B2. Der Metadaten-Lebenszyklus

RDA-DE-Trainings-Workshop
09.06.2017 in Dresden
Heinrich Widmann (DKRZ)

Outline

- Metadaten (MD) : Was, Wozu und Wie ?
- ‚Best practices' : Die FAIR Prinzipien
- Lebenszyklus und Management von MD (und Daten)
  - MD Ingestion Worklflow
    - Standard-Verfahren und Technologien
    - + Übungen → B2FIND-Training
  - EUDAT-B2FIND : Ein interdisziplinärer MD Katalog
  - CMIP5 : Datamanagement für Klimadaten
- Schlußfolgerungen, Ausblick und Diskussion

Einführung zu Metadaten

- MD sind ‘Daten über Daten’
- MD sind ‘Strukturierte Informationen’, die eine ‘Informations-Ressource’ oder ein ‘Datenobjekt’ (DO)
  - beschreiben, erklären und lokalisieren
  - leichter benutzbar und verwaltbar machen
- MD stellen Informationen zur Verfügung, die Daten
  - Sinn geben,
  - mit Konzepten verbinden und
  - mit ‘real world’ Identitäten in Beziehung setzen
- Wir beschränken uns hier auf digitale Forschungsdaten

Daten, Metadaten, Datenschema -1-

Daten

5.1	0.2	Iris-setosa
7.0	1.4	Iris-versicolor
6.3	1.8	Iris-virginica
4.9	0.1	Iris-setosa

Metadaten

```
<metadata>
  <dc:creator>Fisher,
R.A.</dc:creator>
  <dc:title>The use of multiple
measurements in taxonomic
problems</dc:title>
  <dc:subject>Gen.
Statistics</dc:subject>
  <dc:date>1936-09-01</dc:date>
</metadata>
```

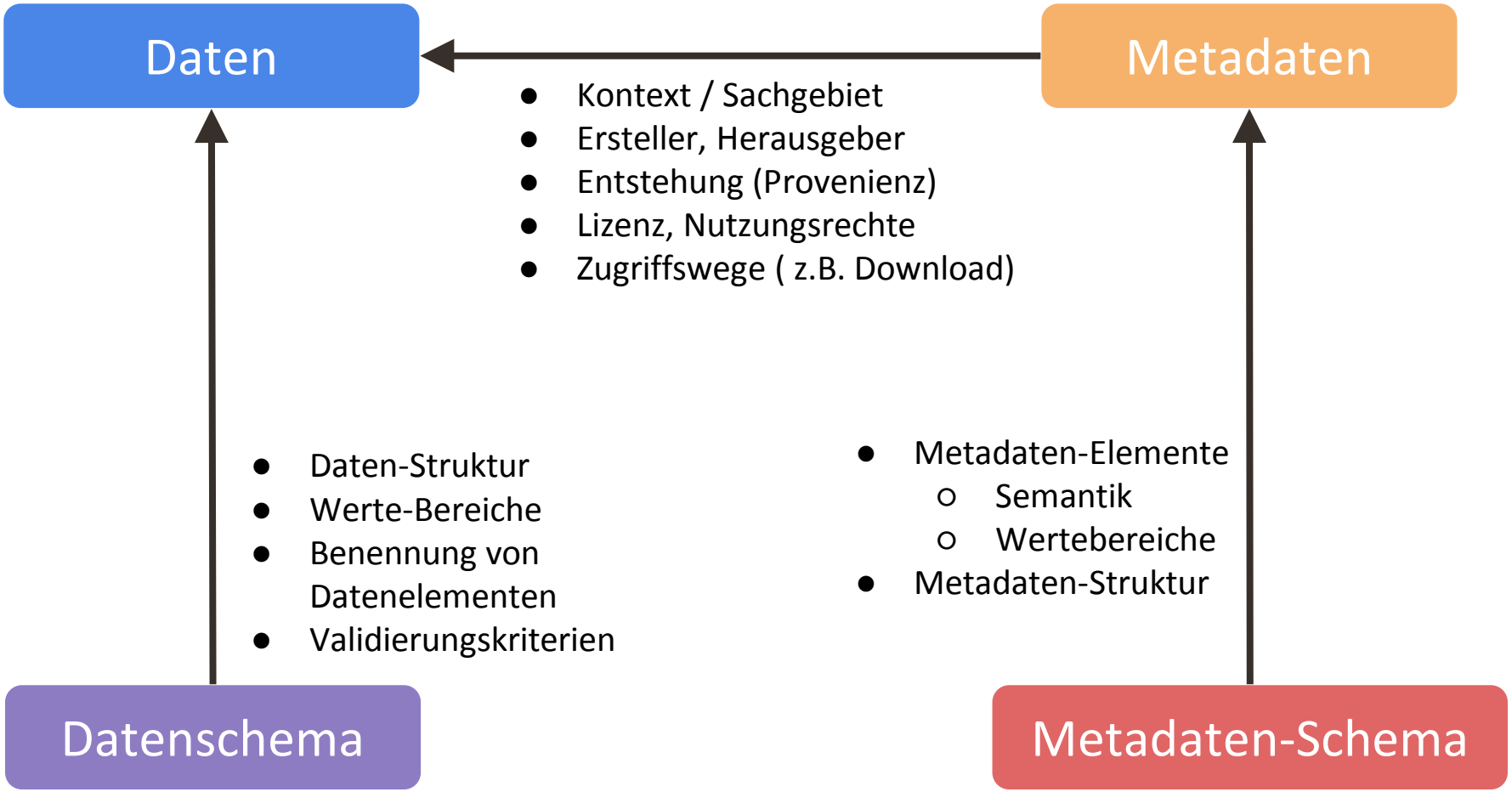
Datenschema

Kelchblatt-Länge [reelle Zahl; cm]	Kronblatt-Breite [reelle Zahl, cm]	Schwertlilien-Art [Text, 4 mögl. Werte]

Metadaten-Schema

dc:creator	An entity primarily responsible for making the resource. [xsd:string]
dc:subject	The topic of the resource. [xsd:string]
dc:date	A point or period of time associated ... [xsd:date]

Daten, Metadaten, Datenschema -2-



Wozu sind Metadaten gut ?

- Veröffentlichung und Verwaltung von Daten
  - **Auffinden/Entdecken** relevanter Daten (durch vereinheitlichte Beschreibungsverfahren)
  - **Abrufen/Beschaffen** der Daten (schnelle Lokalisierung durch verlässlichen Ressourcen-Indikatoren (URIs, URNs, PIDs, usw.)
- Förderung von
  - **Wiederverwendung** von Forschungsdaten
  - fachübergreifender **Interoperabilität**
- Validierung und Qualitätssicherung von Daten

Wie werden MD optimal nutzbar ?

Metadaten sollten

- einem klar definierten **Datenmanagementplan** folgen
- **Standards und Protokolle** erfüllen, die
  - weltweit anerkannt sind und
  - an das **Forschungsfeld angepasst** sind
- Regeln guter wissenschaftlicher Arbeit **(‘Best practices’)** folgen
  - einfach auffindbar (Idealfall: zentral indiziert)
  - unkompliziert und (sofern möglich) ohne Zugriffsbeschränkungen abrufbar

Die **FAIR** Prinzipien

- **F**indability : “Einfachheit, mit der Informationen gefunden werden können”
- **A**ccessibility : “Möglichkeit auf die referenzierten Ressourcen zuzugreifen”
- **I**nteroperability : “Datenaustausch zwischen mehreren Systemen mit unterschiedlichen Strukturen bei minimalem Informationsverlust"
- **R**euseability : “Möglichkeit der Weiter- und Wiederverwendung von Daten, die von anderen erzeugt wurden”

Grad der Interoperabilität -1-

- Es existieren viele **Forschungs- und Anwendungs-spezifische** Standards
- Diese haben innerhalb ihres Anwendungsbereichs ihre Berechtigung zur Erstellung von detaillierten MD-Sätzen
- Diese ‚Proliferation'
  - ist also oft notwendig für die Spezialisierung,
  - erschwert aber die fachübergreifende Kooperation
- Deshalb sollte bevorzugt die **Einbettung in ein gemeinsames, interdisziplinäres ‚Framework'** angestrebt werden, das **‚bestmögliche' Interoperabilität** erlaubt

Allgemeine Metadaten Schemata

Name	***Beschreibung***	***Anwendungs-/ Forschungsfeld***
Dublin Core	Einfacher, leicht verständlicher und sehr weit verbreiteter Metadatenstandards.	domänen-agnostisch, u.a. für den Austausch von MD benutzt (z.B. OAI-PMH)
DataCite	DOIs (= Digital Object Identifiers) als verbindlich vorgegebene Identifizierer	für veröffentlichte, zitierfähige Daten
DCAT	Data Catalog Vocabulary	für interoperable Datenkataloge im Web
PROV	Provenance-Model	Provenance-Modellierung

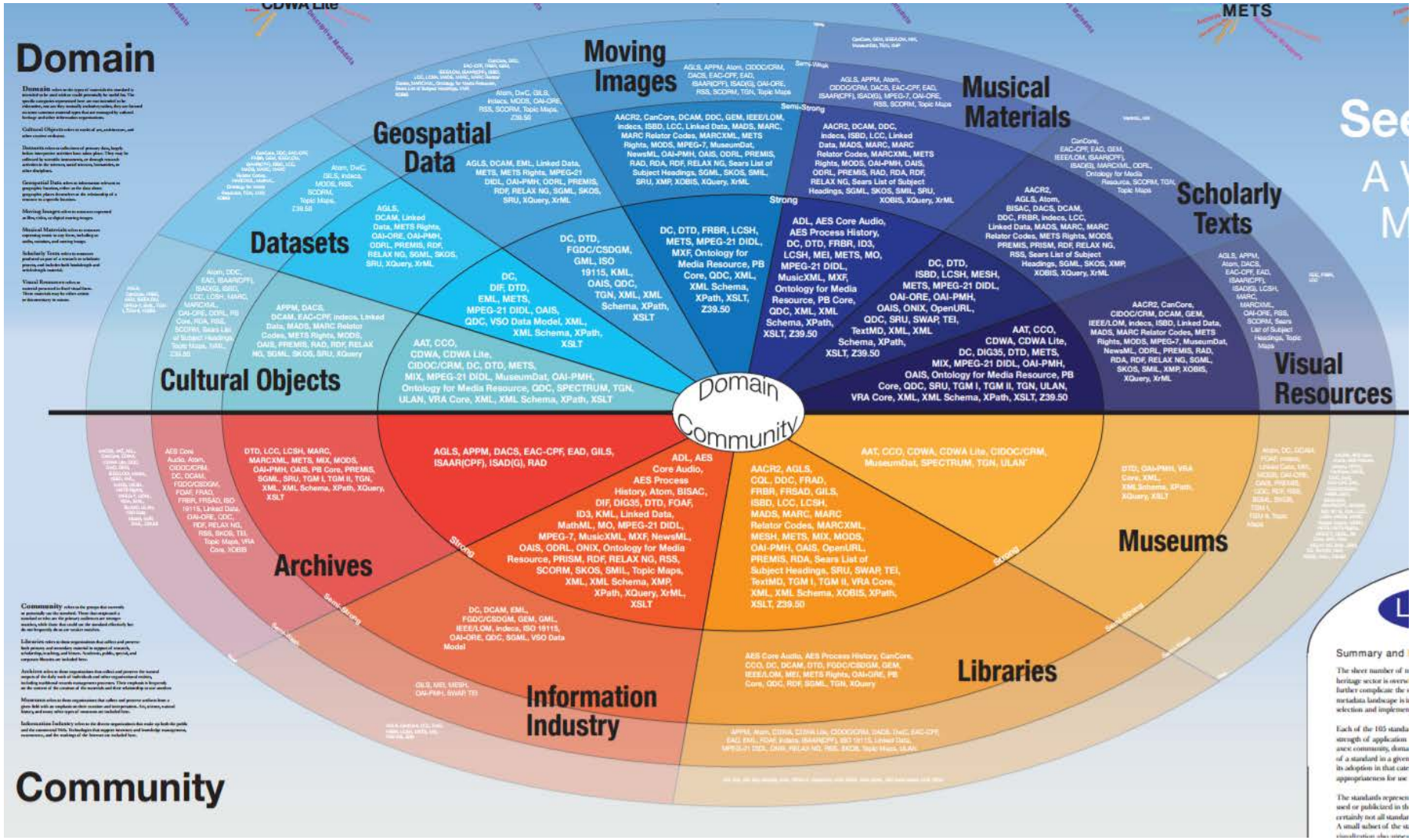
Disziplinäre Metadaten Schemata

Name	***Beschreibung***	***Anwendungs-/Forschungsfeld***
ISO19115 / ISO19139	für georeferenzierte Information (Spatial and Temporal Coverage)	Geowissenschaften
CMDI	Comonent Metadata Infrastructure Framework, mit der Metadaten-Profile beschrieben und wiederverwendet werden können	Linguistik und Sprachwissenschaften
DDI	Data Documentation Initiative : Offener Standard für die Beschreibung von sozial- und wirtschaftswissenschaftlichen Daten. Beschreibung des vollständigen 'Data Life Cycle` mittels XML.	Sozialwissenschaften

