in Kooperation GESIS / ZBW
- da|ra-Metadatenchema: 9 Pflichtfelder und 25 optionale Felder, jeweils mit Unterfeldern (ca. 120 Informationseinheiten)
- Lieferung der Werte von 5 Pflichtfeldern und weiteren optionalen Feldern an DataCite
- Mappings zu DataCite, Dublin Core und DDI

Ansprechpartner: [r.toepfer@zbw.eu](mailto:r.toepfer@zbw.eu)

### Gesamtansicht

Der Außenhandel Deutschlands. Basisdaten für den Zeitraum 1830 bis 2000 gesis

DOI	<a href="https://doi.org/10.4232/1.8358">10.4232/1.8358</a>
ID Studie	ZA8358
Version	1.0.0
Primärforscher	• Sensch, Jürgen
Referenzzeitraum	• 1830 - 2000
Veröffentlichungsdatum	2009
Sprache	Deutsch
Klassifikation	• ZA: <ul style="list-style-type: none"><li>◦ Daten Historischer Studien</li></ul>
Beschreibung	• <b>Zusammenfassung</b> Die vorliegende Datensammlung fasst die wichtigsten Zeitreihen zum Außenhandel aus verschiedenen Primär- und Sekundärquellen ... <a href="#">zeige mehr</a>
Geographischer Raum	• Historische deutsche Gebiete (-1871) (DQDE) Deutscher Zollverein • Deutsches Reich (1871-1945) (DXDE) • Deutschland (DE) • Deutsches Zoll ... <a href="#">zeige mehr</a>
Erhebungsverfahren	Quellen: Daten aus wissenschaftlichen Einzelpublikationen. Veröffentlichungen der amtlichen Statistik des Deutschen Reichs (Statistisch ... <a href="#">zeige mehr</a>
Datensatz	• Anzahl der Einheiten: 170 Anzahl der Variablen: 667
Verfügbarkeit	● Download A - Daten und Dokumente sind für die akademische Forschung und Lehre freigegeben.

Aktualisierung Metadaten 2012-11-21

**Ziel:** publikationsbezogenes Datenarchiv für Fachzeitschriften zu entwickeln

- Systematische Aufbereitung der Herausforderungen, z.B. fehlende Anreizstrukturen
  - Analyse von technischen und organisatorischen Lösungsansätzen
  - Entwicklung eines Metadatenschemas
  - Entwicklung einer Pilotanwendung für innovative publikationsbezogene Datenarchive am Beispiel der Fachzeitschrift „Schmollers Jahrbuch / Journal of Applied Social Science Studies“
  - Kooperationsprojekt des RatSWD, des Instituts INNO-tec der LMU München und der ZBW
  - Blog: [www.edawax.de](http://www.edawax.de); Ansprechpartner: [s.vlaeminck@zbw.eu](mailto:s.vlaeminck@zbw.eu)
-

# MaWiFo - Management wirtschaftswissenschaftlicher Forschungsdaten

Technische Aspekte des Archivierens und Publizierens

MaWiFo

**Ziel:** Untersuchung der IT-technischen Gestaltungsmöglichkeiten für Infrastrukturlösungen zum Management wirtschaftswissenschaftlicher Forschungsdaten

- Forschungsprojekt der ZBW und des Instituts für Informatik der Christian-Albrechts Universität Kiel
- Analyse des Anforderungsbedarfs und der bestehenden Informationssystem
- Forschungsthemen:
  - *Semantikerhaltende Transformation von UML-, OWL- und XML-Modellen auf der Metamodell-Ebene*
  - *Modellgetriebene Generierung von ausführbaren wissenschaftlichen Workflows (Scientific Workflows) zur automatisierten Verarbeitung von Forschungsdaten*
  - *Contributions to Replicability in Computational Research*

