

DataCite Services als Baustein des Forschungsdatenmanagements in wissenschaftlichen Bibliotheken

Katja Pletsch, ZB MED – Informationszentrum Lebenswissenschaften
Brigitte Hausstein, GESIS – Leibniz Institut für Sozialwissenschaften
Britta Dreyer, Technische Informationsbibliothek (TIB)

Hands-On Lab
107. Deutscher Bibliothekartag
14.06.2018



Making research better by enabling people to find, share, use, and cite data

- Forschungsdatenmanagement und Forschungsdatenpublikation
- DataCite und DOI-System
- Einführung DOI-Registrierung
- DOI Fabrica: DataCite's neue Infrastruktur für die DOI-Registrierung
- Praktische Arbeit mit DOI Fabrica
- Granularität und Versionierung
- Weitere Services und Projekte von DataCite

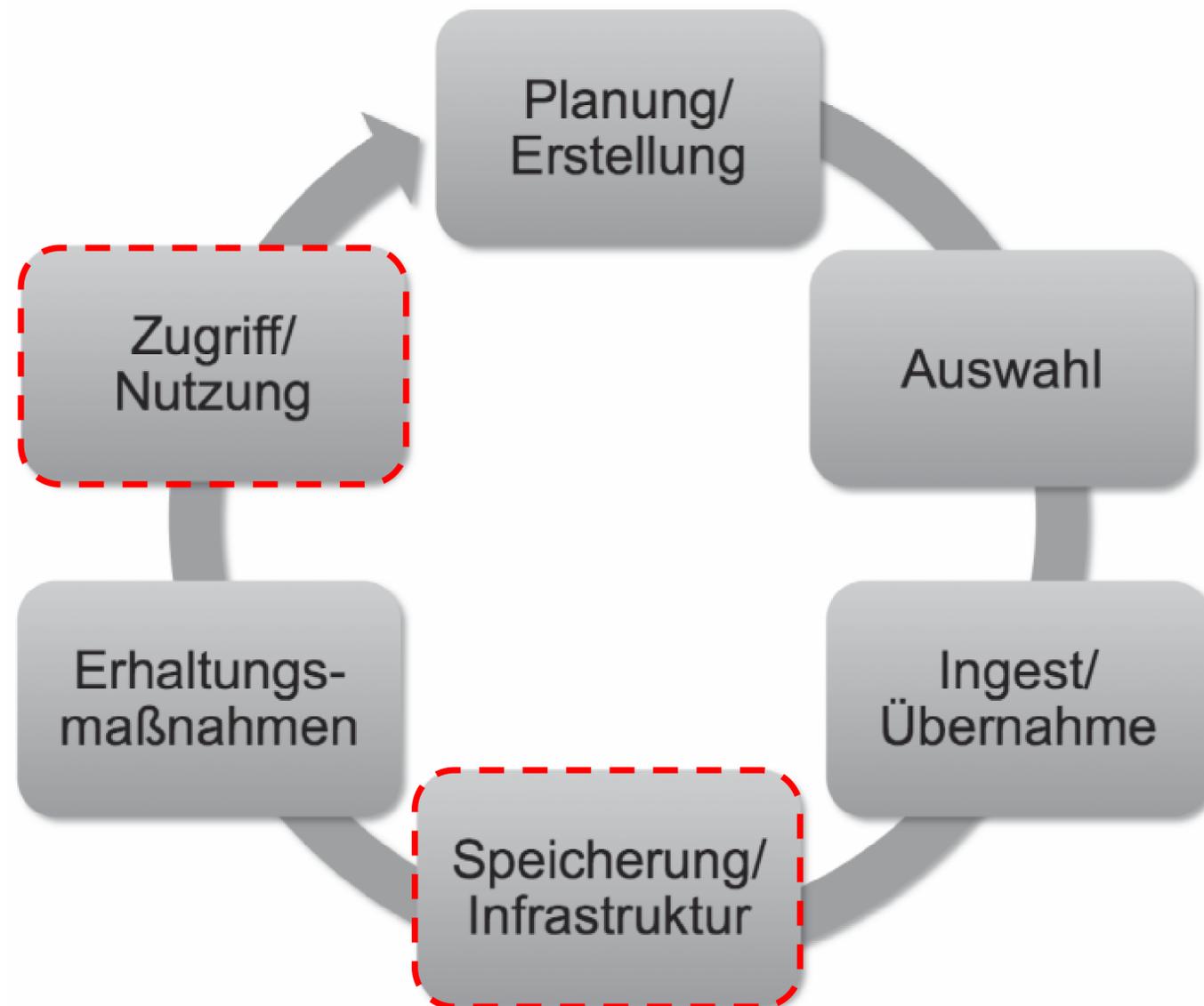
Forschungsdatenmanagement und Forschungsdatenpublikation

Forschungsdatenmanagement



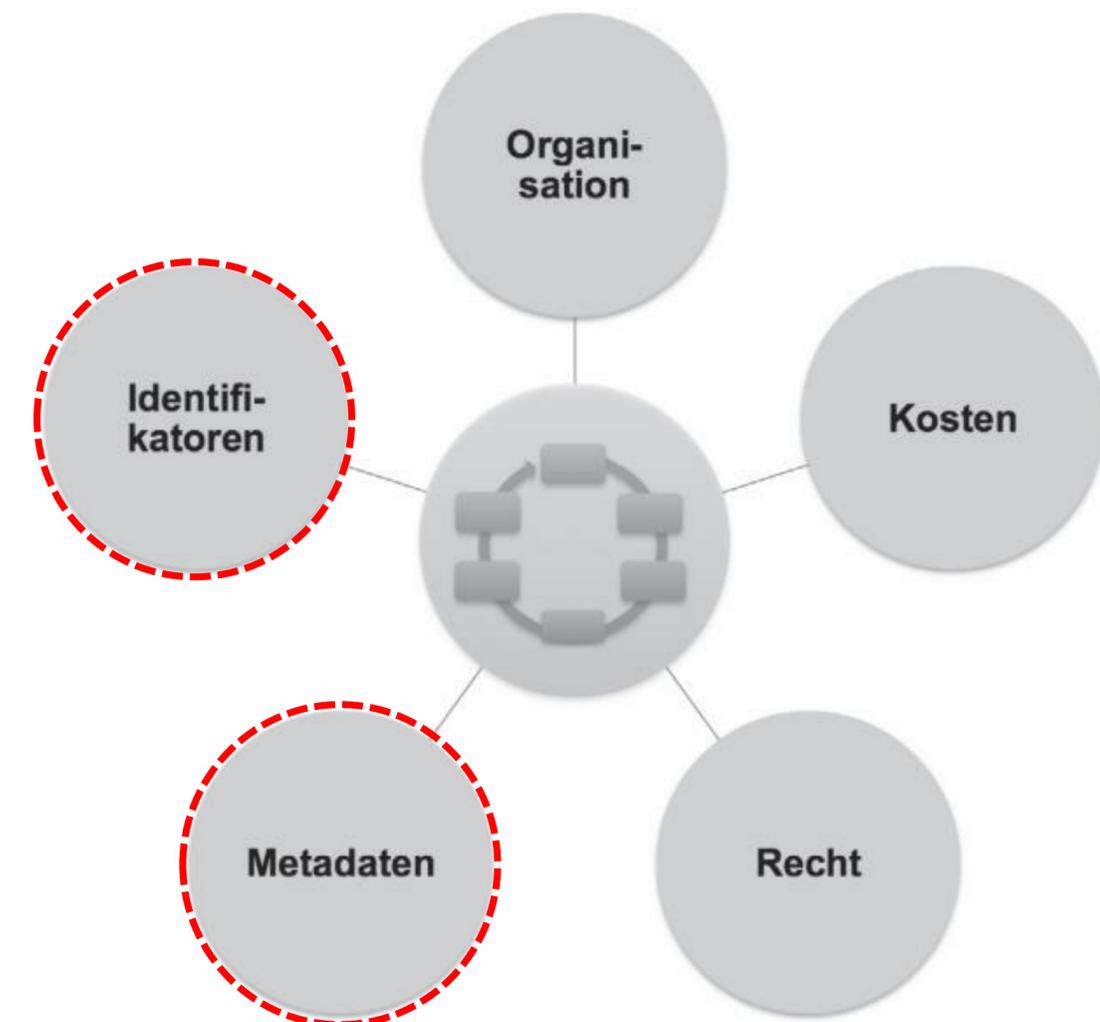
- Geplanter Umgang mit den im Forschungsprozess anfallenden Daten
- Maßnahmen zur Sicherstellung der Nutzbarkeit von Forschungsdaten
 - für eigene/gemeinsame Nutzung im wissenschaftlichen Arbeitsprozess
 - für eine Nachnutzung
 - als Dokumentation der korrekten wissenschaftlichen Arbeitsweise
 - für die Archivierung
- Ziele definieren
 - Qualitätssicherung
 - Replizierbarkeit
 - Nachnutzbarkeit
- Data Management Pläne als Instrument

Lebenszyklus von Forschungsdaten



Jens Ludwig, Harry Enke (Hrsg.): Leitfaden zum Forschungsdaten-Management. 2013

Übergreifende Aufgaben



Forschungsdaten publizieren



- Transparenz und Nachprüfbarkeit der Forschungsergebnisse
- Nachnutzung in neuen Kontexten
- Vermeidung von Doppelarbeit
- Sichtbarkeit der Daten → wissenschaftliche Reputation durch Zitationen
- Erfüllung der Anforderungen von Förderinstitutionen wie DFG und EU oder der eigenen Institution
 - Leitlinien zum Umgang mit Forschungsdaten (DFG)
 - Open Data Pilot, Horizon 2020

Probleme



- Forschungsdaten sind nach wie vor größtenteils nicht öffentlich zugänglich
- Forschungsdaten werden nicht oder uneinheitlich mit Metadaten beschrieben
- Für das Referenzieren und Zitieren von Forschungsdaten hat sich bisher kein Standard durchgesetzt
- „Widerstand“ bzw. wenig Bewusstsein bei den Forschenden

DataCite und DOI-System

- Gegründet am 1. Dezember 2009 in London
- Globales Konsortium getragen von lokalen Einrichtungen
- In Kooperation mit Datenzentren und anderen Einrichtungen, die Daten vorhalten

.....100 Mitglieder in 24 Ländern1500 Datenzentren.... 14 Millionen DOIs.....