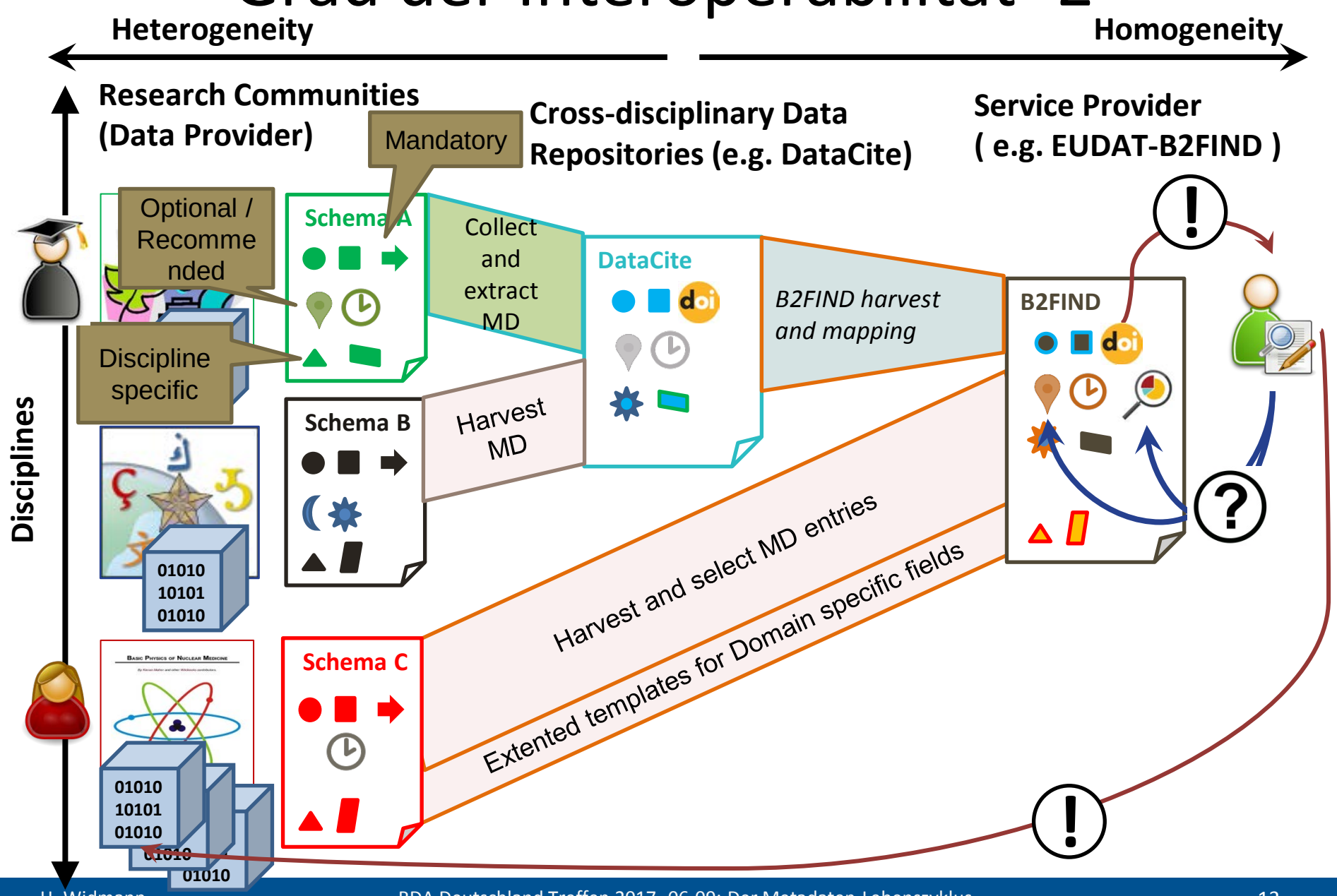
→ Metadata Universe

The Metadata Universe

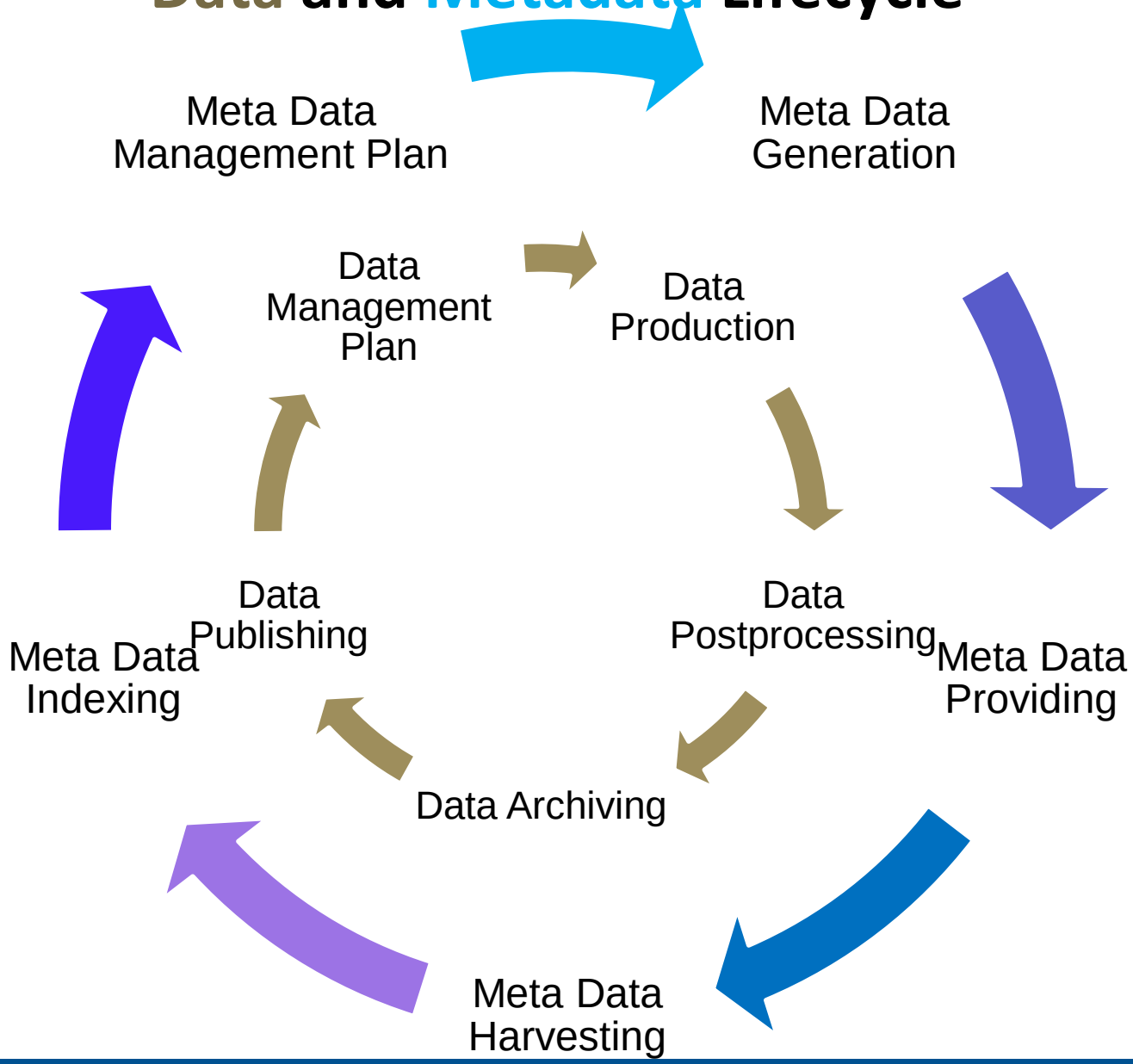
taken from Jenn Riley : http://jennriley.com/metadatamap/seeingstandards.pdf



Grad der Interoperabilität -2-



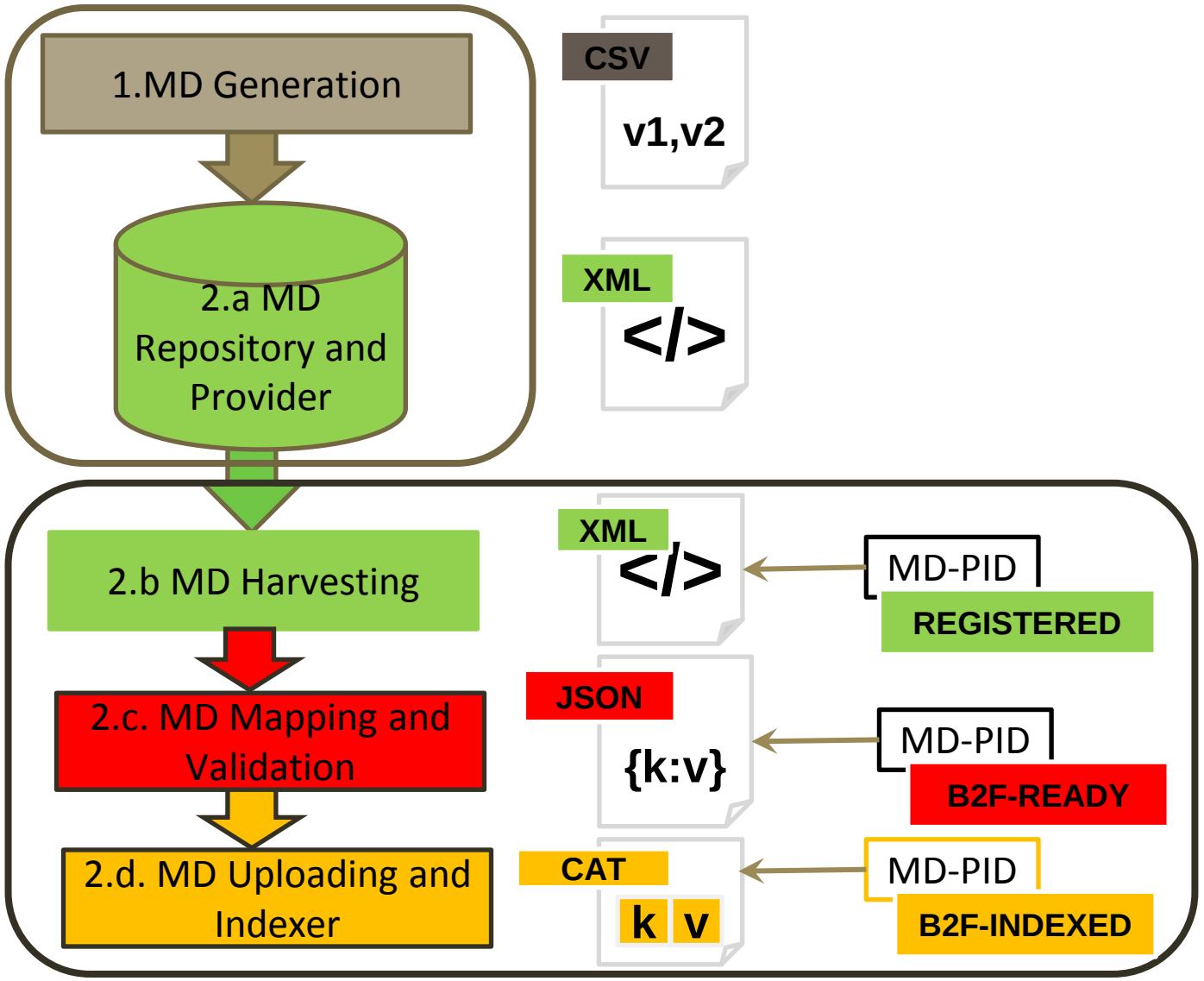
Data and Metadata Lifecycle



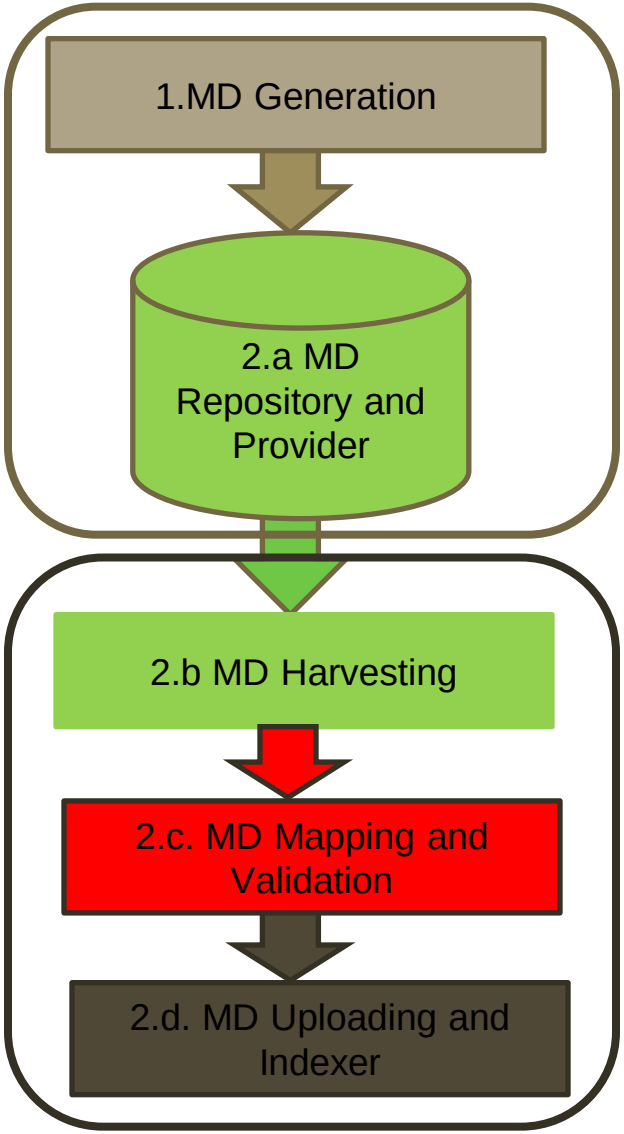
Der MD ‚Lebenszyklus' als Workflow

Data Provider : Community/ Discipline Specific Schema

Service Provider : Unified and Common Schema



MD workflow und B2FIND-Training Module



Modul 1
MD ‘Erzeugen’

Modul 2
MD Anbieten und ‘Harvesten’

Modul 3
MD ‘mappen’ und validieren

Modul 4
MD in Katalog und Portal hochladen

B2FIND-Training

- Das B2FIND-Training ist abrufbar unter
  - https://github.com/EUDAT-Training/B2FIND-Training
- Für einige Anwendungen braucht man eine Linux-Shell und muss sich den Sourcecode vom git repository herunterladen :
  - $ git clone https://github.com/EUDAT-Training/B2FIND-Training

    $ cd B2FIND-Training
- Neben den Übungen zum Metadaten-Workflow, gibt es auch Anleitungen zur Installation eines OAI- und CKAN-Servers, so dass man nicht nur (Meta)Daten-Provider und –Harvester sondern auch sein eigenen MD-Katalog und Suchportal (CKAN) aufsetzen kann.

1. MD Generation

- Dieser Prozess ist sehr spezifisch für jedes Forschungsgebiet und sollte einem Datenmanagementplan folgen
- Metadaten sollten bereits mit/vor der Datenproduktion generiert werden

- Das Ziel ist eine umfassende und eindeutige Datenbeschreibung
- Die Qualität der Metadaten profitiert von der frühen Kontrolle und Validierung
- Bereits hier sollten – soweit möglich - Standards eingehalten werden

1. Übung zu MD Generation

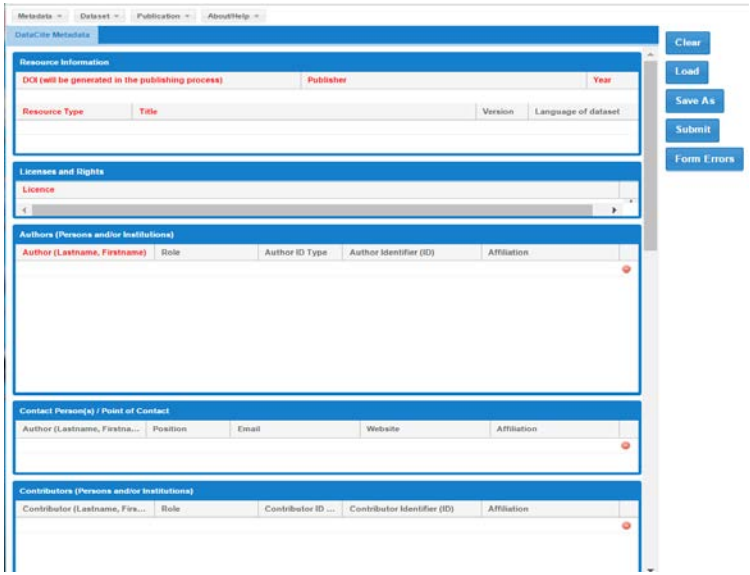
Erzeuge aus vorgegebenen ‚Roh-Metadaten' (Wertelisten, Tabellen, mitgebrachte Beispiele, ...) strukturierte und ‚valide' XML-Dateien im Metadatenformat ‚Dublincore'

a. Verwende den GFZ MD Editor (→ http://dataservices.gfz-potsdam.de/panmetaworks/metaedit/)um ein XML mit DOI zu erzeugen (nur speichern, nicht absenden !)
b. Benutze das script mdmanager.py im Modus ‚g' um aus den vohandenen Beispieldaten (kommasepearierte Liste) ein DublinCore XML zu erzeugen

1.a. Übung zu MD Generartion

- GFZ MD Editor –

Use the GFZ MD Editor
http://dataservices.gfz-potsdam.de/panmetaworks/metaedit
to create metadata for the first dataset in the fish-project sample table
→ https://github.com/EUDAT-Training/B2FIND-Training → samples/RAW_data/sample.csv



Common name	Scientific name	Location	Temperament	Diet	Water	Size	Region of the Aquarium	Breeding
Compressiceps	Haplochromis compressiceps	Lake Tanganyika	Territorial	Omnivore	PH 7.0 - 8.0, Temp. 73 - 77 F	5 inches	Bottom	Hard

taken from D.Ulbricht, K.Elger et al.
→ http://meetingorganizer.copernicus.org/EGU2017/EGU2017-12526-1.pdf