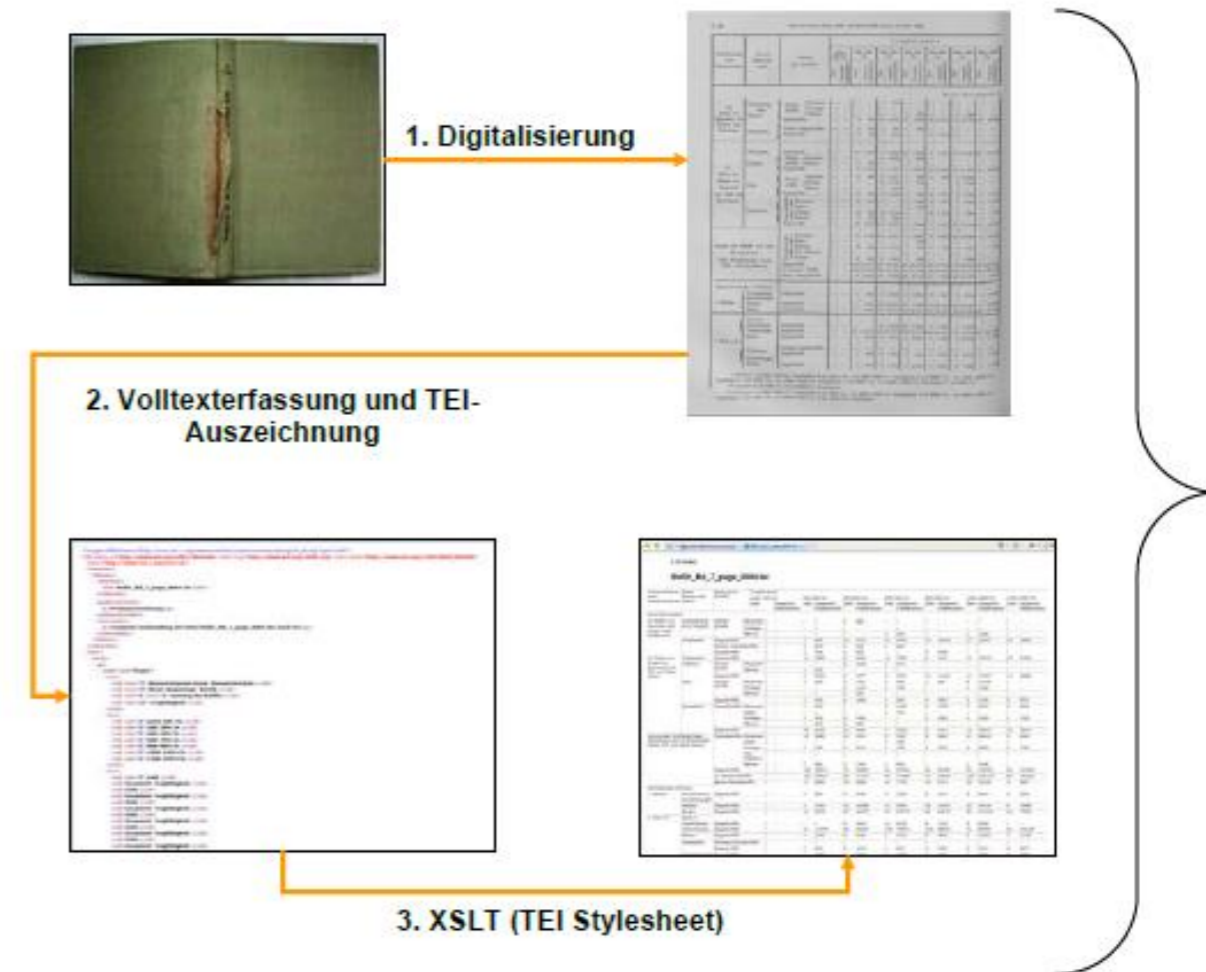
# Digitale Reichsstatistik

## Rekonstruktion von Forschungsdaten

**Ziel:** Digitale Aufbereitung der Statistik des Deutschen Reichs - Alte Folge - (1873 – 1883)

- bislang nur in gedruckter Form vorliegendes statistisches Material nach den Prinzipien des Open Access veröffentlichen
- Statistiken als PDF, als HTML-Dokument und in einer für Tabellenkalkulationsprogramme verarbeitbaren Form

Ansprechpartner: [r.toepfer@zbw.eu](mailto:r.toepfer@zbw.eu)



# Digitale Reichsstatistik

Auswärtiger und überseeischer Warenverkehr des Deutschen Zollgebiets und der Zollausschlüsse sowie Seeschiffsverkehrsverkehr im Jahre 1872 [2] Berlin 1873