Ziele von DataCite



- Fokus: Zugang, Zitierung und Verlinkung von Forschungsdaten und anderen nicht-textuellen Inhalten
- Akzeptanz von Forschungsdaten als relevanter, zitierfähiger Bestandteil des wissenschaftlichen Leistungsausweises stärken
- Unterstützung der Forschenden durch Services, die Forschungsdaten auffindbar, eindeutig identifizierbar und zitierbar machen
- Mithilfe bei der Entwicklung, Implementierung und Förderung von Standards, Workflows und Best Practices
- Metriken für Forschungsdaten

DOI-System



- 1998 Gründung der International DOI Foundation (IDF) als übergeordnete Organisation
- Derzeit 10 Registrierungsagenturen, DataCite ist eine davon
- Das IDF System garantiert vertrauenswürdige Verantwortlichkeiten, einheitliche Standards und Workflows
- DOI ist eine geschützte Marke
- DOI ist eine Vermarktung des technisch geprägten Handle Systems
- DOI-System ist ISO Standard 26324
- DOI steht für Publikation (Zitierfähigkeit, Nachhaltigkeit)

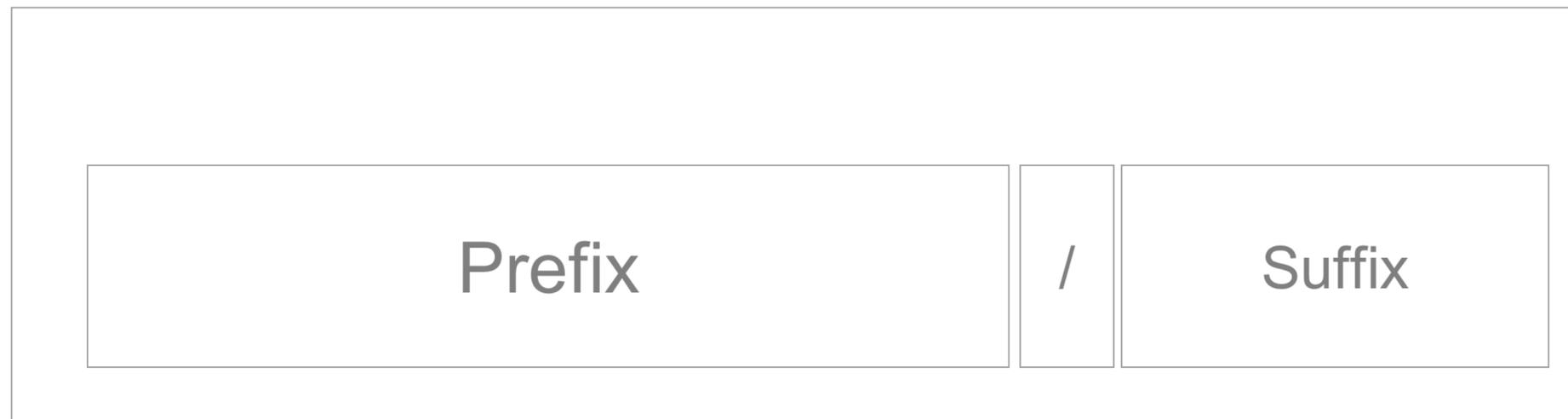


DOI Resolving



- Der Digital Object Identifier DOI als persistenter Identifikator verweist auf das Objekt selbst, ein DOI-Name bleibt unabhängig von URL-Änderungen gültig
- DOI-Namen sind zitierfähig
- DOI-Namen können in jedem Handleserver weltweit aufgelöst werden, ohne Kenntnis der Vergabeinstitution

DOI-Namen



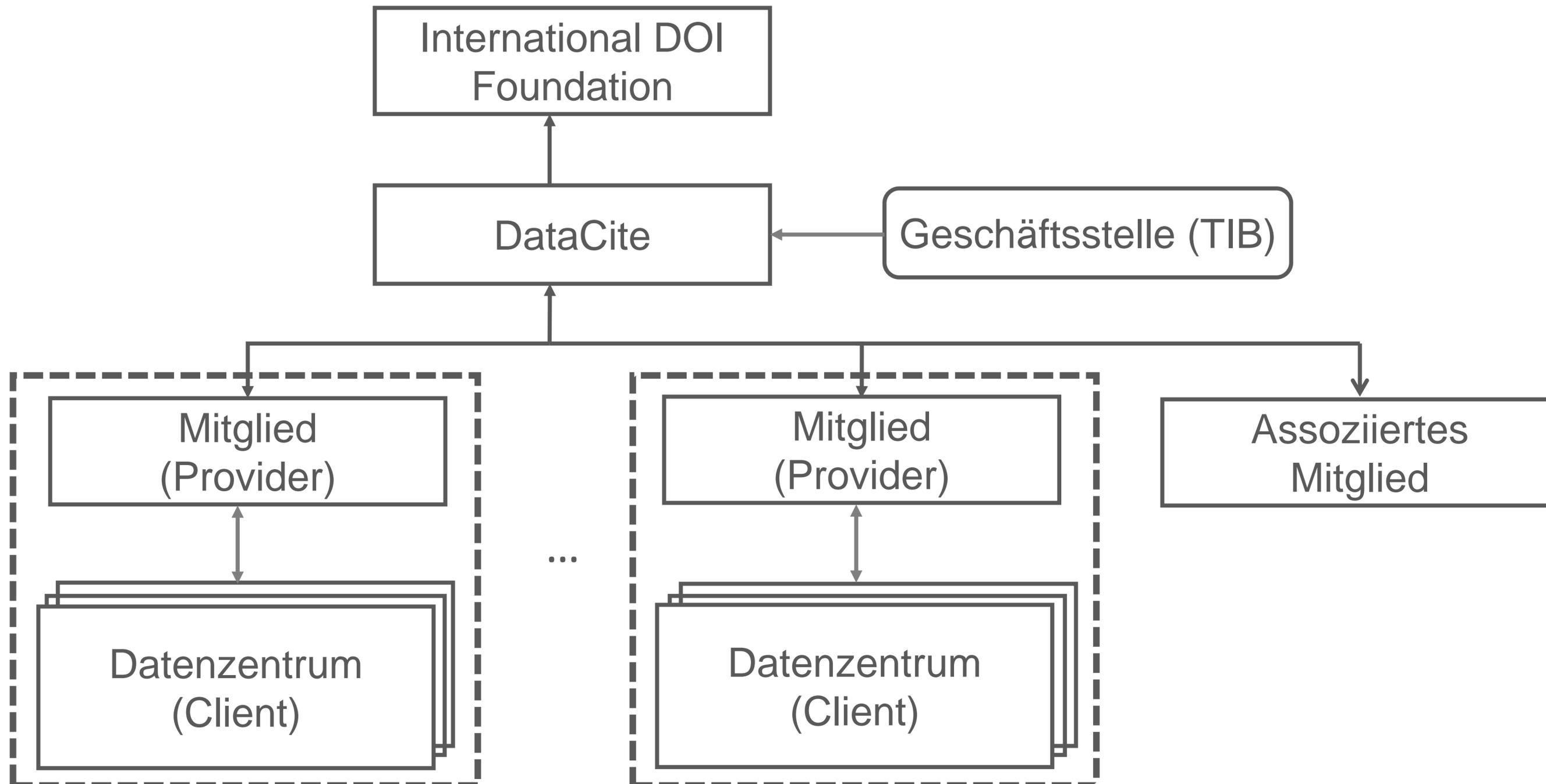
Example:



0-9,
a-z,
A-Z,
-(Bindestrich),
(Punkt),
(Unterstrich),
+ (Plus),
: (Doppelpunkt)
(Schrägstrich)

Achtung: DOI-Namen sind nicht Groß-/Kleinschreibungs-unterscheidend.

DataCite-Struktur



Deutsche DataCite-Mitglieder



Fachliche Aufteilung

- Geisteswissenschaften



- Lebenswissenschaften



- Sozialwissenschaften



- Technik & Naturwissenschaften



- Wirtschaftswissenschaften



Datenzentrum

- Archivierung und Zugänglichkeit für mindestens 10 Jahre
- Qualitätssicherung
- Metadatenzulieferung
- Bereitstellung der sogenannten Landing Pages
- Pflege und Persistenz

DataCite-Mitglied in Kooperation mit DataCite

- Präfix-Vergabe
- dauerhafte Gewährleistung der Auflösbarkeit der DOIs auf die vom Datenzentrum genannten Ziel-URLs
- dauerhafte Bereitstellung der technischen Infrastruktur
- Bereitstellung eines Suchinterface
- Beratung und Support der Datenzentren

Weg zum DOI



- Kooperationsvereinbarung zwischen Datenzentrum und DataCite-Mitglied: in Deutschland DOI-Vergabe für akademische Einrichtungen seit 2013 kostenfrei
- Einrichtung eines Accounts für die DOI-Registrierung für das Datenzentrum durch das DataCite-Mitglied
- Zuweisung eines Präfixes (oder mehrerer Präfixe) durch das DataCite-Mitglied
- Unterstützung des Datenzentrums durch das DataCite-Mitglied beim praktischen Einstieg (Einführung, Metadatenbeispiele, Testumgebung, Schulungen ...)

Einführung DOI-Registrierung

Workflow



Daten



Wissenschaftler

```

<resource xsi:schemaLocation="http://datacite.org/schema/kernel-3 http://s
<identifier identifierType="DOI">10.15466/LERNZdb-84</identifier>
-<creators>
  -<creator>
    <creatorName>McBride, Chris</creatorName>
  </creator>
</creators>
-<titles>
  <title>Global planktonic foraminifera seasonality, supplement to:
  </title>
</titles>
-<publisher>
  Environmental Research Institute (ERI), The University of Waikato
</publisher>
<publicationYear>2014</publicationYear>
+<subjects></subjects>
-<contributors>
  +<contributor contributorType="DataManager"></contributor>
  +<contributor contributorType="HostingInstitution"></contributor>
</contributors>
-<dates>
  <date dateType="Issued">2014-06-04</date>
  <date dateType="Collected">2003-02-13/2003-02-13</date>
</dates>
<language>en</language>
<resourceType resourceTypeGeneral="Dataset">Spreadsheet</resour
+<alternateIdentifiers></alternateIdentifiers>
-<formats>
  -<format>
    application/vnd.openxmlformats-officedocument.spreadsheetml.sheet
  </format>
  
```

Metadaten & URL

DOI



Kataloge/Suchportale



Global planktonic foraminifera seasonality, supplement to:
Jonkers, Lukas; Kucera, Michal (2015): Global analysis of
seasonality in the shell flux of extant planktonic Foraminifera
Biogeosciences, 12(7), 2207-2226

[Jonkers, Lukas / Kucera, Michal](#)
von PANGAEA - Data Publisher for Earth & Environmental Science; 2015

Dokumentinformationen

Format / Umfang: 2 datasets
DOI: [10.1594/PANGAEA.848529](#)
Dokumenttyp: Forschungsdaten
Dokumentformat: Elektronische Ressource
Sprache: Englisch

Abstract

Shell fluxes of planktonic Foraminifera species vary intra-annually in a pattern that appears to follow the seasonal flux maxima in space and among species remains poorly constrained. Thus, although changing seasonal C within a species, this effect is often ignored in the interpretation of Foraminifera-based paleoceanographic record pattern of shell flux variability in 37 globally distributed time series. The existence of a seasonal component in flu This analysis yielded estimates of the number, timing and prominence of seasonal flux maxima. Over 80% of the periodic component, indicating that a considerable part of the intra-annual flux variability is predictable. Temperatu affect differ among species. These different modes of seasonality are distinguishable. Trophic and cultural...