1.b. Übung zu MD Generartion

- ./mdmanager im Modus ‚g‘

Use the script mdmangager.py as described in → https://github.com/EUDAT-Training/B2FIND-Training → 01.b-generate-metadata.md
to create Dublincore XML files from the comma sepearated sample data in **samples/RAW_data/sample.csv**

```
$ ./mdmanager.py --mode g \
    -c fishproject -s samples/RAW_data/sample.csv
    --mdprefix oai_dc  --mdsubset sample --verb comma
$ ...
```

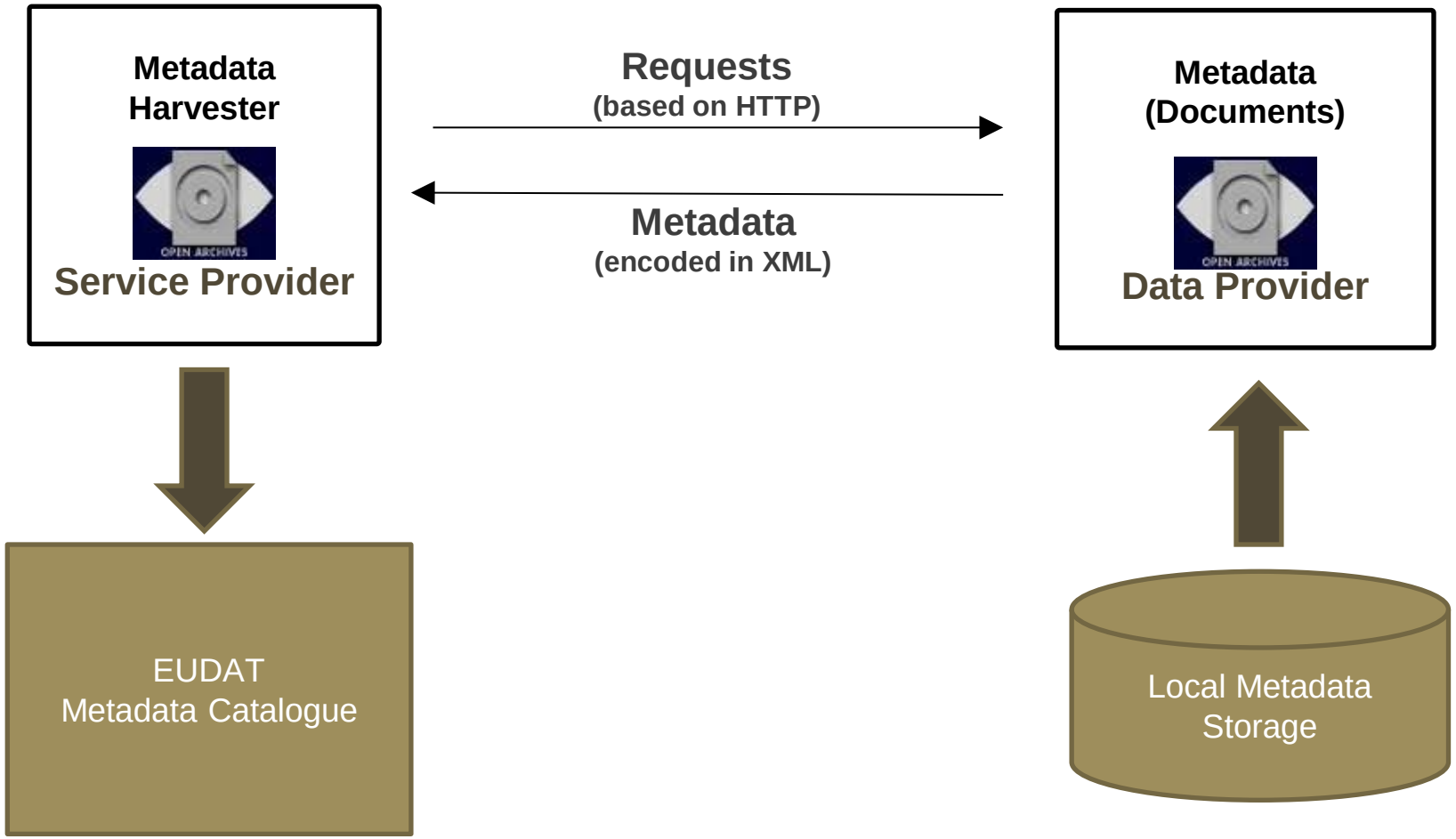
2. MD Providing and Harvesting

- MD ‘Provider’ und ‘Harvester’ müssen installiert werden
- Bevorzugt basierend auf dem Protokol ‘OAI-PMH’
- Der OAI-Provider wird auf Seite des (Meta)Daten-Produzenten aufgesetzt
- Erlaubt Service-Providern das Einsammeln (‘Harvesten’) der Metadaten von den Daten-Providern (communities)
- Daneben gibt es auch andere Scnittstellen (z.B. JSON-API, CSW)

OAI-PMH

- steht für **O**pen **A**rchives **I**nitiative **P**rotocol for **M**etadata **H**arvesting (→ http://www.openarchives.org)
- Ziel: Weltweite Konsilidierung von wissenschaftlichen Archiven
- Ermöglicht freien Zugriff auf Archive (zum. auf deren Metadaten)
- Ist ein einfacher (low-barrier) Mechanismus für die Interoperabilitaät zwischen Repositorien
- Besteht aus sechs ‘verbs’ oder ‘services’, die per HTTP aufgerufen werden
- Bietet konsitente Schnittstelle zwischen Daten- und Service-Anbieter
- Erlaubt leichte Implementation
- Basiert auf wenigen einfachen Protokollen und Standards (HTTP, XML, DublinCore)

Basic functioning of OAI-PMH



OAI benefits

- Interoperability : it is by no means domain specific and based on common metadata schemas
- Widely used : It's a quasi standard tool for providing metadata, for registered data providers (more than 2800 repostitories worldwide) see e.g. at https://www.openarchives.org/Register/BrowseSites
- Simple to install : In the B2FIND-Training we offer a guideline of the software joai. See the list of tools implemented by members of the Open Archives Initiative community at https://www.openarchives.org/pmh/tools/tools.php
- Simple to use : OAI attached great importance to simplicity of the protocol

OAI shortcomings

- Inefficiency : The XML serialisation and deserialisation takes time.
- Reference clash issue : if two records happen to have the same ID value, the envelope is not valid XML.
- Persistence of deletion : OAI-PMH allows three levels of persistence, but most providers promise none.
- Lack of SSL : By a strict reading OAI-PMH standard supports only http: , but not https

OAI-PMH Harvester – Verbs and parameters

Verbs that specify the service being invoked

- **Identify** - used to retrieve information about the repository.
- **ListIdentifiers** - used to retrieve record headers from the repository.
- **ListRecords** - used to harvest full records from the repository.
- **ListSets** - used to retrieve the set structure of the repository.
- **ListMetadataFormats** - lists available metadata formats
- **GetRecord** - used to retrieve an individual record from the repository.

Selective harvesting by parameters

- **identifier** - specifies a specific record identifier.
- **metadataPrefix** - specifies the metadata format of the returned records
- **set** - specifies the set that returned records must belong to.
- **from/until** – returns records created/update/deleted after/before this date
- **resumptionToken** - a token to resume a request where it last left off.

2. Übung zu MD Harvesting

Lassen Sie sich vom OAI-Provider von DataCite alle XML records des Subsets ANDS.CENTRE-1 anlisten, anzeigen bzw. herunterladen

a. Über http im Internet-Browser
b. mit dem Python scripts mdmanager.py im Modus h

2.a. Übung zu MD Harvesting

Harveste vom OAI-Provider von DataCite die MD im subset ANDS.CENTRE-1 per http **im Internet-Browser :**

1. Ermitteln Sie die URL des OAI-Servers von DataCite
2. Rufen Sie die Adresse im Browser auf
3. Wählen Sie als ‚verb' ListIdentifiers oder ListRecords, als set ‚ANDS.CENTRE-1' und als MD-Format Dublincore (mdprefix oai_dc)

Oder geben Sie den vollen http-Request ein :

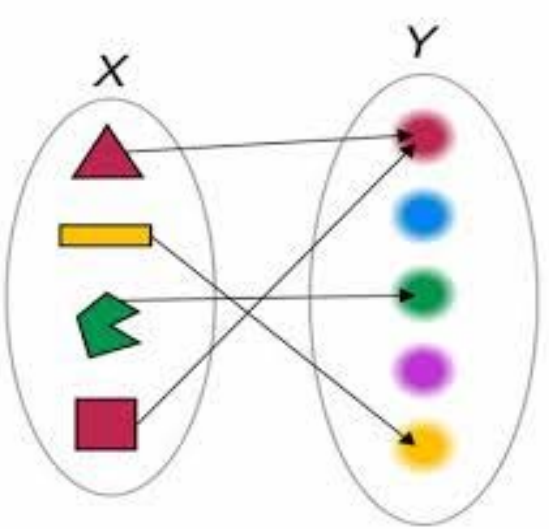
https://oai.datacite.org/oai?verb=ListRecords&metadataPrefix=oai_dc&set=ANDS.CENTRE-1

2.b. Übung zu MD Harvesting

Holen Sie von DataCite alle XML records des Subsets ANDS.CENTRE-1 mit Hilfe des scripts ./mdmangager.py im Modus h
(siehe → https://github.com/EUDAT-Training/B2FIND-Training →
02.b-OAI-harvester.md

```
$ ./mdmanager.py --mode h \
    -c datacite --mdsubset ANDS.CENTRE-1
$ ...
$ ls oaidata/datacite....
```

3a. MD Mapping



Die heterogenen, forschungsspezifischen MD metadata werden weiter prozessiert, homogenisiert und auf das ‘Zielschema’ abgebildet :

- Zerlege die XML –Datensätze und wähle Werte von den MD-Elementen durch spezifische Regeln aus
- Analysiere und ‘parse ‘ die Werte und ordne sie ‘key-value’ Paaren (JSON) zu
- Dabei werden ‘controlled vocabularies’ benützt

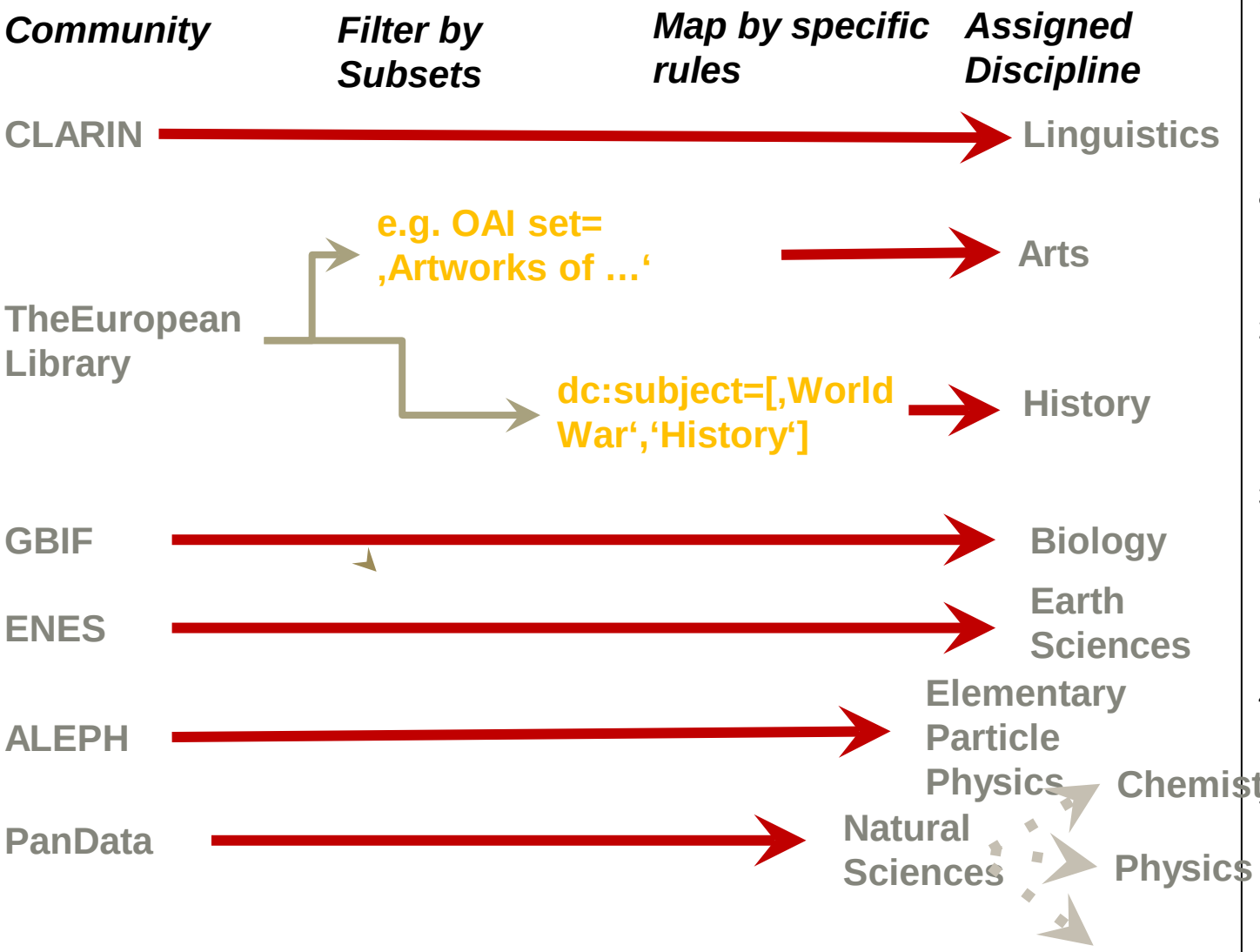
Letztendlich erhält man JSON-Datensätze, die die Spezifikation des B2FIND-Schemas erfüllen und in den B2FIND-Katalog hochgeladen werden können

Mapping and Validation Procedure

The community specific ‘raw’ metadata are processed in the following steps

1. **Select and convert** entries from the harvested XML records by MD format specific XPATH rules
2. **Semantic mapping** : Parse and analyse these lists of entries and map according controlled vocabularies to the B2FIND schema
3. **Validation :** Check and validate the resulting JSON records

Mapping of the Facet ‘Discipline’



B2FIND closed voca for ‚Discipline‘

1. Humanities
   - **1.1 History**
   - **1.2 Linguistics**
   - 1.3 Literature
   - **1.4 Arts**
     - 1.4.1 Performing arts
     - …
   - **1.5 Philosophy**
   - **1.6 Religion**
2. Social sciences
   - 2.1 Anthropology
   - **2.2 Archaeology**
   - ….
   - 2.7 Geography
3. Natural sciences
   - **3.1 Biology**
   - **3.2 Chemistry**
   - **3.3 Earth sciences**
   - **3.4 Physics**
   - …
4. Formal sciences
   - 4.1 Mathematics
   - 4.2 Computer sciences
5. Professions
   - **5.1 Agriculture**
   - ….
   - **5.6 Engineering**
     - 5.6.1 Chemical Eng.
   - 5.12 Library studies
   - 5.13 Medicine

3.b. MD Validation

- Examine each field for
  coverage, consistency and validity
- Semantic validation by using
  - controlled vocabularies
  - standard libraries, e.g.
    iso639 library for ‘Language’

- ‘Technical’ checks, e.g.:
  - Conformance of date-time fields with UTC format
  - Test spatial coverage by geonames.org and consistency of lat/lon coordinates
  - Off- and online checks of URL’s to the data objects (‘Source’, ‘PID’ and ‘DOI’)