aktuelle Seite: [69] - 55

Persistente Identifier (Werk): urn:nbn:de:zbw-dr:sa\_0409

Nummern der Warenverzeichnisse resp. Tarifposition.	Grenzstrecken des Ausgangs.	Menge des Warenansgangs.	Nummern der Warenverzeichnisse resp. Tarifposition.	Grenzstrecken des Ausgangs.	Menge des Warenansgangs.	Nummern der Warenverzeichnisse resp. Tarifposition.	Grenzstrecken des Ausgangs.	Menge des Warenansgangs.
Noch: 406. Terpentinharz.		Ctr. lr.	Noch: 406. Andere Thiere und Mineralien etc.		Ctr. lr.	<b>37. Thiere und thierische Produkte, nicht anderweit genannt.</b>		
Schweiz . . . . .		1	Russland . . . . .		1 413	<b>409. Geflügel und kleines Wildpret aller Art.</b>		
Frankreich . . . . .		130	Oesterreich, Grenze gegen Preussen . . . . .		4 001	Dänemark . . . . .		17
Belgien . . . . .		130	Sachsen . . . . .		38 213	Ostsee . . . . .		5
Niederlande . . . . .		24	Bayern . . . . .		2 799	Russland . . . . .		188
Hamburg . . . . .		97	Württemberg . . . . .		47	Oesterreich, Grenze gegen Preussen . . . . .		23
Zusammen . . . . .		58 460	Schweiz . . . . .		10 348	Sachsen . . . . .		65
<b>404. Andere Harze.</b>		Ctr. lr.	Frankreich . . . . .		263	Bayern . . . . .		400
Dänemark . . . . .		38	Belgien . . . . .		14	Württemberg . . . . .		3
Ostsee . . . . .		1 763	Niederlande . . . . .		1 161	Schweiz . . . . .		2 446
Russland . . . . .		25 175	Bremen . . . . .		29	Frankreich . . . . .		911
Oesterreich, Grenze gegen Preussen . . . . .		14 754	Hamburg . . . . .		1 183	Belgien . . . . .		443
Sachsen . . . . .		20 419	Preuss. Zollausschl. . . . .		37	Niederlande . . . . .		24 109
Bayern . . . . .		3 100	Zusammen . . . . .		61 695	Nordsee . . . . .		245
Württemberg . . . . .		242				Bremen . . . . .		154
Schweiz . . . . .		5 791				Hamburg . . . . .		3 670
Frankreich . . . . .		283				Preuss. Zollausschl. . . . .		987
Belgien . . . . .		14				Zusammen . . . . .		33 666
Niederlande . . . . .		1 236	<b>407. Harzöl: Terpenin, Terpeninöl.</b>		Ctr. lr.	<b>410. Fische, frische und Flusskrebse.</b>		
Bremen . . . . .		1 442	Dänemark . . . . .		137	Dänemark . . . . .		437
Hamburg . . . . .		1 708	Ostsee . . . . .		3 736	Ostsee . . . . .		76
Zusammen . . . . .		75 965	Russland . . . . .		84	Russland . . . . .		53 603
<b>405. Petroleum, Petroleumnaphta.</b>		Ctr. lr.	Oesterreich, Grenze gegen Preussen . . . . .		7 523	Oesterreich, Grenze gegen Preussen . . . . .		1 135
Dänemark . . . . .		4 079	Sachsen . . . . .		5 243	Sachsen . . . . .		72
Ostsee . . . . .		13 954	Bayern . . . . .		2 999	Bayern . . . . .		260
Russland . . . . .		254 206	Württemberg . . . . .		32	Schweiz . . . . .		340
Oesterreich, Grenze gegen Preussen . . . . .		87 054	Belgien . . . . .		22	Frankreich . . . . .		503
Sachsen . . . . .		494 049	Niederlande . . . . .		127	Belgien . . . . .		5 821
Bayern . . . . .		110 341	Nordsee . . . . .		11	Niederlande . . . . .		253
Württemberg . . . . .		3 850	Bremen . . . . .		708	Nordsee . . . . .		55
Schweiz . . . . .		137 611	Hamburg . . . . .		1 761			
			Zusammen . . . . .		23 852			

Die Statistik des Deutschen Reichs

Auswärtiger und überseeischer Warenverkehr des Deutschen Zollgebiets und der Zollausschlüsse sowie Seeschiffsverkehrsverkehr im Jahre 1872 [2]

Titelblatt

Inhalts-Verzeichniss

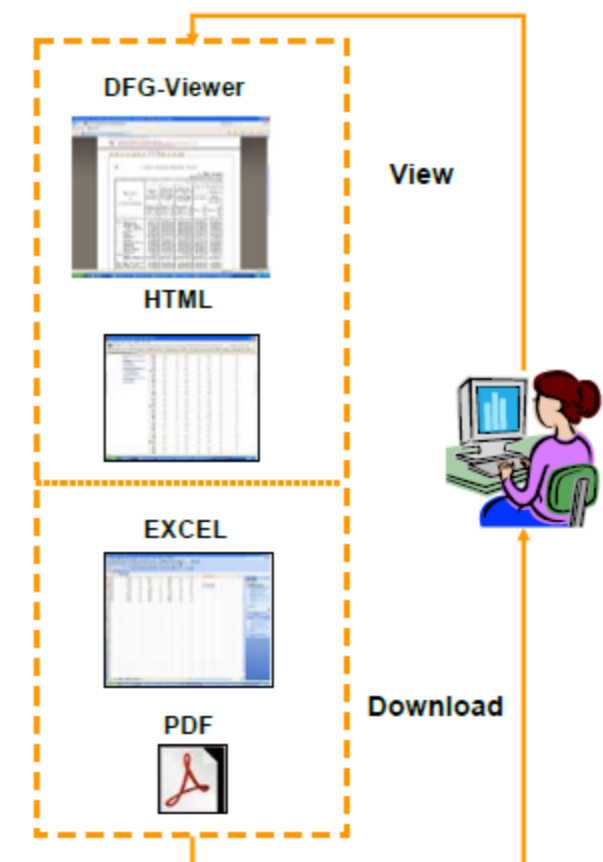
Berichtigungen

Vorbemerkungen

1. Abtheilung. Waaren-Aus- und Durchfuhr des Deutschen Zollgebiets sowie vergleichende Uebersichten der Menge und des Werths der Ein-, Aus- und Durchfuhr und der berechneten Zollerträge im Jahre 1872

I. Haupt-Uebersicht der Waaren-Ausfuhr aus dem freien Verkehr für das Jahr 1872

II. Uebersicht der Waaren-Ausfuhr aus dem freien Verkehr mit Unterscheidung der Grenzstrecken des Ausgangs für das Jahr 1872



- Metadaten: METS/MODS (DFG-Viewer)



# Forschungsdaten der Wirtschaftswissenschaften

Ergebnisse aus den Projekten

EDaWaX

„the current status quo in economics is to not share data or facilitate access to data and codes used for empirical work“