fA_By_FecalAggression

Open Access

Titel: fA_By_FecalAggression
Autor: Holekamp, Kay E. ; Swanson, Eli M. ; Van Meter, Page E.
Inhalt: Data associated with this file show the relationship between maternal fecal androgen their adult peers (n = 14).
Erscheinungsjahr: 2013-05-20T17:57:34Z
Dokumentart: Dataset ; none
Schlagwörter: behavioural flexibility ; constraints ; maternal effects ; morphology ; evolvability
Rechte: <http://creativecommons.org/publicdomain/zero/1.0/>
Nachnutzung: CC0
Beziehungen: doi:10.5061/dryad.305mq
URL: <http://hdl.handle.net/10255/dryad.46380>
Datenlieferant: Dryad Digital Repository (Duke University)

DOI Metadaten



- DOI-Registrierung nur mit Übermittlung von Metadaten möglich
- DataCite Metadata Schema
 - Metadatenschema für die Verlinkung von Daten und anderen Objekten
 - disziplinübergreifend
 - auf Dublin Core basierend
 - wird laufend aktualisiert (Metadata Working Group)
 - <https://schema.datacite.org/>

DataCite Metadata Schema (1)



Pflichtfelder

- Identifier (mit Attribut ‚type‘)
- Creator (mit optionalen Attributen ‚name identifier‘ und ‚affiliation‘)
- Title (mit optionalem Attribut ‚type‘)
- Publisher
- Publication Year
- ResourceType (mit Attribut ‚general type description‘)

Zitation

- Creator (Publication Year): Title. Publisher. (resourceTypeGeneral). Identifier

DataCite Metadata Schema (2)

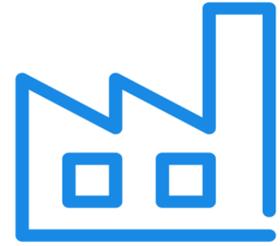


Empfohlene Felder	Optionale Felder
Subject (mit Attribut ,scheme‘)	Language
Contributor (mit Attributen ,type‘, ,name identifier‘ und ,affiliation‘)	Alternateldentifizier (mit Attribut ,type‘)
Date (mit Attribut ,type‘)	Size
RelatedIdentifizier (mit Attributen ,type‘ und ,relation type‘)	Format
Description (mit Attribut ,type‘)	Version
GeoLocation (mit Attributen ,point‘, ,box‘ und ,polygon‘)	Rights
	FundingReference (mit Attributen ,name‘, ,identifizier‘, und ,award related‘)

da|ra Registrierungsagentur für Sozial- und Wirtschaftsdaten



- Eigene Infrastruktur basierend auf DataCite
 - Fachspezifisches, DataCite kompatibles Metadatenschema
 - XML Upload
 - API
 - Suchindex
 - OAI-Schnittstelle
- Betrieben von Gesis und ZBW



DOI Fabrica

DataCite's neue Infrastruktur
für die DOI-Registrierung

Was ist DOI Fabrica?

DOI Fabrica ist das neue Web Frontend für die DOI-Registrierung und -Pflege

- Registrierung & Aktualisierung von DOIs
- Speicherung & Verwaltung von Metadaten
- Alternative/Ergänzung zur Nutzung der API

DataCite Members (Providers) nutzen DOI Fabrica für

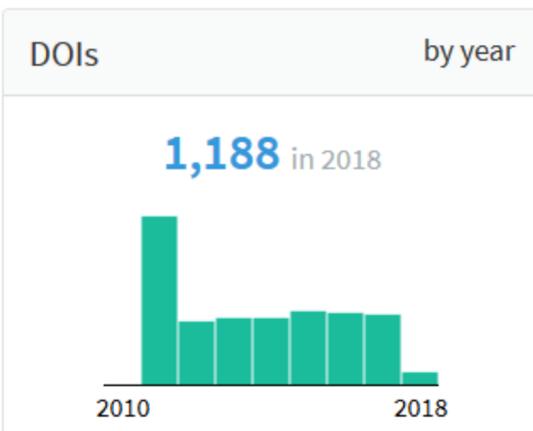
- Verwaltung der Accounts der Datenzentren
- Präfix-Management

Fabrica löst das Metadata Store (MDS) Frontend ab

- MDS Frontend wird bis Ende des Jahres noch parallel betrieben
- MDS API bleibt

Client Dashboard

Info Settings Prefixes DOIs



Atlas of Living Australia

Info Settings Prefixes DOIs

Update Account

Client ID

ANDS.CENTRE63

The Client ID is the unique identifier for each client and can't be changed.

re3data Record (optional)

atlas

Client Name

Alaska Ocean Observing System
Alberta Geological Survey
Allen Brain Atlas

Record created
July 10, 2015

Record last modified
May 2, 2018

Info Settings Prefixes DOIs

Type to search...

Info Settings Prefixes DOIs

2 Prefixes

10.5072

Demo prefix

Added

November 7, 2017

10.3305

Type to search...

Search for (exact) DOI.

53,656 DOIs

Sort by Date Registered ▾

Search

Create DOI (Form)

Create DOI (File Upload)

Was kann DOI Fabrica?

Funktionen des alten MDS Frontend

- DOIs generieren
- Upload von Metadaten
- Aktualisierung bestehender DOIs

NEU

- „States“ für DOIs (Draft, Registered, Findable)
- Upload von Metadaten in verschiedenen Formaten möglich
- Formular zur Registrierung

Draft

- privat
- sind nicht registriert, kein Resolving
- nicht öffentlich auffindbar
- können gelöscht werden
- Metadaten nicht Pflicht
- DOIs mit Test-Präfix 10.5072 sind immer im Status “Draft”

Registered

- aktiv (keine Zurücksetzen auf “Draft” möglich)
- nicht öffentlich auffindbar

Findable

- aktiv
- indexiert in DataCite Search
- Metadaten über OAI-PMH verbreitet
- Status kann in Registered geändert werden

Erzeugen eines DOI per File Upload

NEU

- Integrierte Validierung (grüner Text)
- Möglichkeit der automatischen Generierung von DOI Namen

Create DOI (File Upload)

More information about DOI registration via file upload can be found on our [Support Website](#).

DOI

unique string that identifies the resource and can't be changed. DOIs with demo prefix 10.5072 will always remain in Draft state.

State

- Draft only visible in DOI Fabrica, DOI can be deleted
- Registered registered with the DOI Resolver
- Findable registered with the DOI Resolver and indexed in DataCite Search

The state determines whether a DOI is registered with the DOI Resolver and indexed in DataCite Search. Once in Registered or Findable state, a DOI can no longer be set to Draft state. [More ...](#)

URL

The location of the landing page with more information about the resource.

Metadata

```
<?xml version="1.0" encoding="UTF-8"?>
<resource xmlns="http://datacite.org/schema/kernel-3" xmlns:dip="http://www.dspace.org/xmlns/dspace/dip"
xmlns:dryad="http://purl.org/dryad/terms/" xmlns:dspace="http://www.dspace.org/xmlns/dspace/dim"
xmlns:mets="http://www.loc.gov/METS/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://datacite.org/schema/kernel-3 http://schema.datacite.org/meta/kernel-3/metadata.xsd">
  <identifier identifierType="DOI">10.14454/7tzg-5n54</identifier>
  <version>1</version>
  <creators>
    <creator>
      <creatorName>Ollono, Benjamin</creatorName>
    </creator>
    <creator>
      <creatorName>Durand, Patrick</creatorName>
    </creator>
  </creators>
```

Metadata that describe the resource. Upload a file with metadata in DataCite XML format, or one of the other supported formats. The DOI in the metadata will be replaced with the DOI in the DOI field if different.

Verbesserte Metadatenverarbeitung



- Weitere unterstützte Metadatenformate neben DataCite Format
 - schema.org JSON-LD
 - Citeproc JSON
 - BibTex
 - RIS
- Metadaten werden als DataCite XML gespeichert.
- Direktes Editieren der hochgeladenen Metadaten möglich
- Formular zur Metadateneingabe in DOI Fabrica integriert

Erzeugen eines DOI per Formular

- Bisläng nur Pflichtangaben und zwei optionale Angaben
- Weiter optionale Felder sollen nach und nach integriert werden.

Create DOI (Form)

...

Creators

The main researchers involved in producing the data, or the authors of the publication, in priority order. Please enter one creator per line, with a "Smith, John" format for personal names.

Title

A name or title by which the resource is known.

Publisher

The name of the entity that holds, archives, publishes prints, distributes, releases, issues, or produces the resource.

Publication Year

The year when the resource was or will be made publicly available. Must be between 1450 and 2019.

Resource Type

General

The general type of the resource.

Resource Type

(optional)

A description of the resource, the preferred format is a single term of some detail.

Description

(optional)

All additional information that does not fit in any of the other categories.

DOI-Detailansicht

Info Settings Prefixes **DOIs**

10.14454/3cxn-tm67

URL

<http://www.example.org>

Metadata

Automated quantitative histology reveals vascular morphodynamics during hypocotyl secondary growth Text

Martial Sankar, Kaisa Nieminen, Laura Ragni, Ioannis Xenarios & Christian S Hardtke

Work published 2014

Among various advantages, their small size makes model organisms preferred subjects of investigation. Yet, detailed analysis of numerous developmental processes at cellular level is severely hampered by their scale

<https://handle.test.datacite.org/10.14454/3cxn-tm67>

Citation

Sankar, M., Nieminen, K., Ragni, L., Xenarios, I., & Hardtke, C. S. (2014). Automated quantitative histology reveals vascular morphodynamics during Arabidopsis hypocotyl secondary growth. eLife. <https://doi.org/10.14454/3CXN-TM67>

- Summary View
- Summary View
- DataCite XML
- Schema.org JSON-LD
- Citeproc JSON
- Codemeta
- BibTeX
- RIS
- JATS

- APA
- APA
- Harvard
- MLA
- Vancouver
- Chicago
- IEEE

- Update DOI (Form)
- Update DOI (File Upload)

Findable

DOI registered
April 4, 2018

DOI last updated
May 1, 2018

Nächste Schritte



- Erweiterung des Metadaten-Eingabeformulars
- Statistiken in Form von Berichten
- Link Checker Service

Praktische Arbeit mit DOI Fabrica

Login mit Testzugang unter
<https://doi.test.datacite.org/>

Username: BIBTAG.TEST
Passwort: 2018

Test



- DOIs generieren
 - File-Upload (mit verschiedenen Metadatenformaten)
 - Formular
- DOI-Anzeige
- DOI aktualisieren
- DOI löschen

Feedback



- Fragen? Probleme?
- Usability
- ...
- Was wäre zu verbessern?
- Gibt es weitere Anforderungen aus Sicht ihrer Einrichtung?
- ...