3. Übung zu Mapping und Validierung

Prozessiere die von DataCite im subset ANDS.CENTRE-1 geharvesteten Metadaten

a. Mapping : Erzeuge aus den Dublincore XML-Dateien nach JSON-Dateien im B2FIND-Schema mit mdmanager.py im Modus m
b. Validierung : Checke die erzeugten JSON-Dateien mit mdmanager.py im Modus v

3.a. Übung zu MD Mapping

Erzeuge aus den Dublincore XML-Dateien nach JSON-Dateien im B2FIND-Schema mit mdmanager.py im Modus m
as described in → https://github.com/EUDAT-Training/B2FIND-Training →
03.a-map-metadata.md

```
$ ./mdmanager.py --mode m \
    -c datacite --mdsubset ANDS.CENTRE-1
$ ...
$ ls oaidata/datacite-oai_dc/ANDS.CENTRE-1/json
```

3.b. Übung zu MD Validating

Checke die erzeugten JSON-Dateien mit mdmanager.py im Modus v as described in → https://github.com/EUDAT-Training/B2FIND-Training →

03.b-validate-metadata.md

```
$ ./mdmanager.py --mode v \
    -c datacite --mdsubset ANDS.CENTRE-1
$ ...
$ less oaidata/datacite-oai_dc/ANDS.CENTRE-1/validation.stat
```

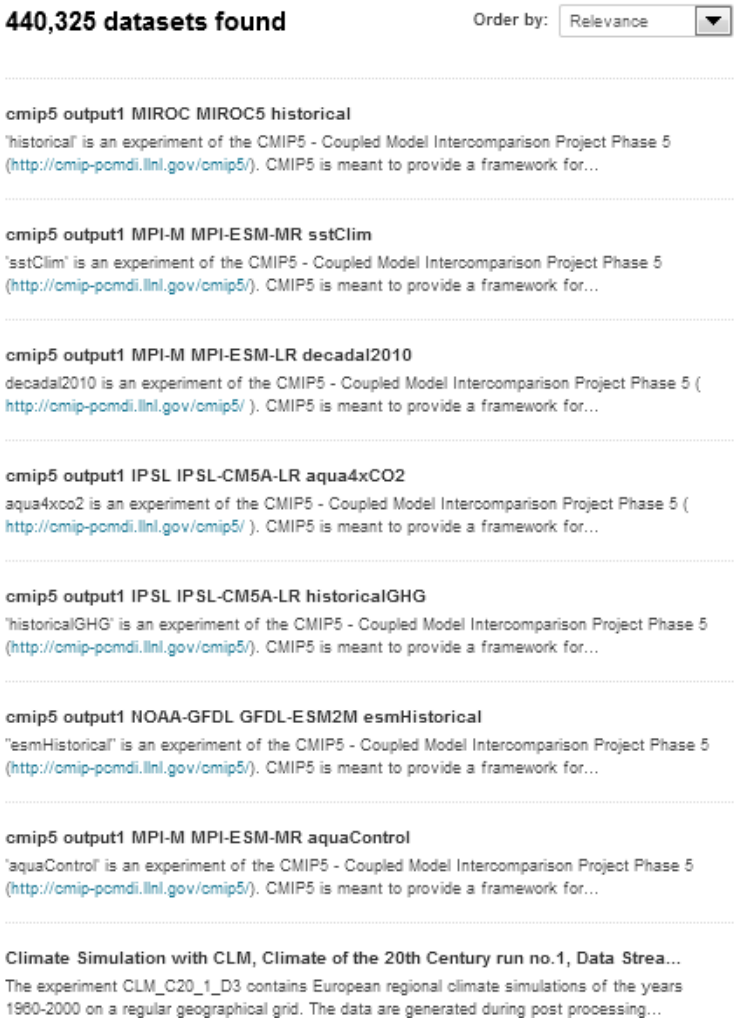
4. MD Uploading

Finally the checked and mapped JSON records are uploaded as datasets to the MD catalogue, which is based on the open source code CKAN.

CKAN

- provides a rich RESTful JSON API and
- uses SOLR for dataset indexing

That enables users to query and search in the catalogue



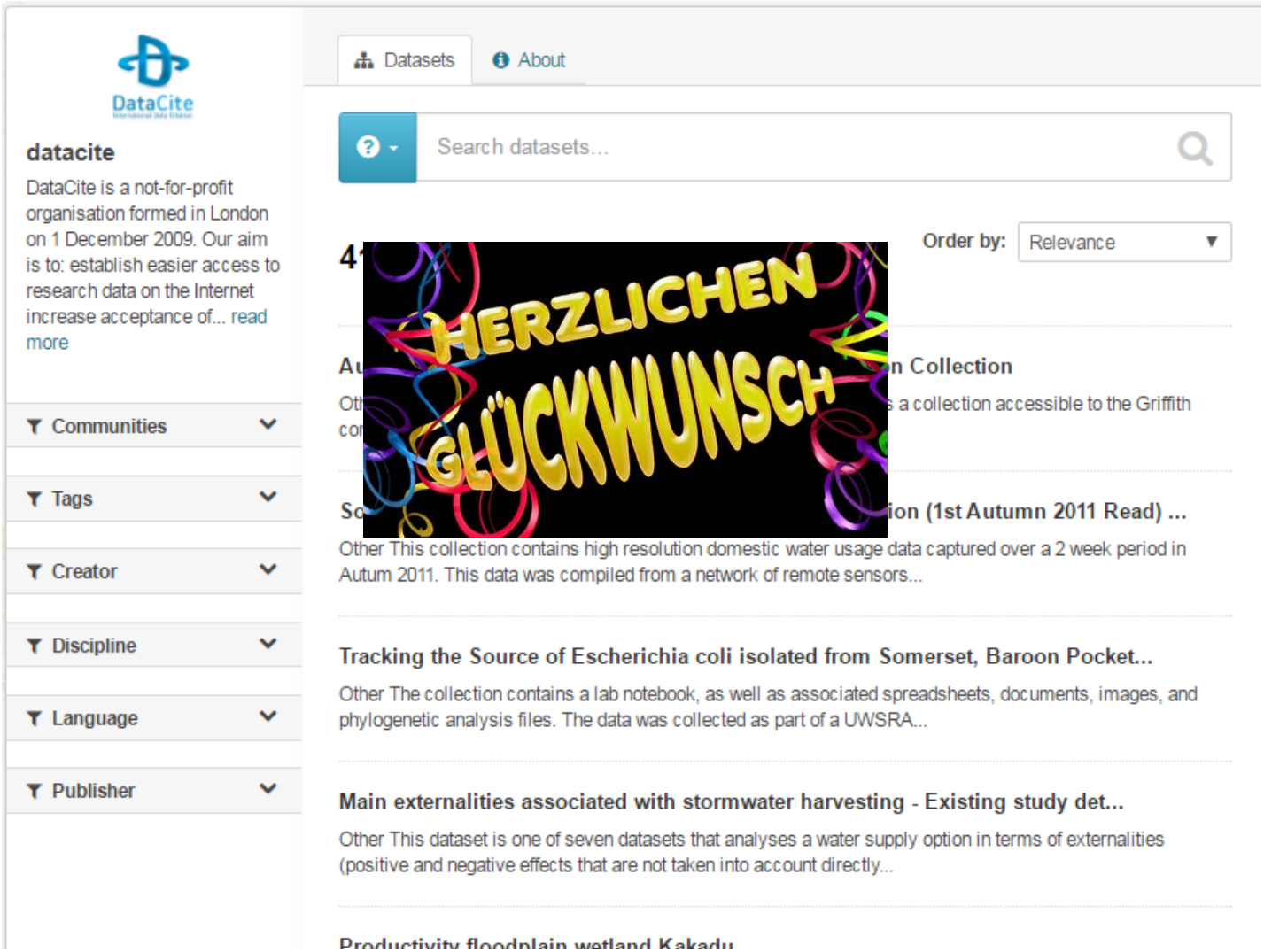
4. Übung zu MD Uploading

Lade die von DataCite stammenden JSON-Dateien im B2FIND-Schema in die B2FIND Trainingsinstanz hoch mit mdmanager.py im Modus m as described in → https://github.com/EUDAT-Training/B2FIND-Training →
04.b-upload-metadata.md

```
$ ./mdmanager.py --mode u \
    -c datacite --mdsubset ANDS.CENTRE-1 /
    –i trng-b2find.eudat.eu
$ ...
```

check http://trng-b2find.eudat.eu/group/datacite ....

Wenn Sie unter http://trng-b2find.dkrz.de/group/datacite die Seite unten angezeigt bekommen, haben Sie es geschafft !!



Fragen und Diskussion zu MD Lifecycle

- Inwieweit entspricht das präsentierte Vorgehen den Arbeitsabläufen in Ihrem Arbeitsumfeld ?
- Welche zusätzlichen Probleme oder Anforderungen haben Sie ?
- Welche Verfahren und Tools verwenden Sie und/oder würden Sie gerne verwenden ?
- Sind on-line Trainings wie das B2FIND-Training nützlich ?

EUDAT-B2FIND
Als Beispiel für interdisziplinäres Metadatenportal

EUDAT -1-

- The project European Data Infrastructure (EUDAT)
  - started in 2011 as an EU funded project
  - is now in 2nd phase EUDAT2020 (Horizon2020 program)
  - will end as development project will end in 03/2018
  - from 2018 on : proposal of EUDAT-EGI-Indigo consortium for H2020 e-Infra-12-2017 call

EUDAT -2-

- **Motivation** : Manage the rising tide of research data
- **Challenge** : Improve Interoperability in a wide cross-disciplinary scope
- **Objective** : Build up a Collaborate Data Infrastructure,
  - following the FAIR principles
  - based on common and generic data services
  - driven by requirements of the research communities

EUDAT B2 Service Suite

→ http://www.eudat.eu/services

Data exchange

Store small-scale data

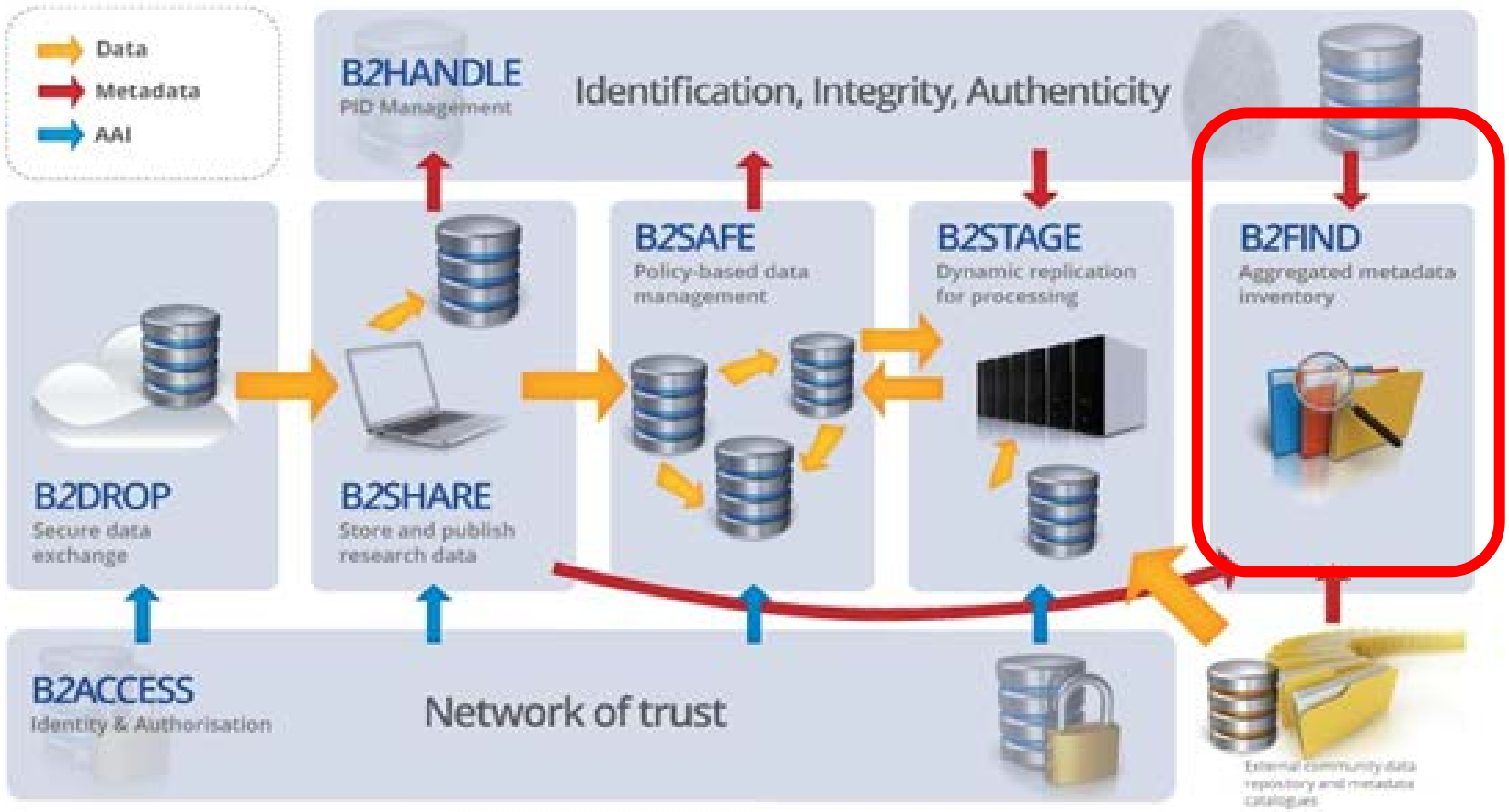
Store large-scale data

Transfer data to HPC

Data discovery

EUDAT Collaborative Data Infrastructre (CDI)

→ https://eudat.eu/eudat-cdi



EUDAT-B2FIND -1-

- **Motivation** : Provide a central discovery service overarching a wide scope of research disciplines
- **Challenge** : Homogenisation and mapping of the various and heterogeneous metadata on a common scheme
- **Objective** : Build up an easy-to-use discovery service
  - ➢ based on a comprehensive joint metadata catalogue
  - ➢ allowing search over cross-disciplinary data