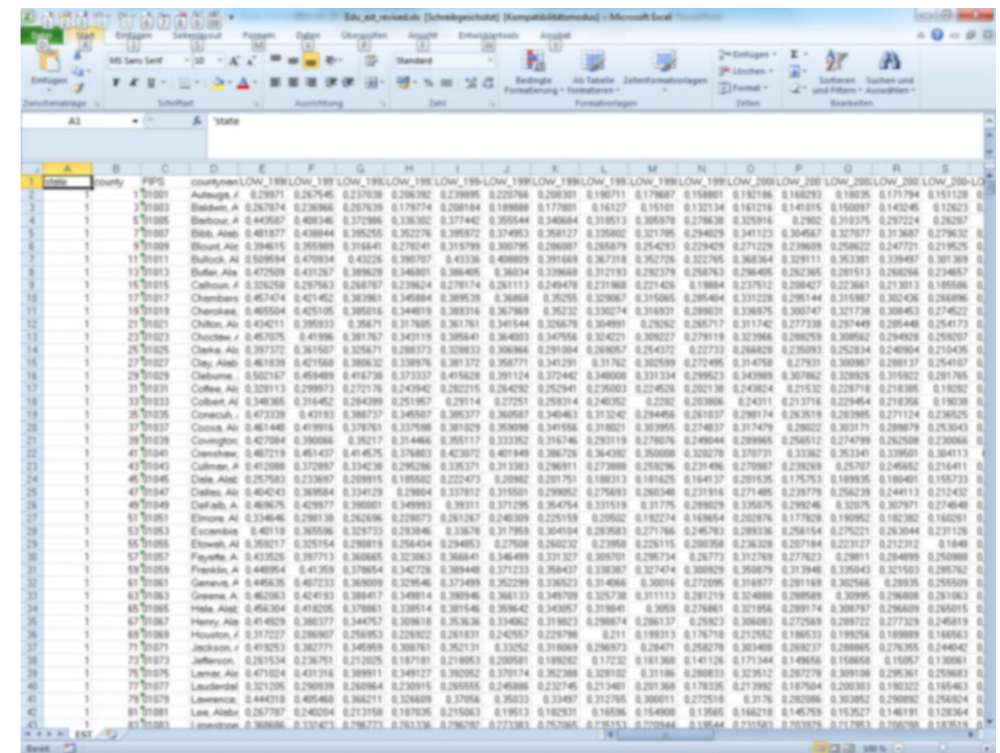
- **89,14 %** der 488 untersuchten Forscher stellen keine Forschungsdaten zur Verfügung
- **8,81%** bieten einen Teil ihrer Daten an
- **2,05%** stellen Ihre Forschungsdaten vollständig zur Verfügung

[http://www.ratswd.de/dl/RatSWD\\_WP\\_215.pdf](http://www.ratswd.de/dl/RatSWD_WP_215.pdf)

# Forschungsdaten der Wirtschaftswissenschaften

## Ergebnisse aus den Projekten: Typen und Eigenschaften der Daten

- Empirische Forschung in den Wirtschaftswissenschaften: Zahlen als Teil von Datenmatrizen und Tabellen
- von einfachen Tabellen bis zu komplizierten Längsschnittdatensätzen
- Daten von nationalen und internationalen statistischen Ämtern
- Rechtliche Aspekte: freie, proprietäre/kommerzielle und vertrauliche Daten
- Konkurrenzsituation in der Forschung
- Keine Publikationskultur mit Metadaten zu Forschungsdaten



The image shows a screenshot of a Microsoft Excel spreadsheet. The spreadsheet contains a large data matrix with columns labeled by country and year (e.g., AUS\_1990, AUS\_1991, etc.) and rows of numerical data. The spreadsheet is titled 'Ea...\_m... (Schreibschutz) [Kompatibilitätsmodus] - Microsoft Excel'. The data is organized in a grid format, with the first column containing country names and the subsequent columns containing numerical values for each year from 1990 to 2010. The spreadsheet is displayed in a standard Excel interface with a menu bar and toolbars.

# Forschungsdaten der Wirtschaftswissenschaften

## MaWiFo



# Forschungsdaten der Wirtschaftswissenschaften

*Befragung von Herausgebern von Fachzeitschriften zu  
Daten-Policies*

EDaWaX

- Eingefordertes Material:
  - 66% Datensätze und andere Materialien;
  - 83% Datensätze,
  - 66% Codes/Skripte;
  - 50% Readme-File;
  - 25% verwendeten Programme.
- Bereitstellung:
  - Webseiten des Herausgebers oder des Verlags
  - Datenrepositorien werden nur in Ausnahmefällen genutzt.
- Metadaten-Anreicherung:
  - 27% reichern an;
  - 46% nicht.
  - Diejenigen, die Metadaten ergänzen, nutzen **keine Metadatenstandards**