Granularität und Versionierung

Granularität Empfehlungen International DOI Foundation



“A DOI name can be assigned to any object, regardless of the extent to which that object might be a component part of some larger entity.

DOI names can be assigned at any desired degree of precision and granularity that a registrant deems to be appropriate.”

DOI Handbook
(https://www.doi.org/doi_handbook/2_Numbering.html#2.3.2)

Granularität

Empfehlungen DataCite I



- 1. Citation:** The current citation and research practices among the client's user community: What is likely to be cited?
- 2. The use of data:** The needs of various stakeholders: how will funders/publishers/administrators etc. use the data?
- 3. The type of resource:** For example a complex dataset may require a more granular identifier structure than a document or image file.
- 4. Sustainability:** The client must be able to maintain each item with a DOI name in accordance with DataCite client responsibilities.

Granularität Empfehlungen DataCite II



Falls erforderlich:

- mehrere DOI-Namen für verschiedene Granularitätsstufen eines Objektes vergeben

und

- in den Metadaten im Element `<relatedIdentifier>` mit *“isPartOf/hasPart”* in Relation setzen

Granularität Beispiele (Sozialwissenschaften)

text

Volkshochschule für die Stadt und den Kreis
Leer: Programm 2. Semester 1966, S. 1



DOI [10.12764/15591_p1](https://doi.org/10.12764/15591_p1)

Version 1.0

Typ der Ressource Text : VHS-Programm

Creator

- Volkshochschule für die Stadt und den Kreis Leer

Übergeordneter Titel

- Programme der Volkshochschule für die Stadt und den Kreis Leer inklusive retrospektiver und zukünftiger Entwicklungen ausgehend von 2004

Veröffentlichungsdatum 2014

Contributor

- Deutsches Institut für Erwachsenenbildung (DIE) / Daten- und Informationszentrum (DIZ) / Servicestellen / Servicestelle Dokumentation und Geschichte (Vertrieb)
- Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen (Datenerhebung)
[zeige weniger](#)

Schlagwörter (frei) Volkshochschule; Kurs; Bildungsangebot; Volkshochschule für die Stadt und den Kreis Leer

Referenzzeitraum

- 1966 - 1966

Geographischer Raum

- Leer, Ostfriesland

Verfügbarkeit ● Download
verfügbar als: paper; xml/mets_mods; pdf; txt;

Verbundene Identifier

- Ist Teil von
DOI: 10.12764/15591
- Wird fortgesetzt von
DOI: 10.12764/15591_p2

Aktualisierung Metadaten: 14.09.2014 | Ausgabennummer: 2 | Registration Date: 21.12.2013

dataset

Gesamtansicht

SOEP-Innovationssample (SOEP-IS), Daten der Jahre 1998-2012



DOI [10.5684/soep.is.2012](https://doi.org/10.5684/soep.is.2012)

Version 2012

Typ der Ressource Datensatz

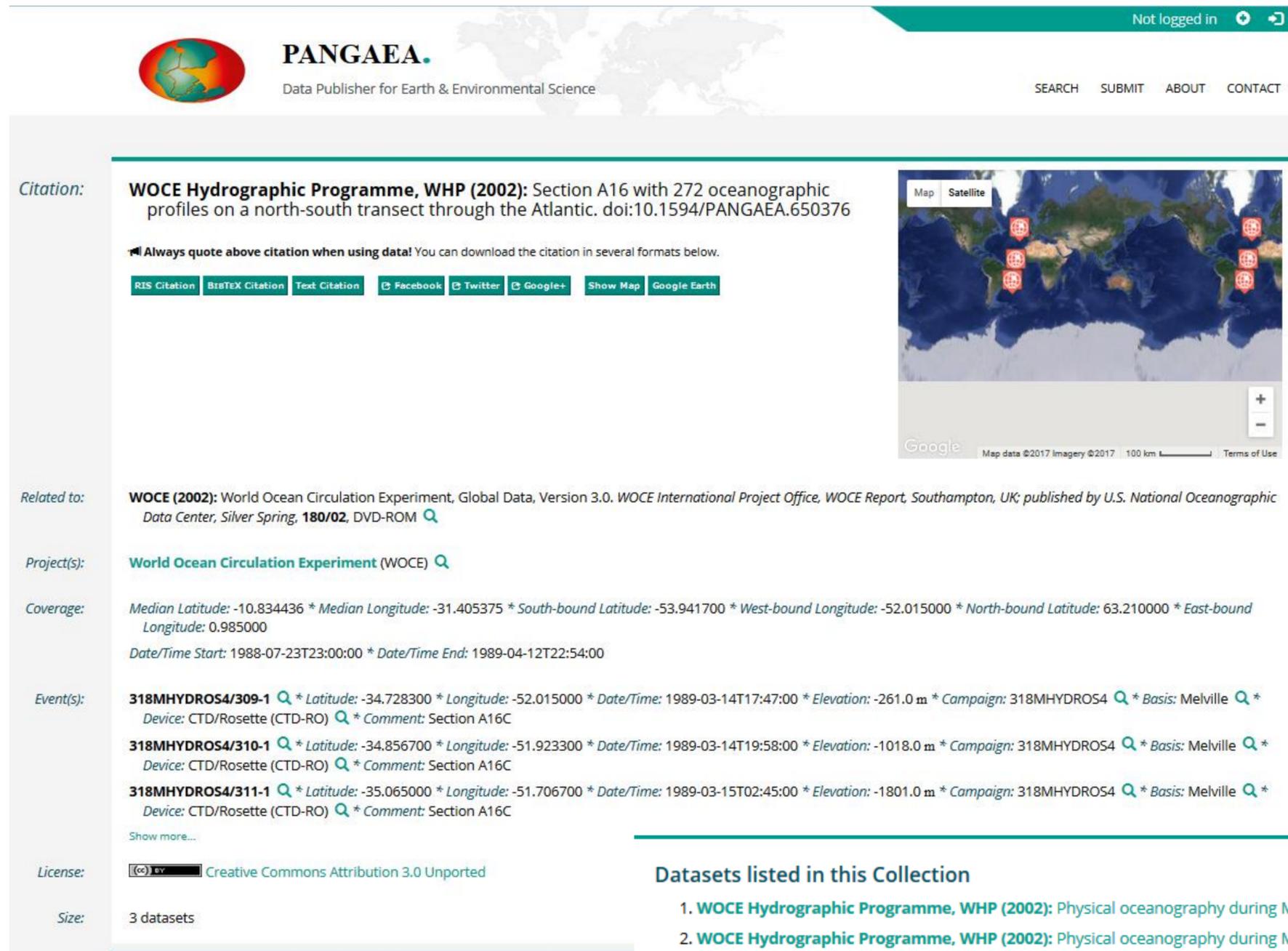
Creator

Daten und Dateinformationen

- Name: soep-is.2012_de_spss.zip
Datenformat: SPSS_de
Daten-Fingerprint: ee0615c883c567ecb2a0a5f72519a70c
Verfahren Fingerprint: MD5
- Name: soep-is.2012_de_stata.zip
Datenformat: Stata_de
Daten-Fingerprint: a0c009f489122ebd5ec18d511630b899
Verfahren Fingerprint: MD5
- Name: soep-is.2012_en_spss.zip
Datenformat: SPSS_de
Daten-Fingerprint: 1e9f3ea964487947c054ca6c138ef4ba
Verfahren Fingerprint: MD5
- Name: soep-is.2012_en_stata.zip
Datenformat: Stata_en
Daten-Fingerprint: 975c9ac2fb38b23de07f5dd9ecb372d9
Verfahren Fingerprint: MD5

[zeige weniger](#)

Granularität Beispiele (Naturwissenschaften)



PANGAEA.
Data Publisher for Earth & Environmental Science

Not logged in

SEARCH SUBMIT ABOUT CONTACT

Citation: **WOCE Hydrographic Programme, WHP (2002):** Section A16 with 272 oceanographic profiles on a north-south transect through the Atlantic. doi:10.1594/PANGAEA.650376

Always quote above citation when using data! You can download the citation in several formats below.

RIS Citation BibTeX Citation Text Citation Facebook Twitter Google+ Show Map Google Earth



Related to: **WOCE (2002):** World Ocean Circulation Experiment, Global Data, Version 3.0. WOCE International Project Office, WOCE Report, Southampton, UK; published by U.S. National Oceanographic Data Center, Silver Spring, 180/02, DVD-ROM

Project(s): **World Ocean Circulation Experiment (WOCE)**

Coverage: Median Latitude: -10.834436 * Median Longitude: -31.405375 * South-bound Latitude: -53.941700 * West-bound Longitude: -52.015000 * North-bound Latitude: 63.210000 * East-bound Longitude: 0.985000
Date/Time Start: 1988-07-23T23:00:00 * Date/Time End: 1989-04-12T22:54:00

Event(s): **318MHYDROS4/309-1** * Latitude: -34.728300 * Longitude: -52.015000 * Date/Time: 1989-03-14T17:47:00 * Elevation: -261.0 m * Campaign: 318MHYDROS4 * Basis: Melville * Device: CTD/Rosette (CTD-RO) * Comment: Section A16C
318MHYDROS4/310-1 * Latitude: -34.856700 * Longitude: -51.923300 * Date/Time: 1989-03-14T19:58:00 * Elevation: -1018.0 m * Campaign: 318MHYDROS4 * Basis: Melville * Device: CTD/Rosette (CTD-RO) * Comment: Section A16C
318MHYDROS4/311-1 * Latitude: -35.065000 * Longitude: -51.706700 * Date/Time: 1989-03-15T02:45:00 * Elevation: -1801.0 m * Campaign: 318MHYDROS4 * Basis: Melville * Device: CTD/Rosette (CTD-RO) * Comment: Section A16C

Show more...

License:  Creative Commons Attribution 3.0 Unported

Size: 3 datasets

Datasets listed in this Collection

- WOCE Hydrographic Programme, WHP (2002):** Physical oceanography during MELVILLE cruise 318MHYDROS4 on section A16C. doi:10.1594/PANGAEA.293952
- WOCE Hydrographic Programme, WHP (2002):** Physical oceanography during MELVILLE cruise 318MSAVE5 on section A16S. doi:10.1594/PANGAEA.293954
- WOCE Hydrographic Programme, WHP (2002):** Physical oceanography during OCEANUS cruise 32OC202_1 and 32OC202_2 on section A16N. doi:10.1594/PANGAEA.293953

272 ozeanographische Profile

Versionierung

- systematische Erfassung von Änderungen an Objekten
- bei jeder **Änderung** → **neue Version** dieses Objektes
- Zugriff auf ältere Versionen eines Objektes
- Nachvollziehbarkeit durch Protokollierung der Änderungen

Versionierung Empfehlungen International DOI Foundation



“The choice of whether to assign a **new DOI name** to a **changed** item which already has a DOI name is one of functional granularity:

→ **it should be possible to identify an entity whenever it needs to be distinguished.**

The IDF does not itself dictate any rules on this.

Individual Registration Agencies may decide to adopt specific rules appropriate for their community.”