EUDAT-B2FIND -2-

B2FIND is EUDAT's Metadata Service and consists of

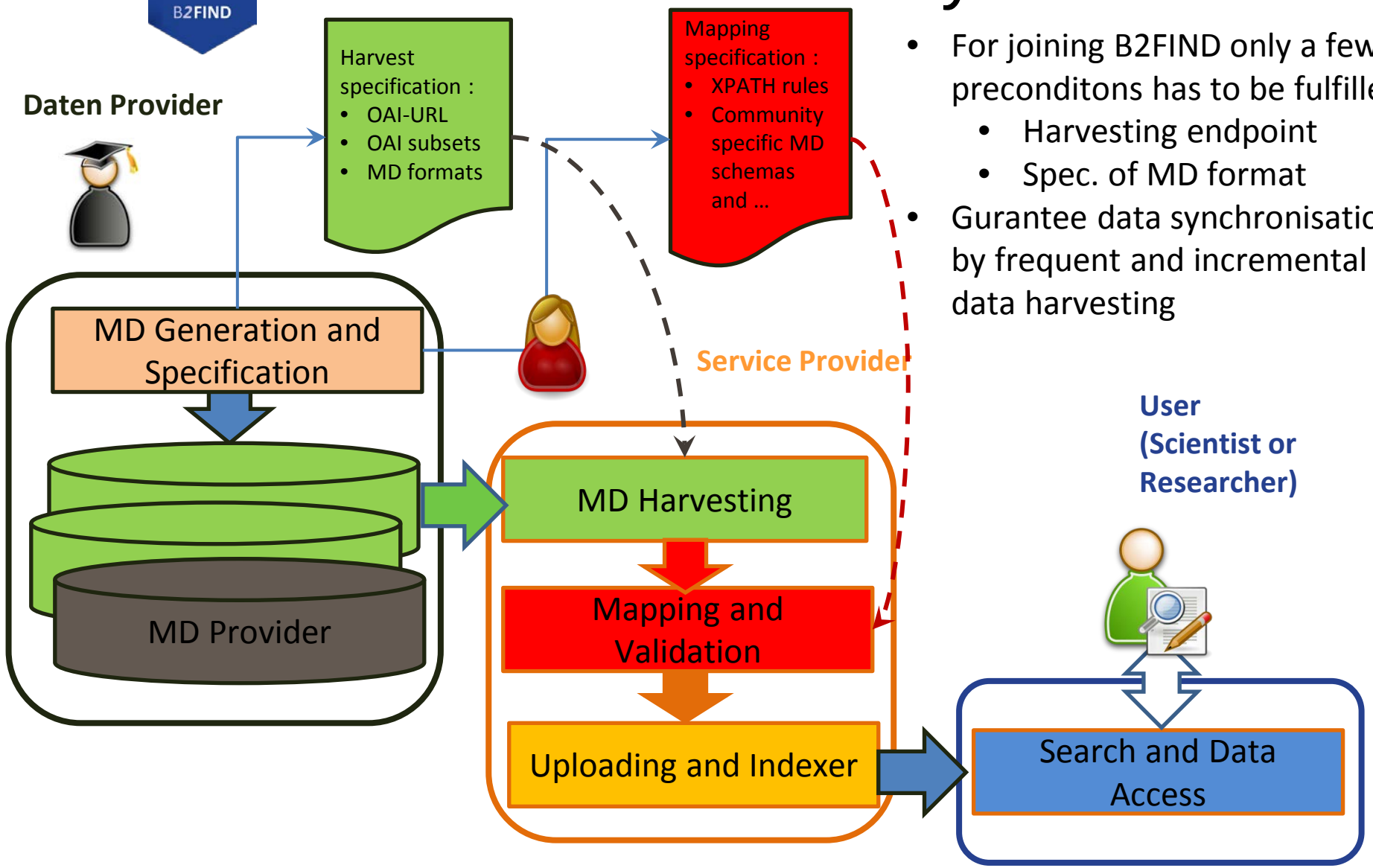
- a comprehensive metadata catalogue that spans a large number of multi formatted datasets
  - harvested from various and heterogeneous sources
  - covering a wide range of highly diverse disciplines
  - mapped to a **common schema (based on DataCite 3.0)**
- an open search portal allowing researchers
  - to find easily collections of scientific resources using standardized facets (e.g. disciplines and coverage)
  - to access data collections through given identifiers

B2FIND is based on the open source software CKAN

Wie die FAIR Prinzipien in EUDAT- B2FIND umgesetzt werden

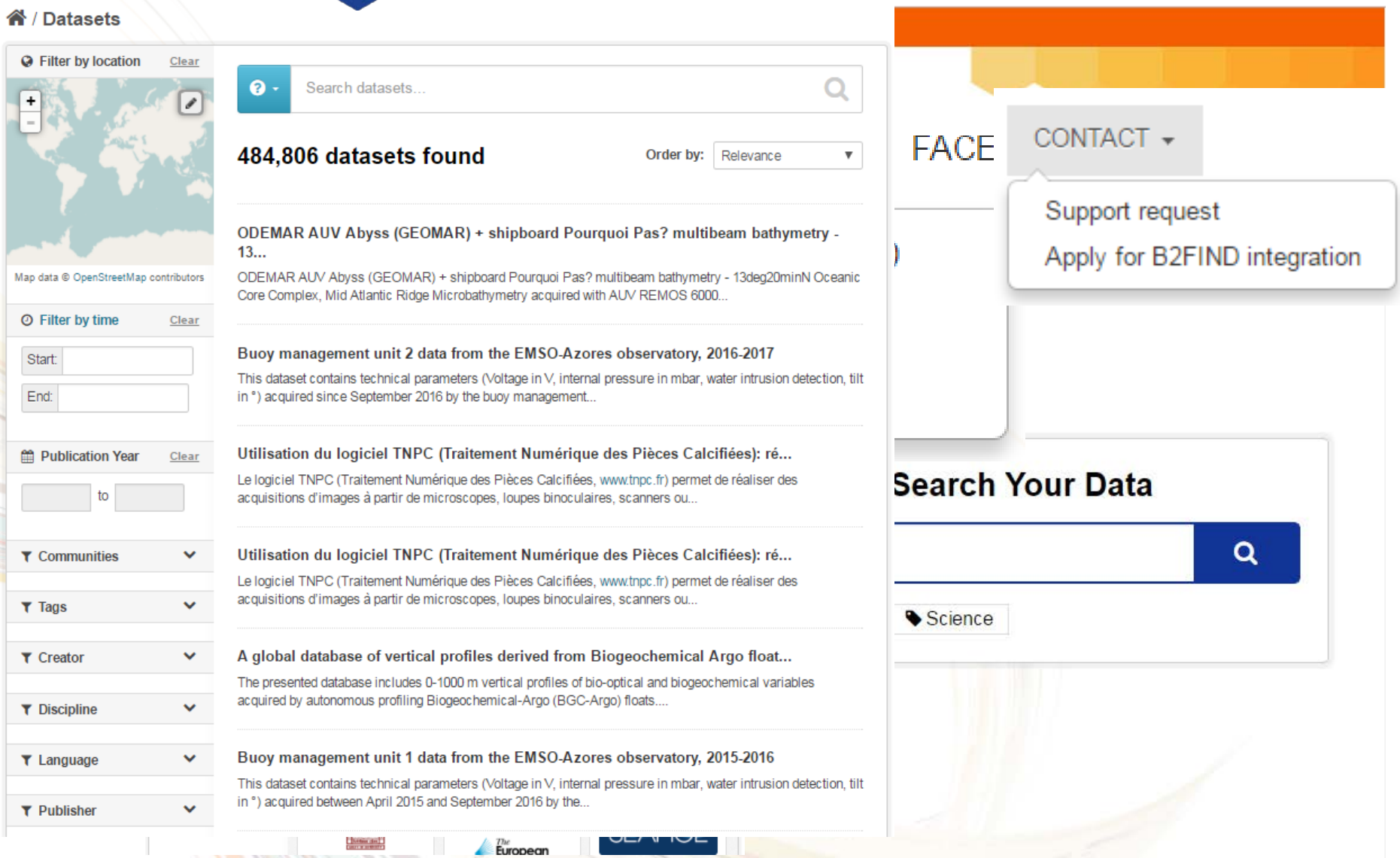
- **F**indability
  - Easy-to-use **Discovery Portal** with powerful search features
- **A**ccessibility
  - Offer **Persistent Identifiers** to resolve and access data objects by humans and machines using standard protocols
- **I**nteroperability
  - **Cross-disciplinary MD catalogue** based on Common Standards and Schema
- **R**euseability
  - Provide information on **data access rights and provenenance**

B2FIND MD Lebenszyklus



- For joining B2FIND only a few preconditons has to be fulfilled
  - Harvesting endpoint
  - Spec. of MD format
- Gurantee data synchronisation by frequent and incremental data harvesting

B2FIND Website



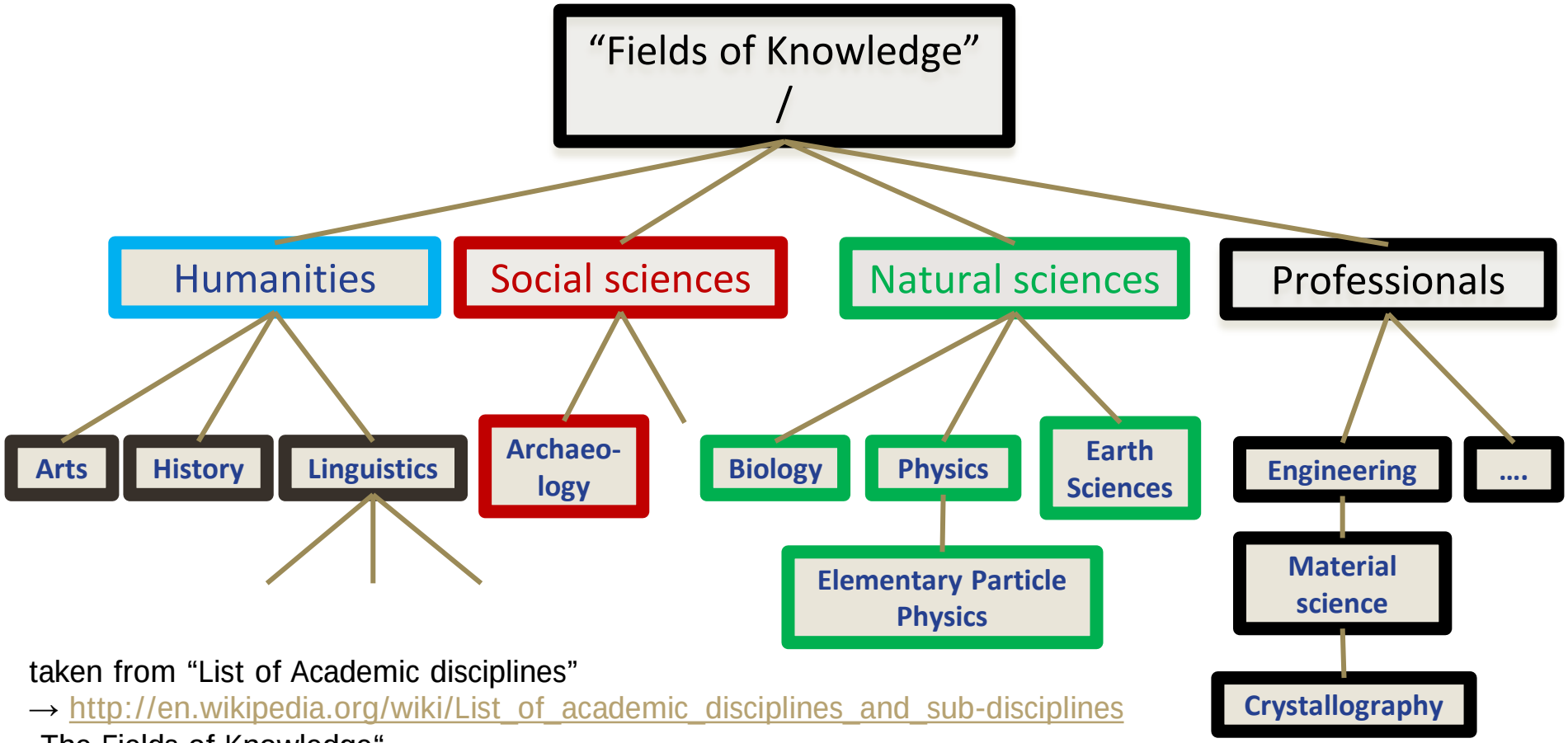
B2FIND MD Schema (extract)

based on DataCite 3.0

Metadata Type	B2FIND Field name	Allowed values	Semantic definition	Level of Obligation	Occurence	
General information	**Title**	Free text (unicode)	A name or title a resource is known	**Mandatory**	1	
	Description	Free text	Additional info	Recommended	0-1	
Data Access	**Source**	Valid URL or URN	Unique link to data resource	**Mandatory (1)**	0-1	1-3
	PID	Persistent Identifier	+ persistent and resolvable		0-1	
	DOI	Digital Object Identifier	+ citable		0-1	
Provenance data	**Creator**	';'-sep. list of names	Main researchers involved in data prod.	Recommended	0-n	
	Discipline	List of values from CV	Field of research (Controlled Vocab)	Recommended	0-n	
	Publication Year	YYYY	The year data are published	Recommended	1	
Formal data	**Temporal Coverage**	Interval of 2 DTimes [Begin, End]	The temporal limits of a date-time	Optional	1-n	
	Spatial	Spatial box or point	The spatial limits of a	Optional	1-n	

‘Disziplinen’ in B2FIND

as CV → https://github.com/EUDAT-B2FIND/md-ingestion/blob/master/mapfiles/b2find_disciplines.tab

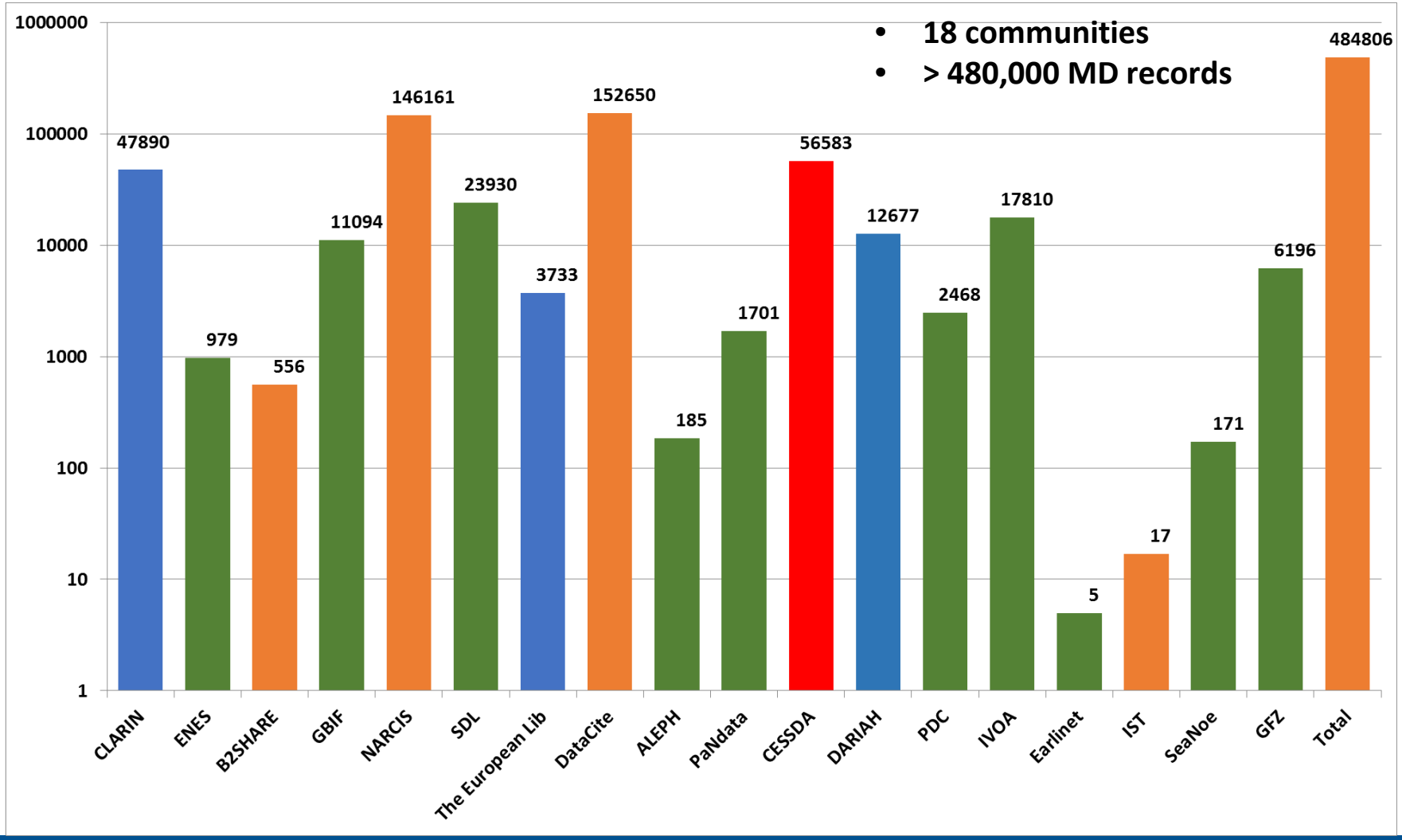


taken from “List of Academic disciplines”
→ http://en.wikipedia.org/wiki/List_of_academic_disciplines_and_sub-disciplines
„The Fields of Knowledge“
→ http://www.thingsmadethinkable.com/item/fields_of_knowledge.php?focus=natural_sciences
„Branches of Science“
→ https://en.wikipedia.org/wiki/Branches_of_science