# Forschungsdaten der Wirtschaftswissenschaften

## *Analyse der Hostingoptionen von Forschungsdatenzentren EDaWaX*

- Hosting:
  - 74% akzeptieren externe Datensätze (meist nur fachlich passende)
  - 25% Software
  - 69% Code
- Metadatenstandards:
  - 70% DDI (Data Documentation Initiative)
  - 29% Dublin Core
- Persistent Identifier:
  - 56% vergeben PIs bei der Speicherung;
  - 33% nicht
- 6% verwenden und verbreiten RDF-Dateien
- 65% der FDZ bieten Unterstützung bei der Metadaten-Generierung;
  - 35% bieten dafür ein Web-Formular an (z.B. Nesstar)

# Arbeitsmodell Metadaten für Forschungsdaten

Abhängig von Aufgabe / Zielsetzung des Informationssystems

## Level 1: Zitierbarkeit herstellen

- Nur wenige Felder notwendig; diese müssen aber vorhanden sein → Pflichtfeld

### DataCite 2.2: Mandatory Properties

ID	Property
1	Identifier (with type attribute)
2	Creator (with name identifier attributes)
3	Title (with optional type attribute)
4	Publisher
5	PublicationYear

### da|ra 2.2.1. Mandatory Properties

ID	Property
0	Typ der Ressource
1	Titel
4	Primärforscher
5	Publikationsagent (mit optionalen Attributen)
6	Registrierungsagentur (mit Attributen)
7	DOI
8	URL
12	Veröffentlichungsdatum
29	Verfügbarkeit (kontrolliert)

# Arbeitsmodell Metadaten für Forschungsdaten

---

Abhängig von Aufgabe / Zielsetzung des Informationssystems

---

## Level 2: Auffindbarkeit unterstützen

- Je nach Zielgruppe / Zielsetzung ergänzende Felder zur Unterstützung der Recherche (z.B. in Fachportalen, Fachdatenbanken, die der Recherche nach Forschungsdaten dienen)

- Grundgesamtheit
- Auswahlverfahren
- Geographische Abdeckung
- Zeitliche Abdeckung
- Schlagwörter und Klassifikationen
- Sprache
- Version
- Sonstige Beteiligte
- Beschreibung / Abstract
- ....

# Arbeitsmodell Metadaten für Forschungsdaten

Abhängig von Aufgabe / Zielsetzung des Informationssystems

## Level 3: Verlinkbarkeit sicherstellen / Zukunftsfähigkeit

- Einbindung von Normdaten von Personen, Institutionen und Themen ermöglicht eine bestandsübergreifende Verlinkung und dient der Zukunftsfähigkeit der Metadaten im Hinblick auf das Semantic Web

Primärforscher		Beispiel
Person	Name des Primärforschers	z.B. Schwarz, Peter
ID Person	Eindeutiger Identifier der Person. Kann von da ra ergänzt werden, wenn nicht vorhanden.	z.B. GND-ID: 123637449 z.B. ORCID ID: G-1442-2009
Vokabular ID Person	Vokabular, aus dem der Identifier stammt	z.B. „ORCID“ z.B. „GND“
URI Normdatensatz	URI des Normdatensatzes	z.B. http://d-nb.info/gnd/123637449



# Arbeitsmodell Metadaten für Forschungsdaten

Abhängig von Aufgabe / Zielsetzung des Informationssystems

## Level 4: Reproduzierbarkeit der Studie durch Metadaten sicherstellen

- Detaillierte Beschreibung des gesamten Entstehungsprozesses der Studie
- Standard zur Abbildung des Lebenszyklus von Forschungsdaten in den Sozialwissenschaften: DDI 3.1

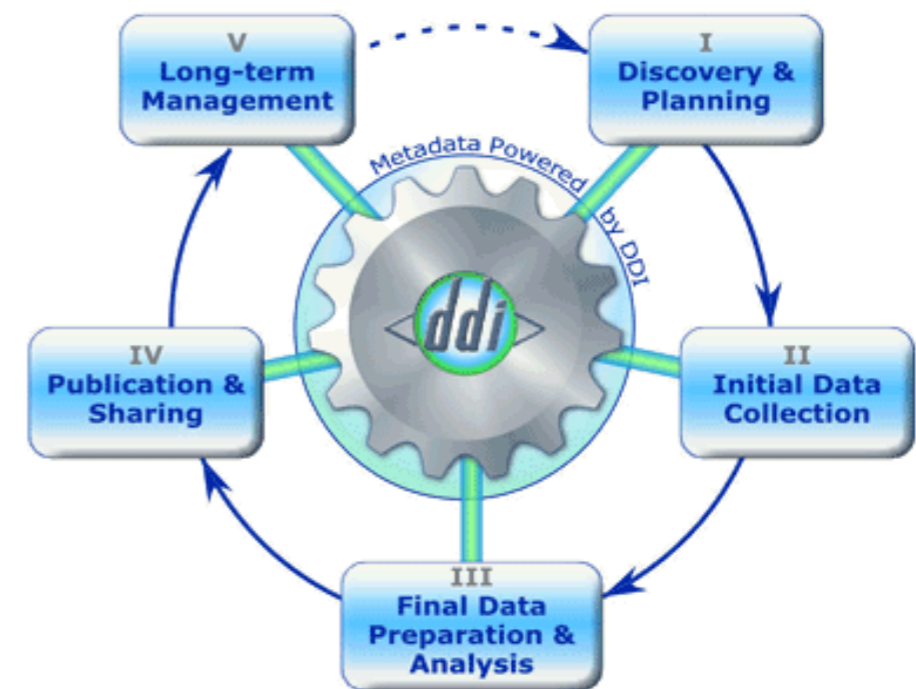
### DDI 3.1 XML Schema Documentation (2009-10-18)