DOI Handbook

https://www.doi.org/doi_handbook/2_Numbering.html#2.3.3

Versionierung

Empfehlungen DataCite



- Ein Objekt mit einem zugewiesenen DOI-Namen solte nicht verändert werden. (Data sharing: **Nachvollziehbarkeit** und **Nachnutzbarkeit**)
- Jede Änderung solte als **neue Version** gespeichert werden und mit einem **neuen DOI-Namen** versehen werden.

Die Verantwortung für die Versionierung liegt beim **Datenanbieter/ Datenzentrum**.

Versionierung

Empfehlung aus den Sozialwissenschaften



- Es gibt sehr unterschiedliche Wege zu versionieren. Wichtig sind: **Einheitlichkeit und Systematik!**
- Zu Beginn der Erzeugung der Daten könnte es sinnvoll sein “Milestone-Versionen” der später zu registrierenden Daten zu definieren und separat zu speichern.
- Änderungen in diesen Milestone-Versionen sollten gut dokumentiert werden und als neue Versionen abgespeichert werden.

Versionierungskonzepte

Empfehlung aus den Sozialwissenschaften



Drei Ebenen-Versionierung: **Major.Minor.Revision**

Die Major-Nummer beginnt mit „1“. Minor- und Revision-Nummern beginnen immer mit „0“. Getrennt durch „.“ Erste Version eines Datenfiles → „1.0.0“.

- 1. Stelle (**Major**): Daten hinzugefügt z.B. neue Variablen, Erhebungswellen, Stichproben etc.
- 2. Stelle (**Minor**): Korrekturen, die die Analyse der Daten beeinflussen (z.B. Änderungen von Werten/Umcodierungen)
- 3. Stelle (**Revision**): Dokumentation geändert oder verbessert (Schreibfehler korrigiert oder Text zugefügt u.ä.)

In Anlehnung an den Standard der Data Documentation Initiative (DDI)

Versionierungskonzepte

Empfehlung aus den Sozialwissenschaften



- Anzahl der **Versionierungsebenen** und ihre Parameter an den jeweiligen Bedarf des Projektkontextes **anpassen**
- außerhalb des Datenfiles **Versionshistorie** dokumentieren
- im Datensatz **Versionsnummer** dokumentieren und bei DOI-Registrierung in Metadatenelementen **<version>** und **<relatedIdentifier>** (*<IsNewVersionOf>* *<IsPreviousVersionOf>*) angeben

Versionierung

Beispiel aus den Sozialwissenschaften



 **Metadata Search beta**

DataCite

Filter

- allocator
- datacentre
- prefix
- resourceType
- contributor
- creator
- publicationYear
- publisher
- language

No active filters. Use the sidebar to filter search results.

9 documents found in 2178ms

Page 1 of 1

[Eurobarometer 74.2 \(2010\)](#)

[version 2.0.0]

[doi:10.4232/1.11588](#) Dataset

European Commission, Brussels DG Communication (COMM.A.1 'Research and Speechwriting')
title: [Eurobarometer 74.2 \(2010\)](#)

[Eurobarometer 74.2 \(2010\)](#)

[version 2.1.0]

[doi:10.4232/1.11600](#) Dataset

European Commission, Brussels DG Communication (COMM.A.1 'Research and Speechwriting')
title: [Eurobarometer 74.2 \(2010\)](#)

[Eurobarometer 74.2 \(2010\)](#)

[version 2.2.0]

[doi:10.4232/1.11626](#) Dataset

Europäische Kommission
title: [Eurobarometer 74.2 \(2010\)](#)

[Eurobarometer 74.2 \(2010\)](#)

[version 1.0.0]

[doi:10.4232/1.10338](#) Dataset

European Commission, Brussels DG Communication (COMM.A.1 'Research and Speechwriting')
title: [Eurobarometer 74.2 \(2010\)](#)

[Eurobarometer 74.2 \(2010\)](#)

[version 1.1.0]

[doi:10.4232/1.10707](#) Dataset

European Commission, Brussels DG Communication (COMM.A.1 'Research and Speechwriting')
title: [Eurobarometer 74.2 \(2010\)](#)

Options |

da|ra
Registrierungsagentur für
Sozial- und Wirtschaftsdaten

Hilfe | Kontakt | Anmelden
Schnellsuche in Metadaten

Mein da|ra | Für Datenzentren | Für Forscher | Für Verlage | Über uns | Aktuell

Suchergebnis

Erweiterte Suche | Hilfe zur Suche

Ihre Suche: Titel: Eurobarometer 74.2 (2010)
Aktive Filter: **Datenanbieter: GESIS Datenarchiv**

Treffer: 5
Ihre Suche ergab 5 Treffer in allen Ressourcen.

dataset	dataset	dataset	dataset
Eurobarometer 74.2 (2010) gesis			

Resources
[registriert \(5\)](#)

Resource types
[Datensatz \(5\)](#)

Datenanbieter
 GESIS Datenarchiv

Creator
[European Commission, Brussels DG Communication \(COMM.A.1 'Research and Speechwriting'\) \(4\)](#)
[Europäische Kommission \(1\)](#)

Contributor
[Capacent, Reykjavik, Island \(5\)](#)
[Emor, Tallinn, Estonia \(5\)](#)
[Kadem, Nicosia, Turkish Cypriot Community \(5\)](#)
[MISCO, Valletta, Malta \(5\)](#)
[Puls, Zagreb, Croatia \(5\)](#)

[mehr >](#) [alle >](#)

Verfügbarkeit
[Lieferbar \(4\)](#)
[Download \(1\)](#)

Versionierung

Beispiel aus den Sozialwissenschaften



gis  | [Contact](#) | [Email to Data Service](#)

GESIS Data Catalogue 2.1

[Home](#) | [Search](#) | **[Browse](#)** | [Overview](#) | [News](#) | [About](#)

[Login](#) | [Shopping Cart](#)

- Groups
- Products
- Studies**
- Geographic Codes

ZA5449: Eurobarometer 74.2 (2010)

[Bibliographic Citation](#) | [Content](#) | [Methodology](#) | [Data & Documents](#) | [Errata & Versions](#) | [Publications](#)

[Groups](#)

You have requested the DOI for version 1.1.0, but the current version is 2.2.0!

Citation 	<i>European Commission (2013): Eurobarometer 74.2 (2010). TNS OPINION & SOCIAL, Brussels [Producer]. GESIS Data Archive, Cologne. ZA5449 Data file Version 2.2.0, doi:10.4232/1.11626</i>
Study No.	ZA5449
Title	Eurobarometer 74.2 (2010)

Member of the  Leibniz Association

Versionierung

Beispiel aus den Sozialwissenschaften



gesis | Contact | Email to Data Se

GESIS Data Catalogue 2.1

Home Search **Browse** Overview News About

Login Shopping Ca

Groups
Products
Studies
Geographic Codes

ZA5449: Eurobarometer 74.2 (2010)

Bibliographic Citation Content Methodology Data & Documents **Errata & Versions** Publications

Groups

Errata in current version Version 2.2.0 Version 2.1.0 Version 2.0.0 Version 1.1.0 Version 1.0.0

Errata in current version	2014-2-20	Media use index (V726)	The derivation for the MEDIA USE INDEX as documented in Standard Eurobarometer 74 (report page 20, footnote 8), does not consider DK answers. Accordingly DK answers in one or more input variables (QD3: V550 to V554) were recoded to MISSING VALUE code 9 in the index variable (V726). In order to reproduce the original index variable as provided by TNS and used for the official tabular reports, please run the SPSS instructions as provided in the patch syntax file 'ZA5449_patch_qd3_v2-2-0.sps'.

Member of the Leibniz Association

da|ra Registration agency for social and economic data

Versionierung

Beispiel aus den Sozialwissenschaften



Metadata Search beta Eurobarometer 74.2 Search

No active filters. Use the sidebar to filter search results.
 9 documents found in 2178ms
 Page 1 of 1

related/identifier

- IsNewVersionOf:DOI:10.4232/1.10707
- IsNewVersionOf:DOI:10.4232/1.10338

[version 2.1.0]
 doi:10.4232/1.11600 Dataset
 European Commission, Brussels DG Communication (COMM.A.1 'Research and Speechwriting')
 title: Eurobarometer 74.2 (2010)

Eurobarometer 74.2 (2010)
 [version 2.2.0]
 doi:10.4232/1.11626 Dataset
 Europäische Kommission
 title: Eurobarometer 74.2 (2010)

Eurobarometer 74.2 (2010)
 [version 1.0.0]
 doi:10.4232/1.10338 Dataset
 European Commission, Brussels DG Communication (COMM.A.1 'Research and Speechwriting')
 title: Eurobarometer 74.2 (2010)

Eurobarometer 74.2 (2010)
 [version 1.1.0]
 doi:10.4232/1.10707 Dataset
 European Commission, Brussels DG Communication (COMM.A.1 'Research and Speechwriting')
 title: Eurobarometer 74.2 (2010)

Options |

da|ra Registrierungsagentur für Sozial- und Wirtschaftsdaten

Hilfe | Kontakt | Anmelden
 Schnellsuche in Metadaten

Mein da|ra Für Datenzentren Für Forscher Für Verlage Über uns Aktuell

Suchergebnis

Erweiterte Suche
 Hilfe zur Suche

Ihre Suche: Titel: Eurobarometer 74.2 (2010)
 Aktive Filter: Datenanbieter: GESIS Datenarchiv

Treffer: 5
 Ihre Suche ergab 5 Treffer in allen Ressourcen.

dataset Eurobarometer 74.2 (2010) gesis
 DOI: 10.4232/1.10707

Verbundene Identifier

- Ist neue Version von DOI: 10.4232/1.11763
- Ist neue Version von DOI: 10.4232/1.11341

dataset Eurobarometer 74.2 (2010) gesis

DOI: 10.4232/1.11600
 Version: 2.1.0
 Creator: European Commission, Brussels DG Com ... zeige mehr
 Referenzzeitraum: 2010-11-11 - 2010-12-01
 • 20 ... zeige mehr
 Veröffentlichungsdatum: 2013-04-05
 Sprache: Englisch
 Verfügbarkeit: Lieferbar

Contributor

- Capacent, Reykjavik, Island (5)
- Emor, Tallinn, Estonia (5)
- Kadem, Nicosia, Turkish Cypriot Community (5)
- MISCO, Valletta, Malta (5)
- Puls, Zagreb, Croatia (5)

mehr > alle >

dataset Eurobarometer 74.2 (2010) gesis

DOI: 10.4232/1.11588
 Version: 2.0.0
 Creator: European Commission, Brussels DG Com ... zeige mehr
 Referenzzeitraum: 2010-11-11 - 2010-12-01

Verfügbarkeit

- Lieferbar (4)
- Download (1)

Versionierungskonzepte

Beispiel Naturwissenschaften: High Energy Physics



- Große Mengen an Experimentaldaten, die verschiedene Verarbeitungsstufen bis hin zu extrahierten Tabellen in Publikationen umfassen
- Diese Versionen stehen nicht immer in linearer Beziehungen zu einander, sondern in Form verzweigter Strukturen
- Dafür bietet das DataCite Metadata Schema die Beziehungen “IsDerivedBy” and “IsSourceOf” an.