B2FIND MD Catalogue

- Ingestion Status -

Humanities
Social Sciences
Natural Sciences
Cross Discipline



B2FIND Discovery Portal

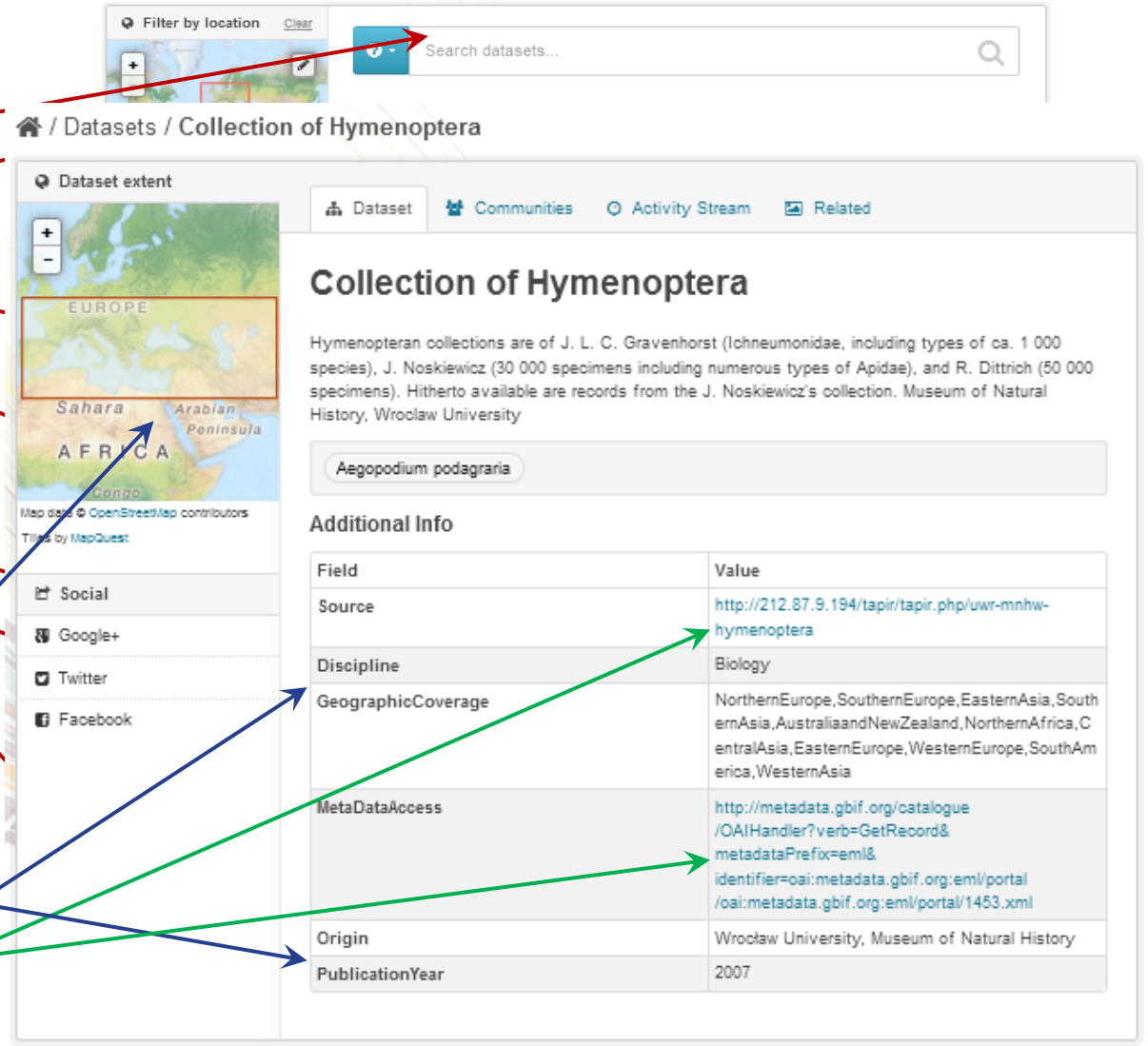
→ http://b2find.eudat.eu

B2FIND provides 'faceted' search for

- Free text
- Geo spatial
- Temporal coverage
- Publication year
- Textual facets as
  - Tags
  - Creator
  - Discipline etc.

Dataset view provides display of metadata :

- Spatial extent
- Table of field-value pairs
- Links to data resources

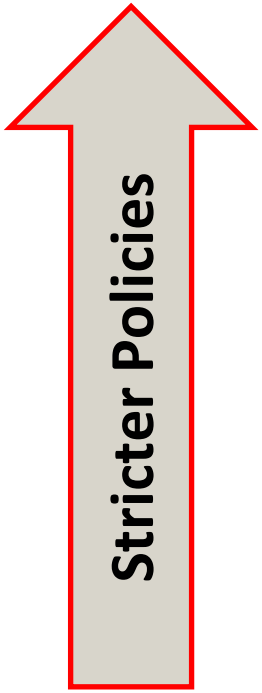


Data Access in B2FIND

- B2FIND provides different identifiers (e.g. PID, DOI and/or other URL) that refer to the underlying data object
- Each reference may
  - ‘resolve’ the data object (direct view or download of the data object)
  - (in most cases) lead to a landing page or another metadata view
- This ‘level of resolvement’ depends on the access policies and the data structure of the data provider
  - Often log-in is required to access and download the data files
  - Landing pages can also be used to refer to a collection or aggregation of datasets (e.g. CERA model)

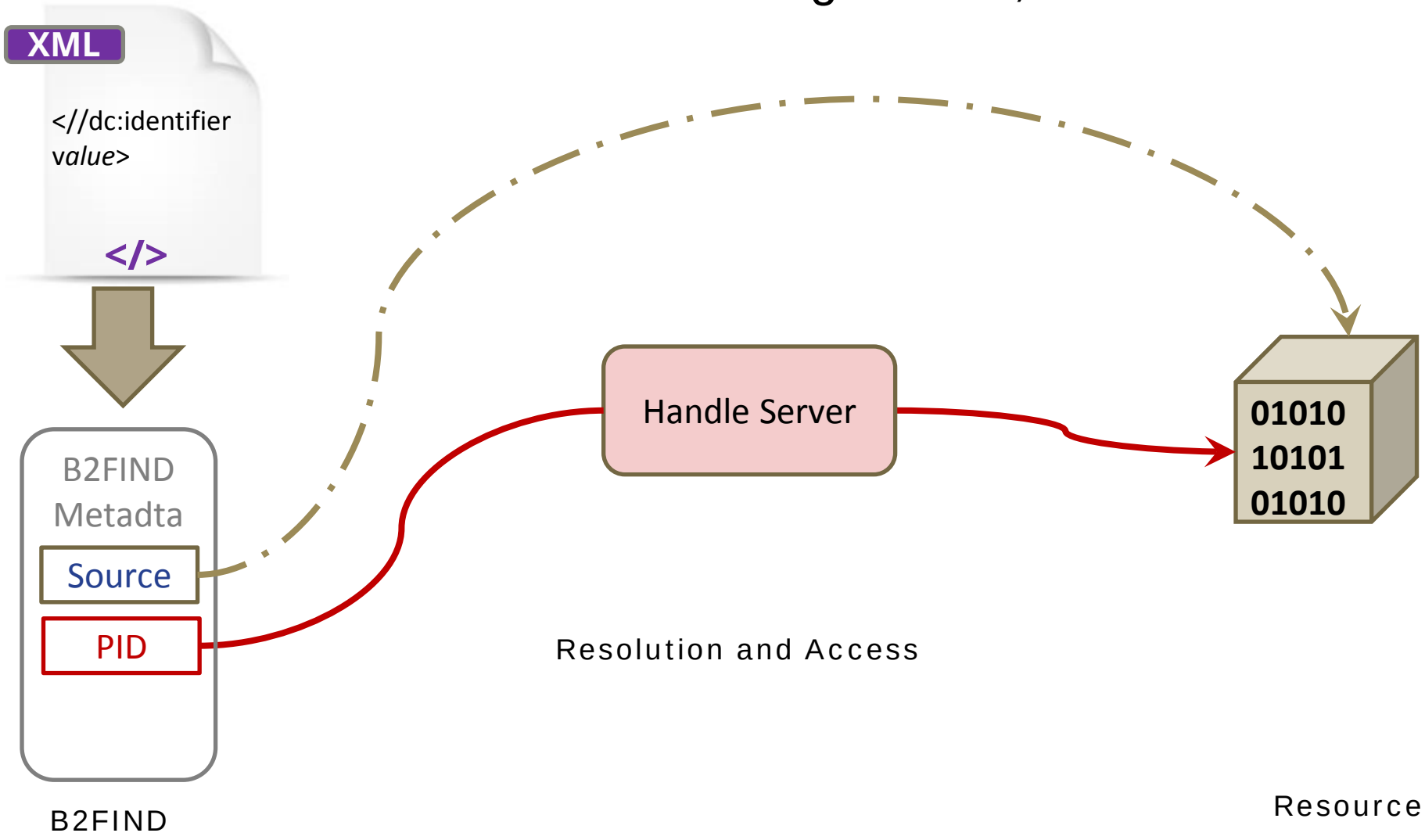
Data Identifiers

Type	Unique	Persistent	Resolvable	Citable
DOI	✓	✓	✓	✓
PID	✓	✓	✓	x
URL (Source)	✓	?	?	x



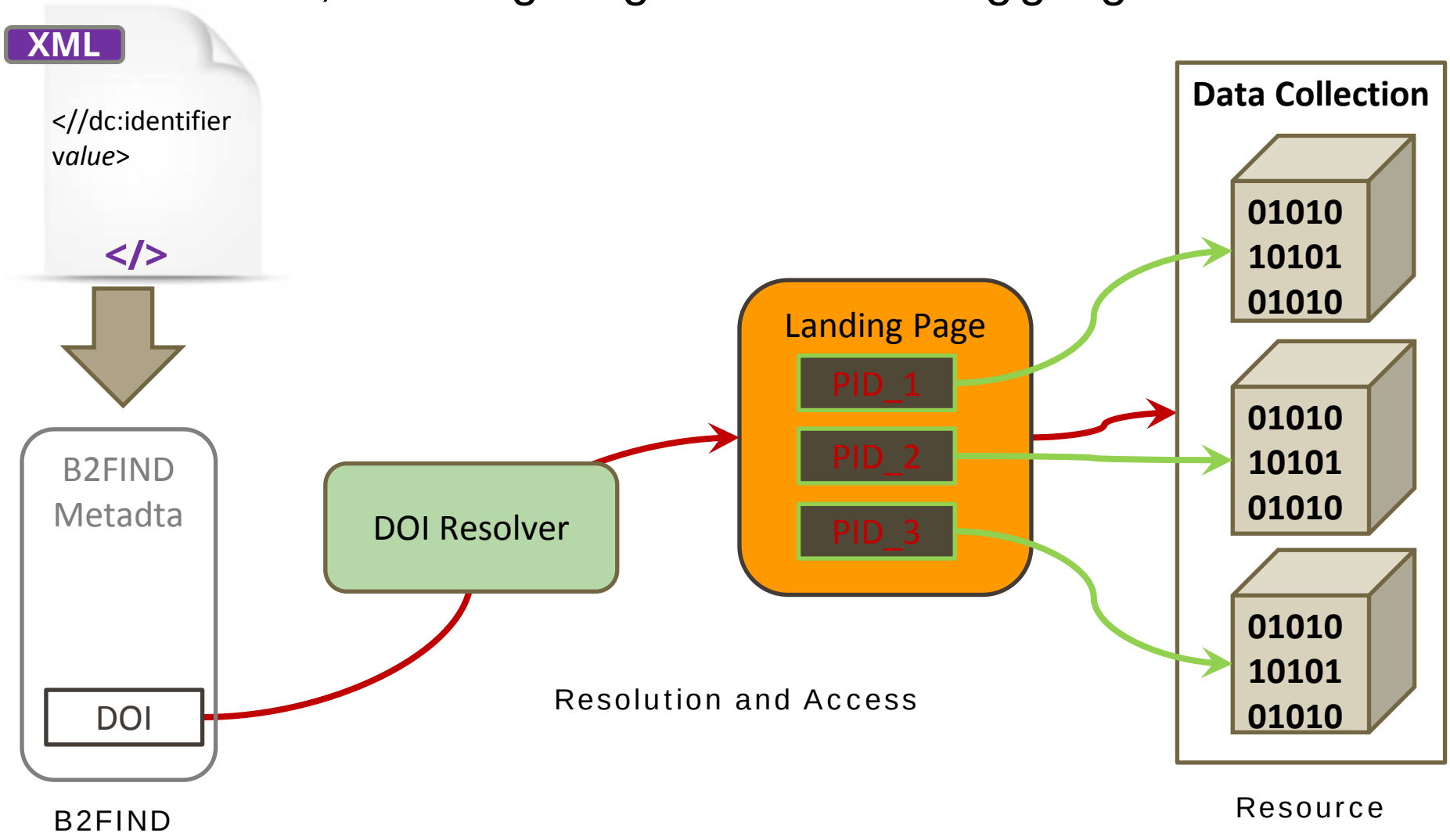
Datenzugriff

Persistente Identifizierung durch ‚Handles'