© DDI Alliance 2009 - licensed under the terms of the [GNU Lesser General Public License](#)

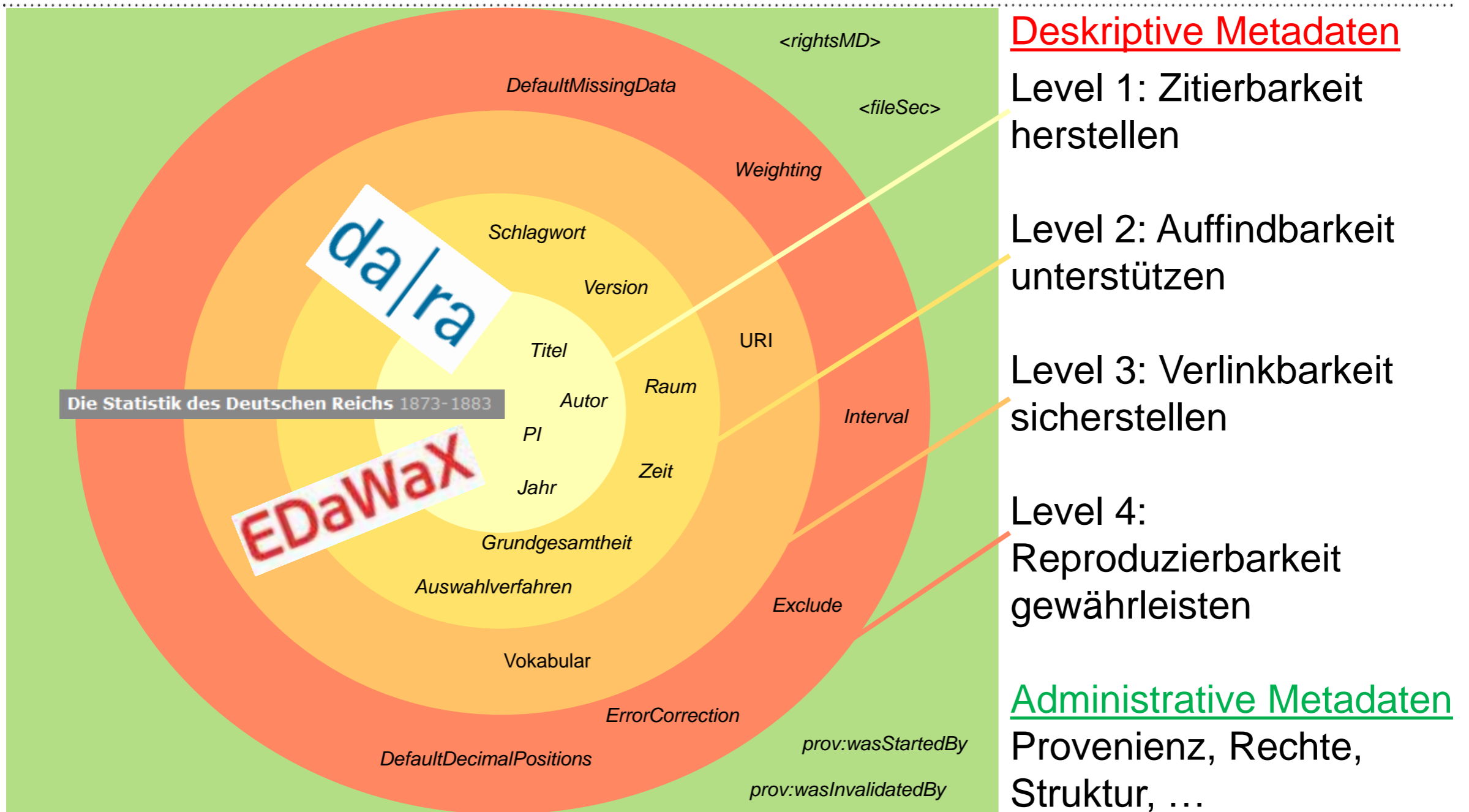
#### [All Components](#)