Versionierung

Beispiel: High Energy Physics (HEP-Data)



- **HEP-Data:** Datenbank für publikationsreife Tabellen die am Ende der HEP provenance pipeline stehen
- Jeder Record besteht aus den Tabellen einer Publikation
- Jede Tabelle hat einen DOI-Namen
- Jeder Record hat einen DOI-Namen und ist verbunden (relatedIdentifier) mit dem DOI-Namen der Publikation und den IDs aus den Repositorien arXiv und INSPIRE

<https://hepdata.net>

Versionierung

Beispiel: High Energy Physics (HEP-Data)



HEPData search results for 'energy'. The interface shows a search bar with 'energy' entered, a search button, and a 'Reset' link. Below the search bar, there are filters for 'Max results', 'Sort by', and 'Reverse order'. The results show 'Showing 25 of 3616 results'. A date histogram is visible, and a specific result is highlighted: 'A Measurement of J / psi decay widths' by 'The BES collaboration' (Bai, J.Z.; Chen, G.P.; Chen, H.F.; et al.), published in 'Phys.Lett. B355 (1995) 374-380, 1995'. The DOI '10.17182/HEPDATA.28500' is also visible.

Metadata Search beta

DataCite
10.17182/hepdata.28500 Search

- Filter
- allocator
- datacentre
- prefix
- resourceType
- publicationYear
- publisher
- language

No active filters. Use the sidebar to filter search results.

1 documents found in 55ms
Page 1 of 1

A Measurement of J / psi decay widths
[version 1]
doi:10.17182/HEPDATA.28500 Collection : Collection
Bai, J.Z. • Chen, G.P. • Chen, H.F. • Chen, S.J. • Chen, S.M. • (et. al.)
doi: **10.17182/HEPDATA.28500**

Page 1 of 1

DataCite Metadata Search • Query Time: 55ms • API Query : xml, json, csv • OAI-PMH

Metadata Search beta

DataCite
10.17182/hepdata.28500 Search

- Filter
- allocator
- datacentre
- prefix
- resourceType
- publicationYear
- publisher
- language

No active filters. Use the sidebar to filter search results.

1 documents found in 156ms
Page 1 of 1

Table 1
A Measurement of J / psi decay widths
[version 1]
doi:10.17182/HEPDATA.28500.V1/T1 Dataset : Dataset
Bai, J.Z. • Chen, G.P. • Chen, H.F. • Chen, S.J. • Chen, S.M. • (et. al.)
doi: **10.17182/HEPDATA.28500.V1/T1**

Page 1 of 1

DataCite Metadata Search • Query Time: 156ms • API Query : xml, json,

relatedIdentifier

- IsSupplementTo:URL: <http://www.inspirehep.net/record/39870>
- IsSupplementTo:DOI: ~~10.1016/0370-2693(95)00712-T~~
- IsSupplementedBy:DOI: **10.17182/hepdata.28500.v1/t1**

<https://hepdata.net>

Versionierungskonzepte

Beispiel: Software - Zenodo-GitHub



- Erstellung eines snapshot der Software aus dem GitHub code repository und Speicherung in Zenodo
- Jedes code release (Versionierung ist in GitHub implementiert) erhält einen neuen DOI-Namen

Aber:

- Link zum Repository (als “**relatedIdentifier**”) wird z.Z. noch nicht an DataCite übermittelt
- Zenodo vergibt keine DOI-Namen für die Kollektion aller Releases/Versionen

Dynamische Daten



- „Open time series“ wie z.B. Messwerte, die ständig ergänzt, korrigiert bzw. kalibriert werden (z.B. Satellite missions, Ocean profiler usw.)
- Faktendatenbanken (z.B. Zeitreihendatenbanken, Echtzeitdatenbanken der Bundesbank und Statistischen Ämtern usw.)

Persistente Identifikation von dynamischen Daten



1. *(Versionierung: Speicherung eines neuen Datensatzes nach jeder Änderung)*
2. DOI-Registrierung des Datenprojektes, Zitation mit Angabe des Zugriffszeitpunkts (z.B. kontinuierlich aktualisierte Datensätze in einer Datenbank, wobei bereits enthaltene Daten nicht verändert werden)
3. Definieren von bestimmten Abschnitten (zeitlich – *snap shot*, inhaltlich – *slice*) des Datensatzes, für die die Daten archiviert und registriert werden können.
4. *(DOI-Registrierung von mit den Daten verwandten Dokumenten – „data products“, Zitation der Dokumente)*

Weitere Services und Projekte von DataCite

Persistent identifiers: connecting research



Researchers & Contributors



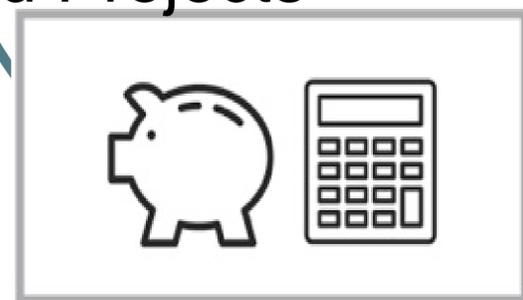
Data and Software



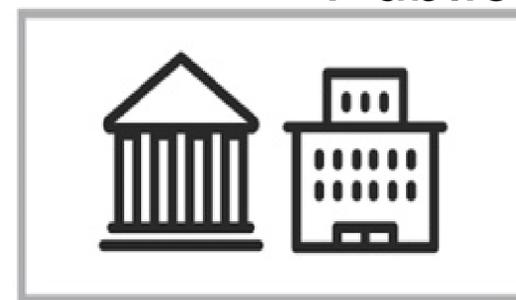
Grants and Projects



Publications



Funders



Institutions



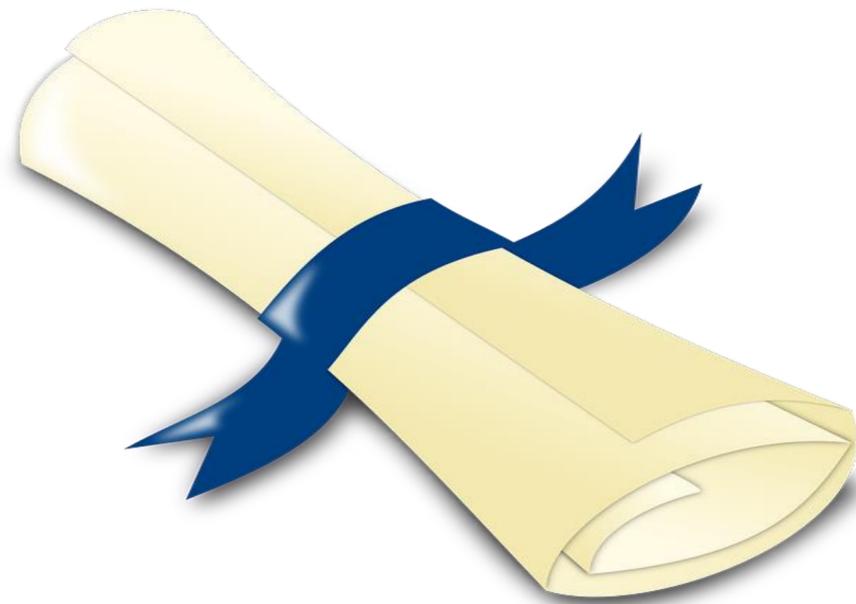
[ORG iD WG](https://www.crossref.org/ORGIDWG/)

Joint Declaration of Data Citation Principles – Force11(2014)



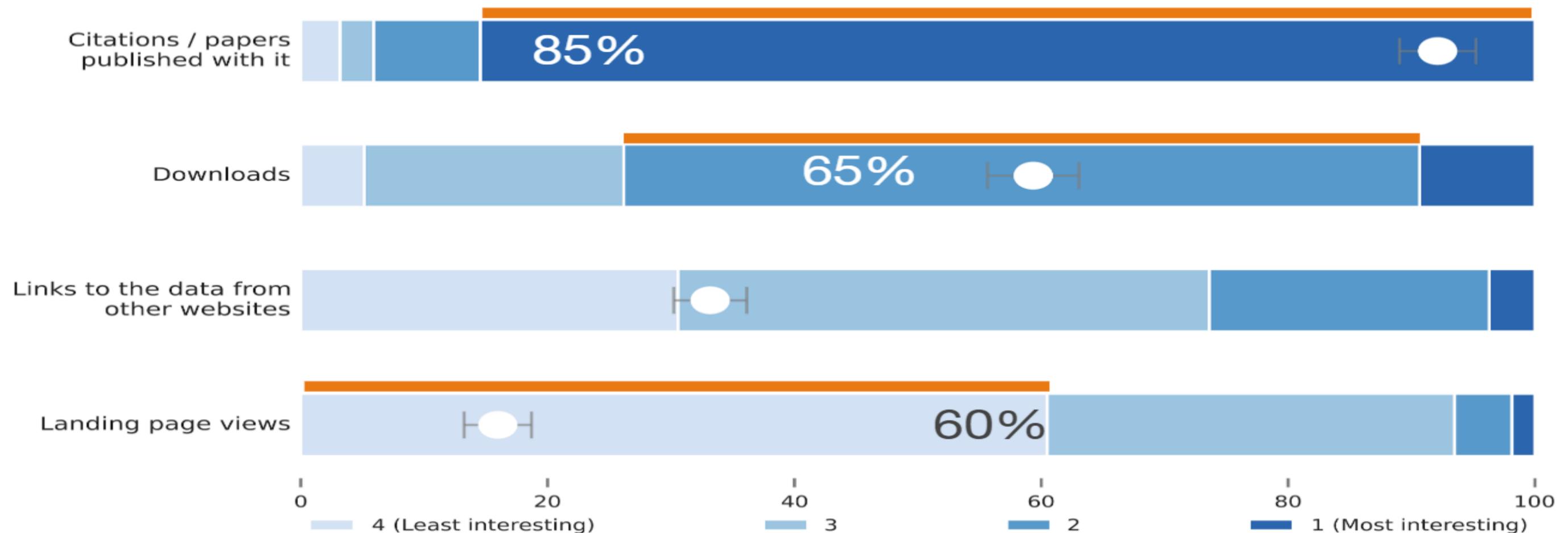
2. Credit and Attribution

Data citations should facilitate giving scholarly credit and normative and legal attribution to all contributors to the data, recognizing that a single style or mechanism of attribution may not be applicable to all data.