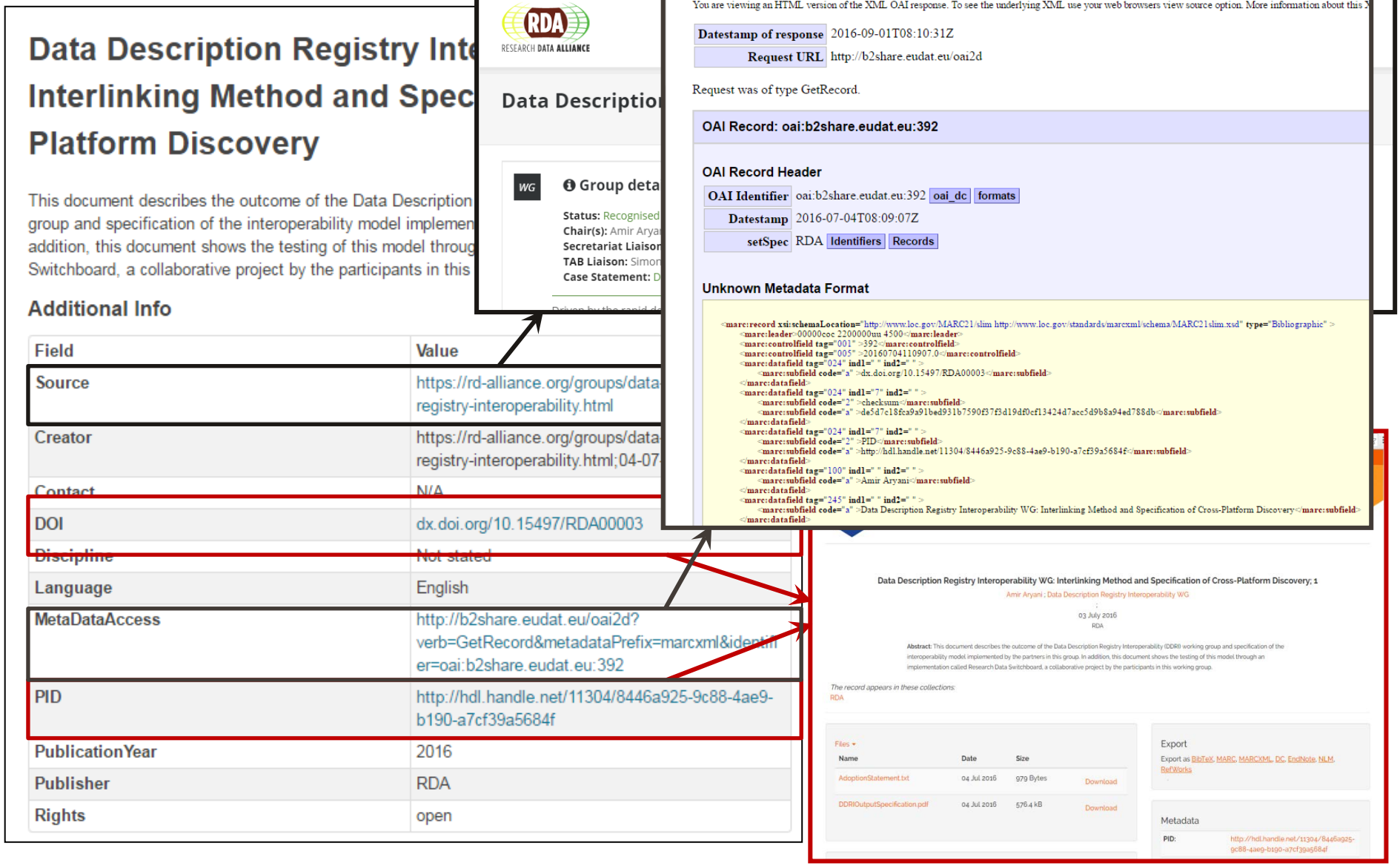


Datenzugriff

DOI, Landing Page und Data Aggregation



Beispiel 1



Data Description Registry Inte
Interlinking Method and Spec
Platform Discovery

This document describes the outcome of the Data Description
group and specification of the interoperability model implemen
addition, this document shows the testing of this model throug
Switchboard, a collaborative project by the participants in this

Additional Info

Field	Value
Source	https://rd-alliance.org/groups/data registry-interoperability.html
Creator	https://rd-alliance.org/groups/data registry-interoperability.html;04-07
Contact	N/A
DOI	dx.doi.org/10.15497/RDA00003
Discipline	Not stated
Language	English
MetaDataAccess	http://b2share.eudat.eu/oai2d?verb=GetRecord&metadataPrefix=marcxml&identifier=oai:b2share.eudat.eu:392
PID	http://hdl.handle.net/11304/8446a925-9c88-4ae9-b190-a7cf39a5684f
PublicationYear	2016
Publisher	RDA
Rights	open

https://rd-allianc

RESEARCH DATA SHARING WITHOUT

RESEARCH DATA ALLIANCE

Data Descriptio

WG

Group deta

Status: Recognised
Chair(s): Amir Arya
Secretariat Liaison
TAB Liaison: Simon
Case Statement:

https://b2share.eudat.eu/oai2d?verb=GetRecord&metadataPrefix=marcxml&identifier=oai:b2share.eudat

OAI 2.0 Request Results

Identify | ListRecords | ListSets | ListMetadataFormats | ListIdentifiers

You are viewing an HTML version of the XML OAI response. To see the underlying XML use your web browsers view source option. More information about this

Datestamp of response	2016-09-01T08:10:31Z
Request URL	http://b2share.eudat.eu/oai2d

Request was of type GetRecord.

OAI Record: oai:b2share.eudat.eu:392

OAI Record Header

OAI Identifier	oai:b2share.eudat.eu:392 oai_dc formats
Datestamp	2016-07-04T08:09:07Z
setSpec	RDA Identifiers Records

Unknown Metadata Format

<marc:record xsi:schemaLocation="http://www.loc.gov/MARC21/slim http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd" type="Bibliographic" >
<marc:leader>00000coc 2200000uu 4500</marc:leader>
<marc:controlfield tag="001" >392</marc:controlfield>
<marc:controlfield tag="005" >20160704110907.0</marc:controlfield>
<marc:datafield tag="024" ind1=" " ind2=" " >
<marc:subfield code="a" >dx.doi.org/10.15497/RDA00003</marc:subfield>
</marc:datafield>
<marc:datafield tag="024" ind1="7" ind2=" " >
<marc:subfield code="2" >checksum</marc:subfield>
<marc:subfield code="a" >de5d7c18fca9a91bed931b7590f37f3d19df0cf13424d7acc5d9b8a94ed788db</marc:subfield>
</marc:datafield>
<marc:datafield tag="024" ind1="7" ind2=" " >
<marc:subfield code="2" >PID</marc:subfield>
<marc:subfield code="a" >http://hdl.handle.net/11304/8446a925-9c88-4ae9-b190-a7cf39a5684f</marc:subfield>
</marc:datafield>
<marc:datafield tag="100" ind1=" " ind2=" " >
<marc:subfield code="a" >Amir Aryani</marc:subfield>
</marc:datafield>
<marc:datafield tag="245" ind1=" " ind2=" " >
<marc:subfield code="a" >Data Description Registry Interoperability WG: Interlinking Method and Specification of Cross-Platform Discovery</marc:subfield>
</marc:datafield>

Data Description Registry Interoperability WG: Interlinking Method and Specification of Cross-Platform Discovery; 1

Amir Aryani ; Data Description Registry Interoperability WG

03 July 2016
RDA

Abstract: This document describes the outcome of the Data Description Registry Interoperability (DDRI) working group and specification of the interoperability model implemented by the partners in this group. In addition, this document shows the testing of this model through an implementation called Research Data Switchboard, a collaborative project by the participants in this working group.

The record appears in these collections:
RDA

Files

Name	Date	Size	
AdoptionStatement.txt	04 Jul 2016	979 Bytes	Download
DDRIOutputSpecification.pdf	04 Jul 2016	576.4 kB	Download

Export

Export as BibTeX, MARC, MARCXML, DC, EndNote, NLM, RefWorks

Metadata

PID: http://hdl.handle.net/11304/8446a925-9c88-4ae9-b190-a7cf39a5684f

Beispiel 2

http://b2find.eudat.eu/dataset/81d9b405-f980-5c28-9bca-3a8a1a8b3ffa

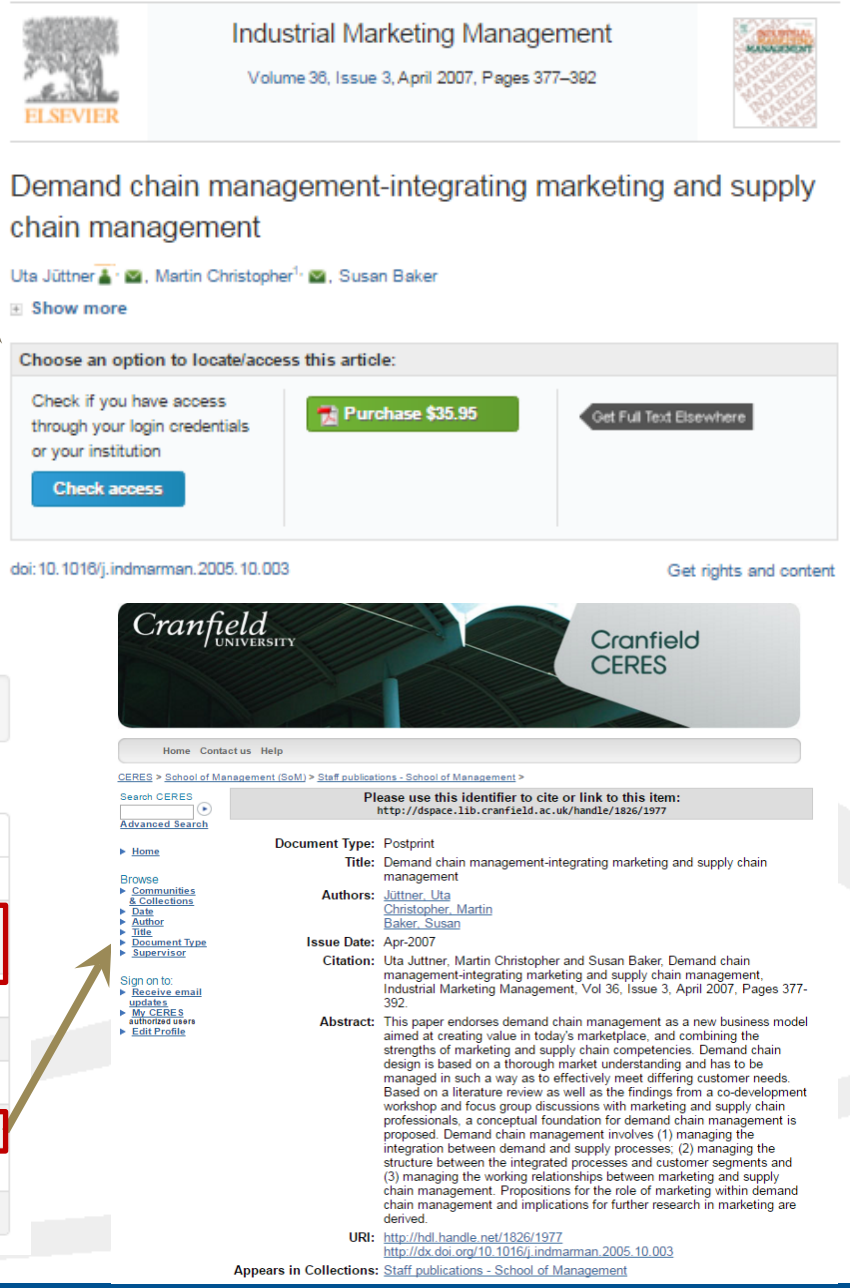
Demand chain management-integrating marketing and supply chain management

This paper endorses demand chain management as a new business model aimed at creating value in today's marketplace, and combining the strengths of marketing and supply chain competencies. Demar chain design is based on a thorough market understanding and has to be managed in such a way as to effectively meet differing customer needs. Based on a literature review as well as the findings from a co development workshop and focus group discussions with marketing and supply chain professionals, a conceptual foundation for demand chain management is proposed. Demand chain management involve (1) managing the integration between demand and supply processes; (2) managing the structure between the integrated processes and customer segments and (3) managing the working relationships between marketing and supply chain management. Propositions for the role of marketing within dem chain management and implications for further research in marketing are derived.

Postprint

Additional Info

Field	Value
Creator	Baker, Susan;Christopher, Martin;Jüttner, Uta
DOI	http://dx.doi.org/doi:doi:10.1016/j.indmarman.2005.10.003
Discipline	Marketing
Format	application/pdf;242936 bytes
Language	English
PID	http://hdl.handle.net/1826/1977
PublicationYear	2007
Publisher	Elsevier



Industrial Marketing Management

Volume 36, Issue 3, April 2007, Pages 377–392

Demand chain management-integrating marketing and supply chain management

Uta Jüttner, Martin Christopher, Susan Baker

Show more

Choose an option to locate/access this article:

Check if you have access through your login credentials or your institution

Check access

Purchase $35.95

Get Full Text Elsewhere

doi:10.1016/j.indmarman.2005.10.003

Get rights and content

Cranfield University

Cranfield CERES

Home Contact us Help

CERES > School of Management (SoM) > Staff publications - School of Management >

Search CERES

Advanced Search

Home

Browse

Communities & Collections

Date

Author

Title

Document Type

Supervisor

Sign on to:

Receive email updates

My CERES authorized users

Edit Profile

Please use this identifier to cite or link to this item:
http://dspace.lib.cranfield.ac.uk/handle/1826/1977

Document Type: Postprint

Title: Demand chain management-integrating marketing and supply chain management

Authors: Jüttner, Uta; Christopher, Martin; Baker, Susan

Issue Date: Apr-2007

Citation: Uta Juttner, Martin Christopher and Susan Baker, Demand chain management-integrating marketing and supply chain management, Industrial Marketing Management, Vol 36, Issue 3, April 2007, Pages 377-392.

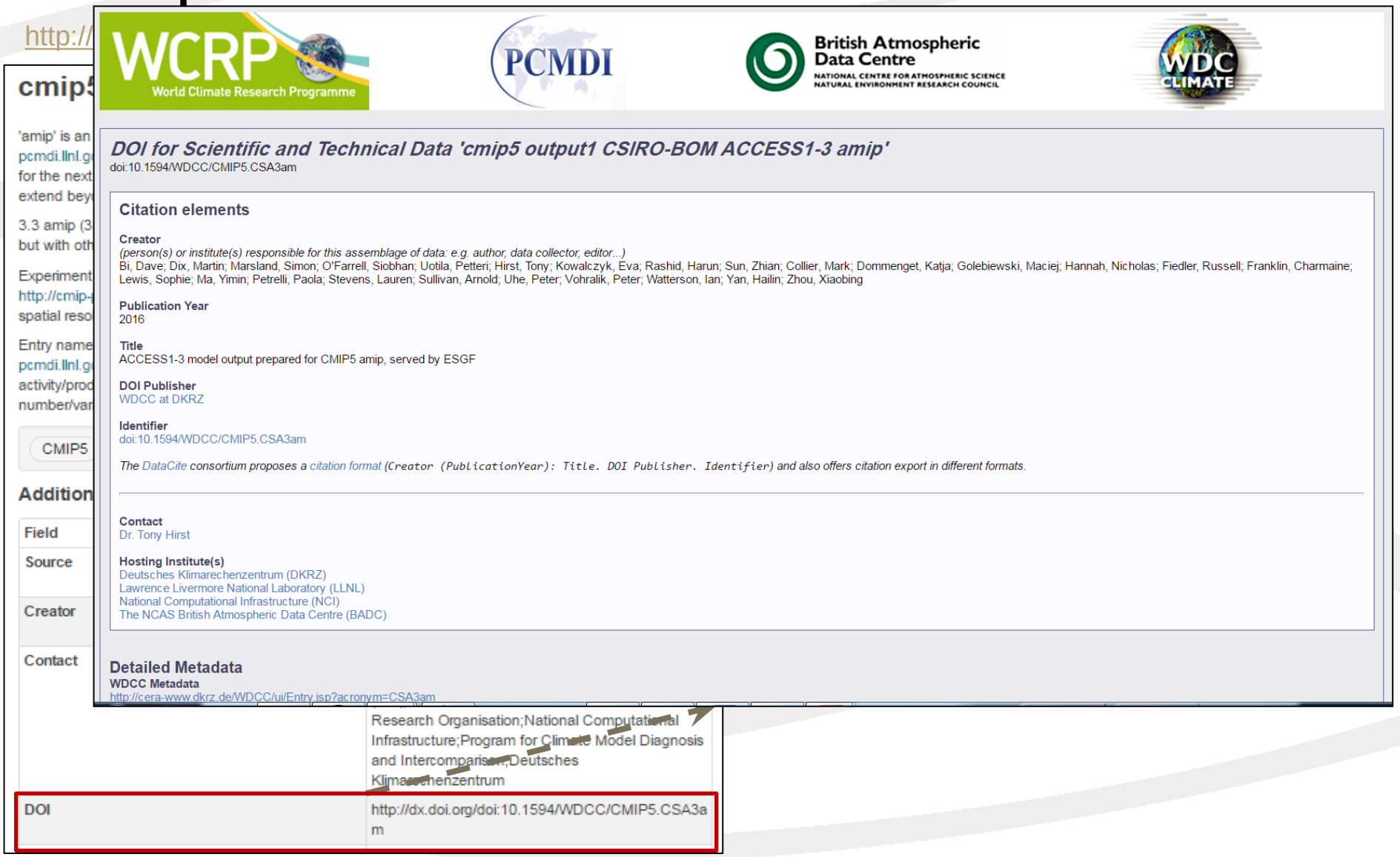
Abstract: This paper endorses demand chain management as a new business model aimed at creating value in today's marketplace, and combining the strengths of marketing and supply chain competencies. Demand chain design is based on a thorough market understanding and has to be managed in such a way as to effectively meet differing customer needs. Based on a literature review as well as the findings from a co-development workshop and focus group discussions with marketing and supply chain professionals, a conceptual foundation for demand chain management is proposed. Demand chain management involves (1) managing the integration between demand and supply processes; (2) managing the structure between the integrated processes and customer segments and (3) managing the working relationships between marketing and supply chain management. Propositions for the role of marketing within demand chain management and implications for further research in marketing are derived.