#### All Elements (846)

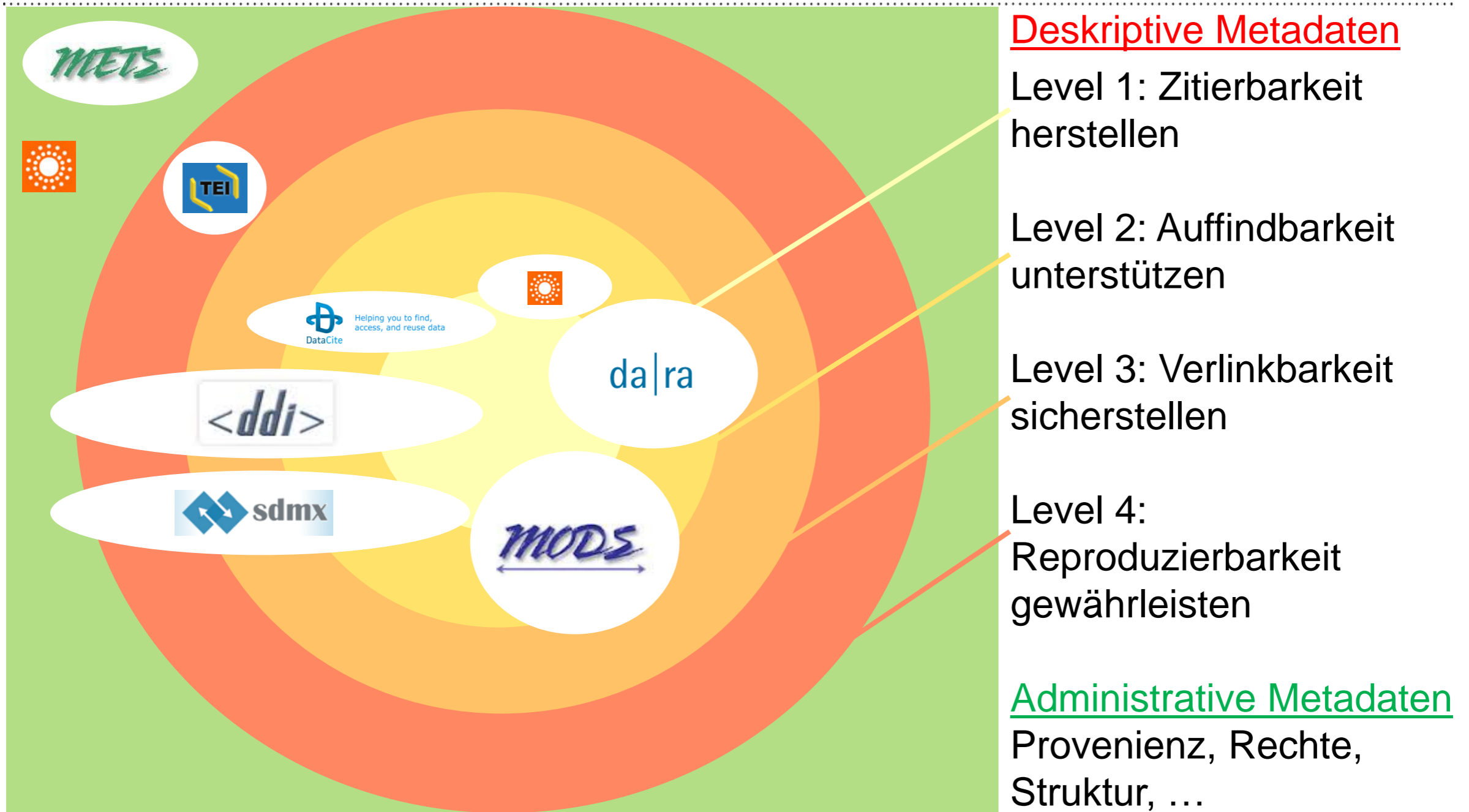
[a](#)  
[abbr](#)  
[Abbreviation](#)  
[Abbreviation](#) (in [SeriesStatement](#))  
[Abstract](#)  
[Abstract](#)  
[Access](#)  
[AccessConditions](#)  
[AccessPermission](#)  
[AccessRestrictionDate](#)  
[AccessTypeName](#)  
[acronym](#)  
[ActionToMinimizeLosses](#)  
[Add](#)  
[Address](#)  
[address](#)  
[AddressMatchType](#) (in [GeographicDomain](#))  
[AdministrativePurpose](#)  
[AgencyOrganizationReference](#)  
[AgencyOrganizationReference](#)



# Arbeitsmodell Metadaten für Forschungsdaten



# Arbeitsmodell Metadaten für Forschungsdaten

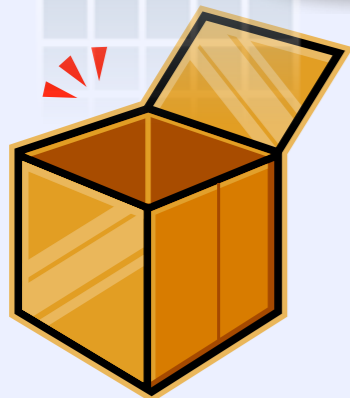


# Forschungsdaten-Metadaten im Semantic Web

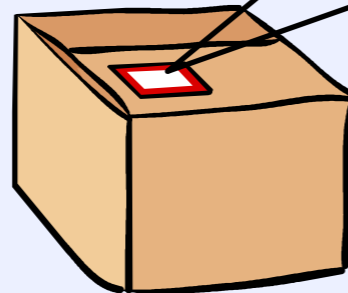
## Klassische Retrieval-Systeme vs. Semantic Web

### Repository

```
62 define lab (ID = !charend('/')  
63 /x = !charend('/')  
64 /iota = !cmdend )  
65  
66 compute my_ID = !ID.  
67 compute my_x = !x.  
68 compute my_iota = !iota.  
69 compute a = 0.  
70 compute b = 0.  
71 exe.
```