Making Data Count (2014–2015)

Survey results (247 researchers responded)



Kratz, J. E., & Strasser, C. (2015). Making data count. *Scientific Data*, 2, 150039.
<https://doi.org/10.1038/sdata.2015.39>

Event Data (2015)



- ✓ DOI reference from reference list of the article
- ✓ Metadata

Relation types =
isSupplementedBy
references

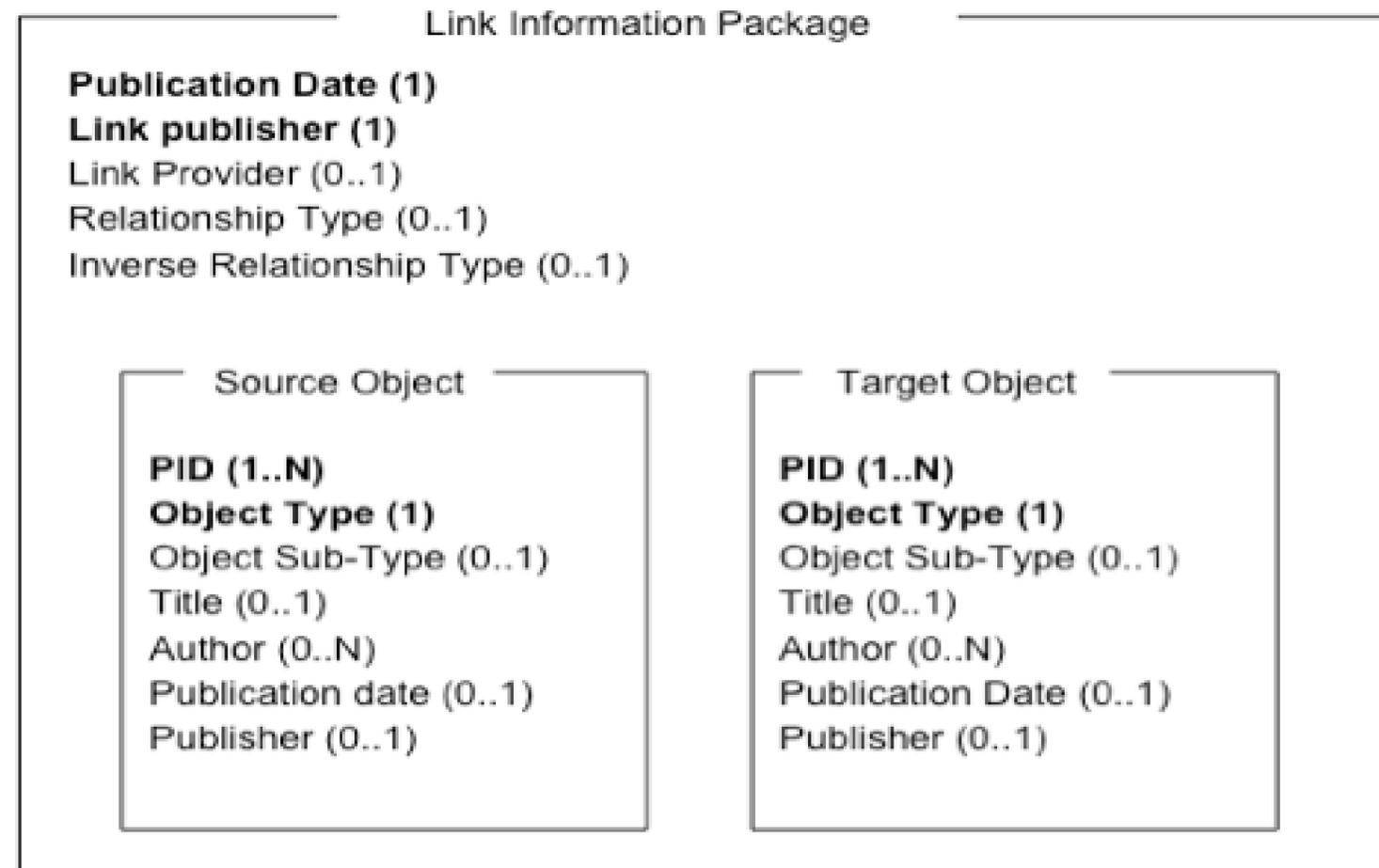
Relation types =
isSupplementedTo
isReferencedBy

- *Only the publisher of the DOI can add the relations to the metadata*
- *Crossref and DataCite provide relations between their DOIs as a service – but don't add them to the metadata*

The staging (preview) instance of the Scholix endpoint:

<http://api.eventdata.crossref.org/v1/events/scholix?rows=100>

Object types =
✓ Dataset
✓ Software





Scholarly Link Exchange

(2016)



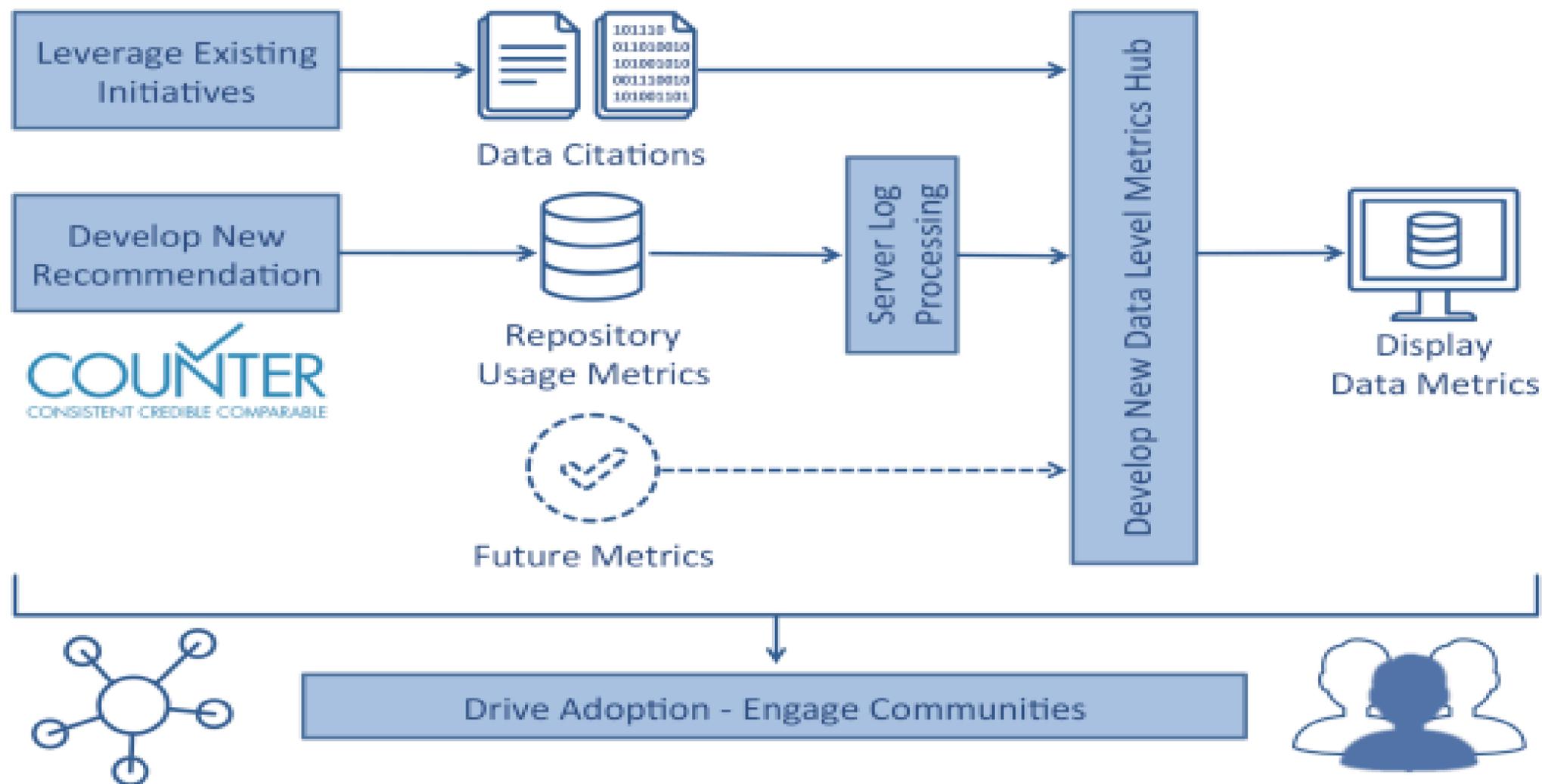
→ *RDA/WDS Scholarly Link Exchange Working Group*

Scholix exchange framework for data-article information links

- A consensus among publishers, datacentres, and global/ domain service providers to work collaboratively and systematically to improve exchange of data-literature link information
- Information model: conceptual definition of what is a Scholix scholarly link
- Link metadata schema: metadata representation of a Scholix link
- Options for exchange protocols (forthcoming)



MAKE
DATA
COUNT (2017–2019)



Lowenberg, Daniella; Cruse, Patricia; Chodacki, John; Fenner, Martin (2018): PIDapalooza Make Data Count Slides. figshare. Presentation. <https://doi.org/10.6084/m9.figshare.5818002.v1>

Success Stories

Researchers

As a researcher I want to know how many people use my data so I can understand and promote impact **so that my work receives appropriate recognition.**

Repositories

As a repository I want to have a standardized way of tracking metrics and citations so I can see impact of datasets **so that I know if I need to provide better discovery channel and optimize discovery services.**

Publishers

As a publisher I want to see the full picture of research output related to the articles I am publishing **so that I can see the impact of published work.**

Funders

As a funder I want to know if people are making use of the data that I have funded **so that I can see my return on investment.**

Analysis Scripts and Data for SAMPL5 Distribution Coefficients Submission Analysis and Reference Calculations,

Bannan, Caitlin C
Burley, Kalistyn H
Chiu, Michael
Gilson, Michael K
Mobley, David L

Publication date: June 18, 2016
Publisher: UC Irvine
<https://doi.org/10.7280/D1988W>

Citation

Bannan, Caitlin C et al. (2016), Analysis Scripts and Data for SAMPL5 Distribution Coefficients Submission Analysis and Reference Calculations,, UC Irvine Dash, Collection, <https://doi.org/10.7280/D1988W>

 Download the dataset ~ 1 GB

 Download Data Publication (PDF)

Versions

> June 18, 2016

Metrics

 41 views

 8 downloads

 1 citations

Minimum viable product (MVP):

- Views and Downloads: Internal logs are processed against the Code of Practice and send standard formatted usage logs to a [DataCite hub](#) for public use and eventually, aggregation.
- Citations: Citation information is pulled from Crossref Event Data.