URI: http://hdl.handle.net/1826/1977
http://dx.doi.org/10.1016/j.indmarman.2005.10.003

Appears in Collections: Staff publications - School of Management

Beispiel 3

http://cera-www.dkrz.de/WDCC/CMIP5/Compact.jsp?acronym=CSA3am



Herausforderungen an EUDAT[-B2FIND]

- Balanceakt zwischen der Bereitstellung von generischen, disziplin-agnostischen Diensten und forschungsspezifische Bedürfnisse
- Skalierbarkeit und Performance
- Verbesserte Nutzerfreundlichkeit
- Uptake von Metadaten von B2SAFE
- Granularität und Grad der Aggregation

Weitere Entwicklungen in B2FIND

- Nutzen des Potential des ‚Semantic Web' (Linked Open Data) + Annotationen von Datensätzen
- Automatisierung des MD-Mappings
- Einbinden von (domainspezifischen) Ontologien und Thesauri
- Hierarchische Suche in Taxonomien (z.B. in ‚Tree of Disciplines' oder/und in ‚Levels of Aggregation')

Zusammenfassung zu B2FIND

- EUDAT-B2FIND
  - ist ein operativer MD Service und basiert auf Standards und Richtlinien wie den FAIR Prinzipien
  - bietet ein Suchportal über einen einheitlichen und weitgefächerten Suchraum mit vielen Facetten und Funktionalitäten
  - basiert auf einem umfassenden interdisziplinären MD Katalog, der Forschungsdaten aus vielen heterogenen und domain-spezifischen Quellen kombiniert
- Verbesserte Interoperabilität wird durch Homogenisierung auf ein vereinheitlichtes MD Schema erreicht

Daten- und Metadaten-Management für CMIP5

CMIP5 - History

1. 1988 : Das UN Umweltprogramm UNEP und die WMO gründete den Weltklimarat (IPCC), um „eine klare wissenschaftliche Sicht auf den aktuellen Stand des Wissens über den Klimawandel und seine […] Auswirkungen zu geben“.
2. IPCC beurteilt und bewertet die wissenschaftlichen, technischen und sozioökonom. Informationen, die für das Verständnis des Klimawandels relevant sind und erstellt die sog. Sachstandsberichte (AR‘s)
3. 2008 : Für den fünften Sachstandsbericht (IPCC AR5) wurde das internationale Modelvergleichsprojekt CMIP5 beauftragt die Koordination von Klimasimulationen zu übernehemen

CMIP5 in a nutshell

Das Climate Model Intercomparision Project (CMIP) ist ein Framework für globale Klimasimulationen unter der Schirmherrschaft des World Climate Research Programs (WCRP). In CMIP5 wurde eine Reihe von Standard-Modelexperimenten gefördert um

- zu evaluieren, wie realistisch die Modelle die nahe Vergangenheit reproduzieren können,
- Abschätzungen für den zukünftigen Klimawandel zu liefern und
- die Gründe für die Unterschiede in den Modellresultaten zu verstehen

Herausforderungen

- Große Datenvolumen (core data ~ 4 PB)
- Sehr viele Dateien (ca. 150000 Datensätze)
- Viele Modellierergruppen (>20) liefern Daten unterschiedlicher Art, Qualität und Struktur

➢ Standardisierung der Datenbeschreibung (Metadaten !)

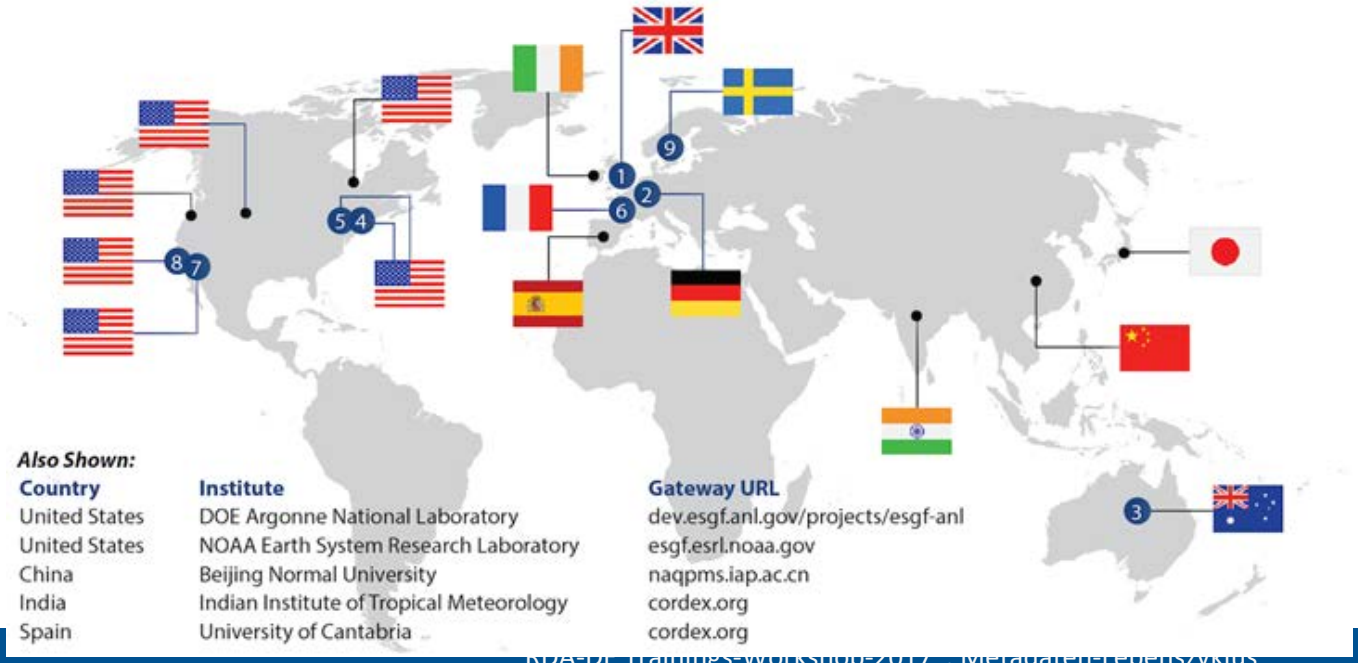
- Internationale Kooperation und globaler Archivverbund
- Lange Produktionszyklen (3 bis 4 Jahre)
- Interdisziplinäre Datennutzung über die Klimamodellierer-Community hinaus erfordert
  - Schnelle und zuverlässigen Datenzugriff
  - Ausführliche Dokumentation der Daten (Metadaten!)
  - Wahrung der Datenqualität
- Langzeitverfügbarkeit

ESGF

Earth System Grid Federation

A Peer-to-Peer Enterprise

The ESGF peer-to-peer enterprise system provides services and resources essential for global-scale Earth system science. This system is developed, deployed and maintained by an international multi-agency federation. ESGF's open source, operational code base disseminates petabytes of data including model simulations, observational, and reanalysis data for research assessments.

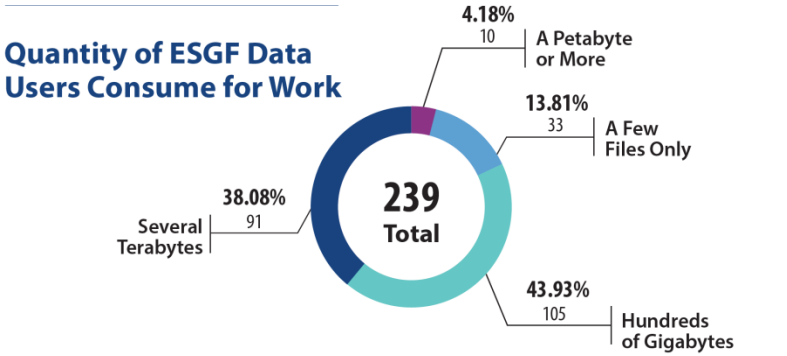


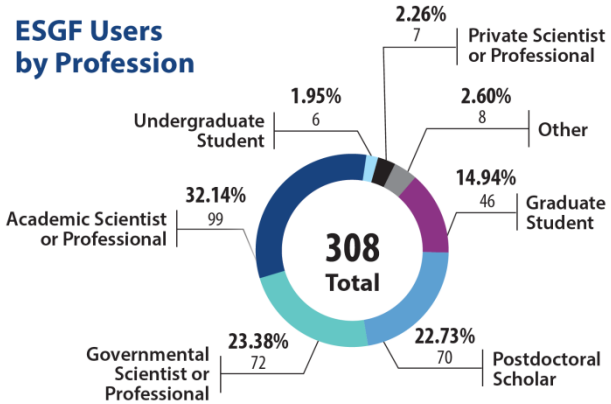
1	CEDA	UK
2	DKRZ	Germany
3	ANU NCI	Australia
4	NOAA GFDL	U.S.
5	NASA GSFC	U.S.
6	IPSL	France
7	NASA JPL	U.S.
8	DOE LLNL	U.S.
9	LiU	Sweden

ESGF
Earth System Grid Federation
A Peer-to-Peer Enterprise

ESGF: Digital Footprint

- Supports **>700,000** datasets from universities as well as national and international laboratories. **~4 million** datasets downloaded.
- Manages **>5 PB** of data in the total ESGF federated archive, which is expected to expand to >40 PB of uncompressed data, distributed across >25 projects and **~70** model intercomparison projects (MIPs).
- Services **18** highly visible national and international geophysical data products, including CMIP3, **CMIP5**, and soon CMIP6.



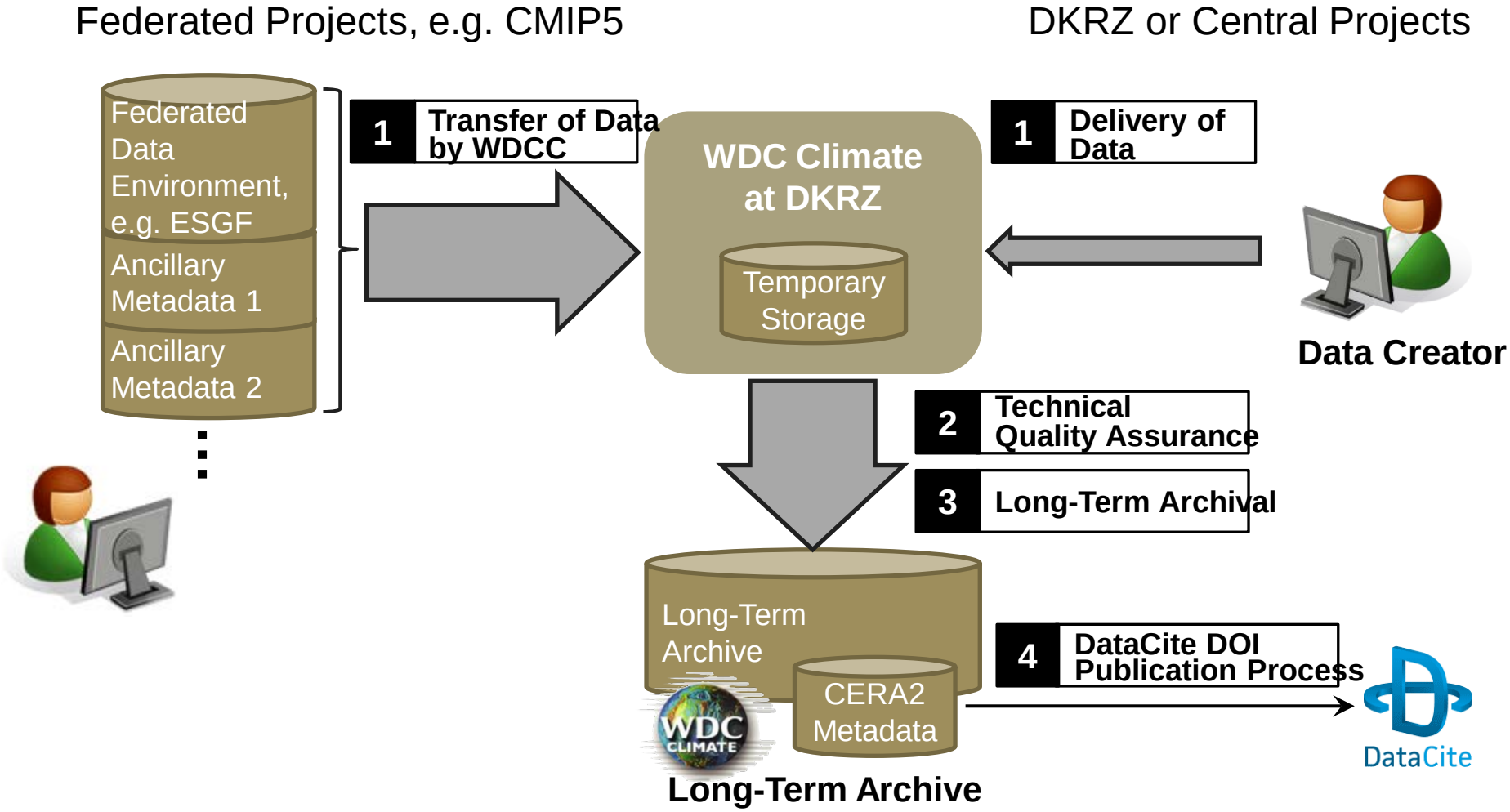


CMIP5 Metadaten

Verschiedene Ebenen : Metadaten, die

- die Simulationen und Experimente beschreiben befinden sich im ES-DOC (→ https://es-doc.org/)
- die Daten (Output der Klimasimulationen) beschreiben findet man in den Headern der Output-Dateien selbst (netCDF files)
- die CMIP5 Experimente als ganzes beschreiben findet man z.B. über B2FIND, aber insbesondere in der CERA-Datenbank des WDCC (→ https://cera-www.dkrz.de/WDCC/ui/cerasearch/)

WDCC's Long Term Archival



Danke für die Aufmerksamkeit

- Fragen bitte an
  - → widmann@dkrz.de
  - → https://eudat.eu/support-request
- Metadata standards
  - RDA : http://rd-alliance.github.io/metadata-directory/standards/
  - DCC : http://www.dcc.ac.uk/resources/metadata-standards
- EUDAT-B2FIND: http://b2find.eudat.eu
- B2FIND-Training: https://github.com/EUDAT-Training/B2FIND-Training
- CMIP5: http://cmip-pcmdi.llnl.gov/cmip5/