Der Außenhandel Deutschlands. Basisdaten für den Zeitraum 1830 bis 2000		gesis
DOI	<a href="https://doi.org/10.4232/1.8358">10.4232/1.8358</a>	
ID Studie	ZA8358	
Version	1.0.0	
Primärforscher	• Sensch, Jürgen	
Referenzzeitraum	• 1830 - 2000	
Veröffentlichungsdatum	2009	
Sprache	Deutsch	
Klassifikation	• ZA: ◦ Daten Historischer Studien	
Beschreibung	• <b>Zusammenfassung</b> Die vorliegende Datensammlung fasst die wichtigsten Zeitreihen zum Außenhandel aus verschiedenen Primär- und Sekundärque ... <a href="#">zeige mehr</a>	
Geographischer Raum	• Historische deutsche Gebiete (-1871) (DQDE) Deutscher Zollverein Deutsches Reich (1871-1945) (DXDE) Deutschland (DE) Deutsches Zoll ... <a href="#">zeige mehr</a>	
Erhebungsverfahren	Quellen: Daten aus wissenschaftlichen Einzelpublikationen. Veröffentlichungen der amtlichen Statistik des Deutschen Reichs (Statistisch ... <a href="#">zeige mehr</a>	



- Für jede Art von Forschungsdaten
- Inkonsistenzen durch Trennung von Daten und Metadaten
- Reproduktion erschwert



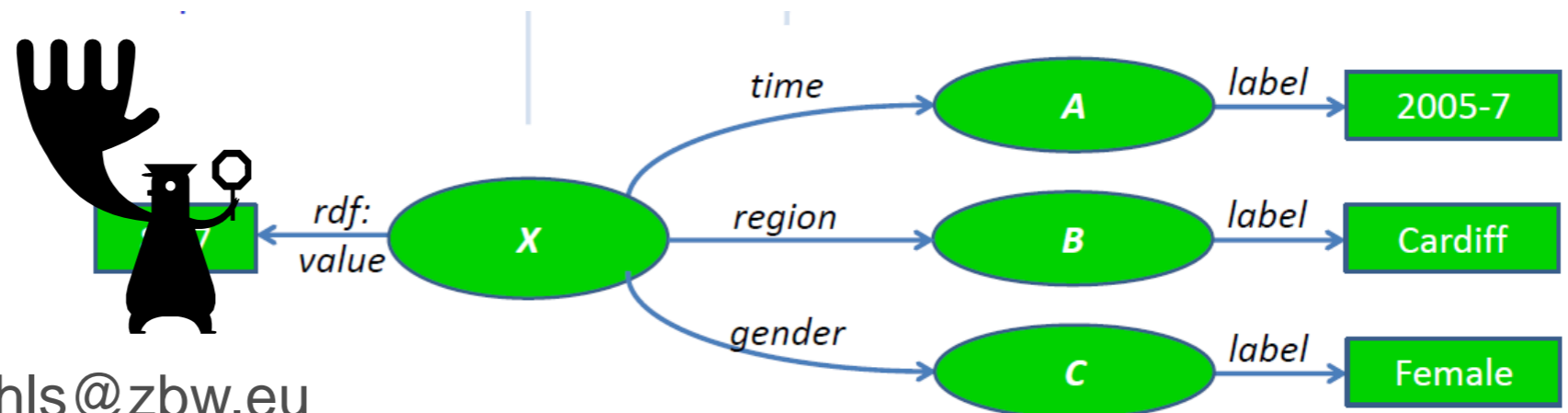
# Forschungsdaten-Metadaten im Semantic Web

Referenz-Modell für Forschungsdaten

MaWiFo

Resource Description Framework (RDF) + RDF Data Cube Vocabulary

- Jedes Datenelement ist referenzierbar
  - Für verteilte Daten aus verschiedenen Quellen
  - Erweiterbar
  - Werte können geschützt werden
- ➔ Executable Paper: Referenzierung + Code im Quelltext der Publikation



Ansprechpartner: [d.bahls@zbw.eu](mailto:d.bahls@zbw.eu)

# Fazit:

## Forschungsdaten in den Wirtschaftswissenschaften

---

- Derzeit wenig Bereitschaft Forschungsdaten zur Verfügung zu stellen
- Keine zentralen (oder gut vernetzte dezentralen) Daten-Repositoryen
- Metadaten:
  - Disziplin-spezifisch, vielfältige Standards, kein Regelwerk
  - Metadatenerfassung durch Forschende: Anreize?
- Workflow:
  - Wie können Publikationen und Forschungsdaten im Prozess der Veröffentlichung verknüpft werden?
- Rolle der Bibliotheken:
  - Mitarbeit bei der Entwicklung von Forschungsumgebungen
  - Entwicklung von Publikationsplattformen
  - Unterstützung / Beratung der Forscher
  - Langzeitarchivierung, Dokumentation und Verbreitung
  - ganz praktisch: z.B. Normierung von Entitäten

# Zitieren – Recherchieren - Reproduzieren

---

Forschungsdaten in den Wirtschaftswissenschaften

---

Vielen Dank für Ihr Aufmerksamkeit!

Fragen ?



[k.jeude@zbw.eu](mailto:k.jeude@zbw.eu)