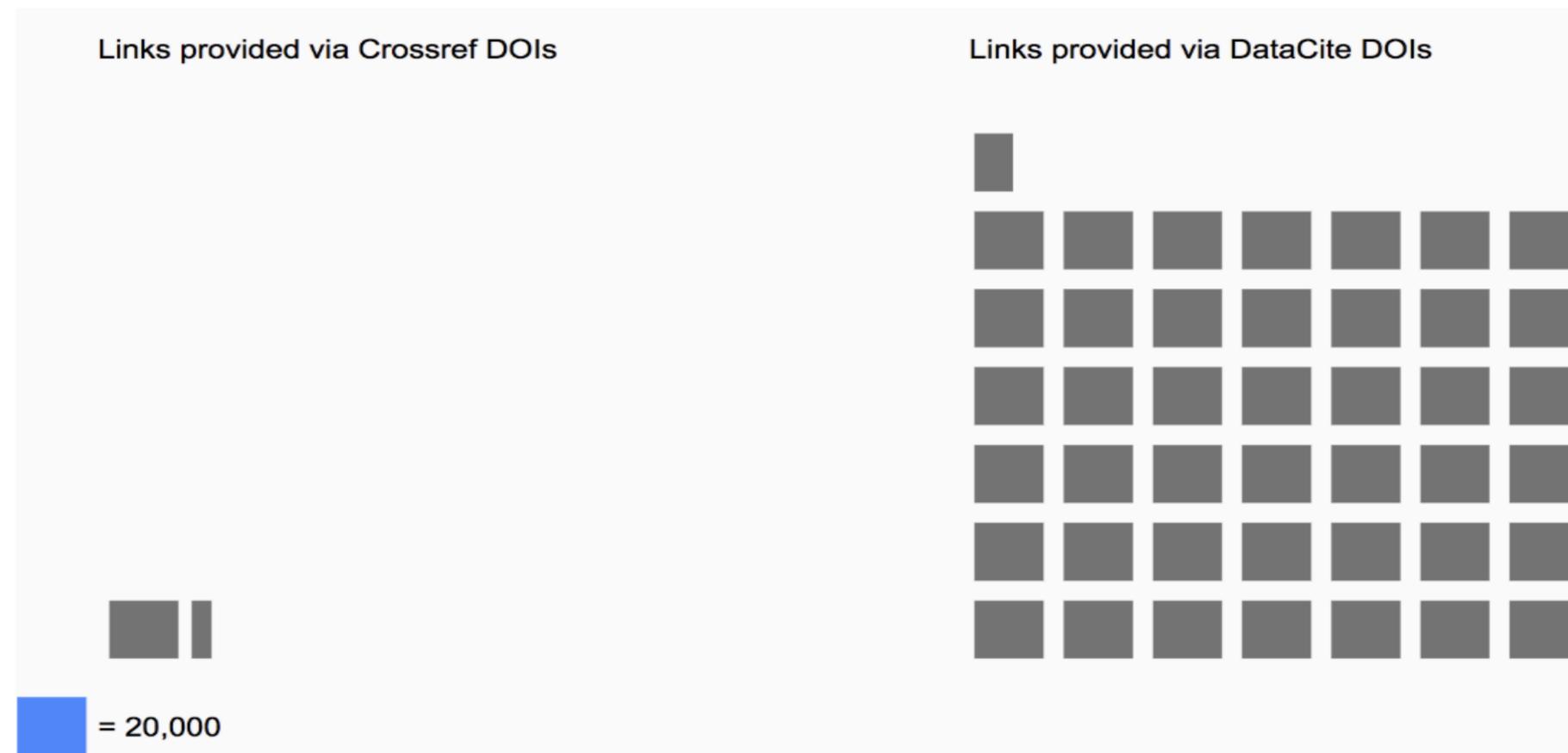
Future versions:

- details about where the data are being accessed
- volume of data being accessed
- citation details
- social media activity

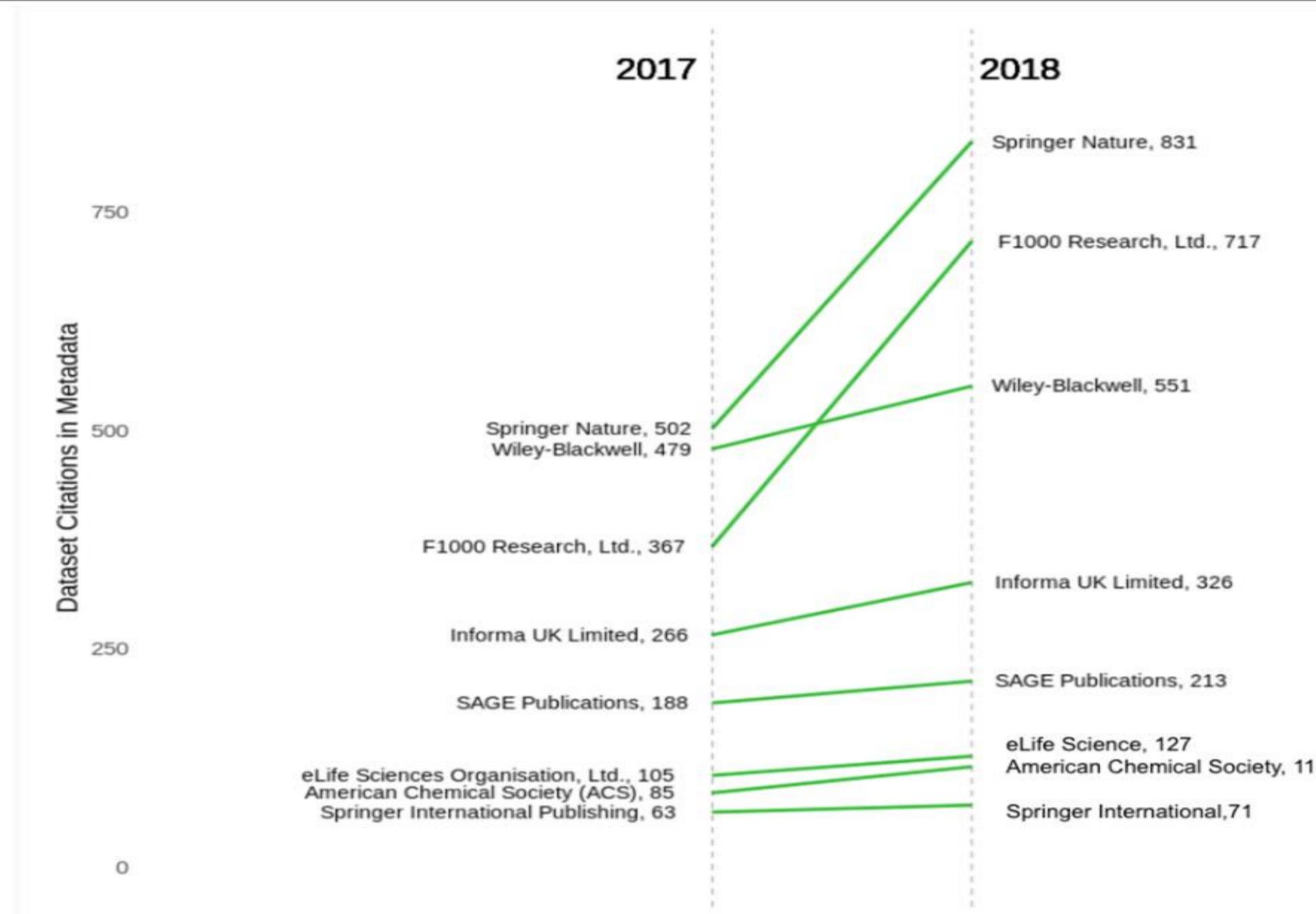
Lowenberg, Daniella; Budden, Amber; Cruse, Patricia (2018): It's Time to Make Your Data Count! <https://doi.org/10.5438/pre3-2f25>

Snapshot March 2018

- 870,000 links between Crossref DOIs and DataCite DOIs
 - ✓ 850,000 links originating from DataCite DOIs
 - ✓ 22,000 links originating from Crossref DOIs



There is improvement



Garza, K., Fenner, M. (2018). Glad You Asked: A Snapshot of the Current State of Data Citation.

<https://doi.org/10.5438/h16y-3d72>

FREYA (2017–2020)



- EU funded, 3 years, 5 million, 11 partners

Goals:

- ✓ Harmonizing PID resolution and metadata;
- ✓ improve discoverability of research outputs and associated researchers;
- ✓ expose links between research outputs, and between research outputs and associated researchers;
- ✓ integrate other identifiers
- ✓ Build a search interface for all DOI metadata
<https://www.project-freya.eu/en>

ORCID



PANGAEA.



Hindawi

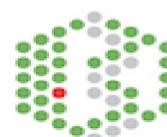


AUSTRALIAN NATIONAL DATA SERVICE



Science & Technology
Facilities Council

EMBL-EBI



Data Archiving and Networked Services
DANS



Zusammenfassung

- Viele Verlage/Zeitschriften haben keine Policy, wie Forschungsdaten zu referenzieren sind.
- Forschungsdaten werden hauptsächlich im Aufsatz/Artikel selbst zitiert und nicht in die Referenzliste eingetragen.
- Viele Forschungsdaten werden ohne DOI publiziert, sondern mit community-spezifischen Identifiern, wie z. B. Accession Nummern in den Lebenswissenschaften.
- Die Koordination der Publikation von Aufsatz/Artikel und Forschungsdaten ist schwierig, insbesondere wenn es parallel passieren muss.

Nächste Schritte

1. Metriken zu Forschungsdaten und die entsprechenden Zitierungen sollten auf der Landingpage der Forschungsdaten erscheinen.
2. Alternativen zum Sammeln der Forschungsdatenzitate, wie z.B. Text Mining nach DOIs sollten betrachtet werden.
3. Artikel-Forschungsdaten Publikationsworkflows müssen vereinfacht und standardisiert werden. ([Enabling FAIR Data project](#))
 - a) DataCite Integration mit metadata lookup in Publikations Tools, z.B. um Referenzlisten zu validieren, und
 - b) Koordination von DataCite und Crossref bei der DOI Registrierung am Anfang des Publikationsworkflows nach Manuskriptannahme, jedoch vor der Publikation.
4. Entwicklung eines Repository Auswahl Tools für die Forschenden auf Basis der re3data Informationen. ([Enabling FAIR Data project](#))

Garza, K., Fenner, M. (2018). Glad You Asked: A Snapshot of the Current State of Data Citation.

<https://doi.org/10.5438/h16y-3d72>

Was könnten Sie tun?

- Der [How-To Guide](#) von der California Digital Library unterstützt bei der Implementierung von Make Data Count.
- Tipps and Tools (e.g. a Python Log Processor) im How-To-Guide und Github <https://github.com/CDLUC3/Make-Data-Count>
- Die [DataCite documentation](#) enthält die Informationen zu den Event Data APIs
- Einladung zu einem Webinar: **How to implement Make Data Count at your repository and learn more about the project** ---- am Dienstag, den 10. Juli, um 17:00 Uhr ----
Webinar link: <http://bit.ly/2xJEA4n>

Kontakt



- Katja Pletsch
doi-service@zbmed.de
Tel.: 0221 478 7085
- Brigitte Hausstein
brigitte.hausstein@gesis.org
Tel.: 0221 476 94 596
- Britta Dreyer
doi-info@tib.eu
Tel.: 0511 762 17642