UNIVERSITÄT ZU LÜBECK

Information Systems

CS4130-KP06

Prof. Dr. Sylvia Melzer

SoSe2026

Information Systems & Research Data Management

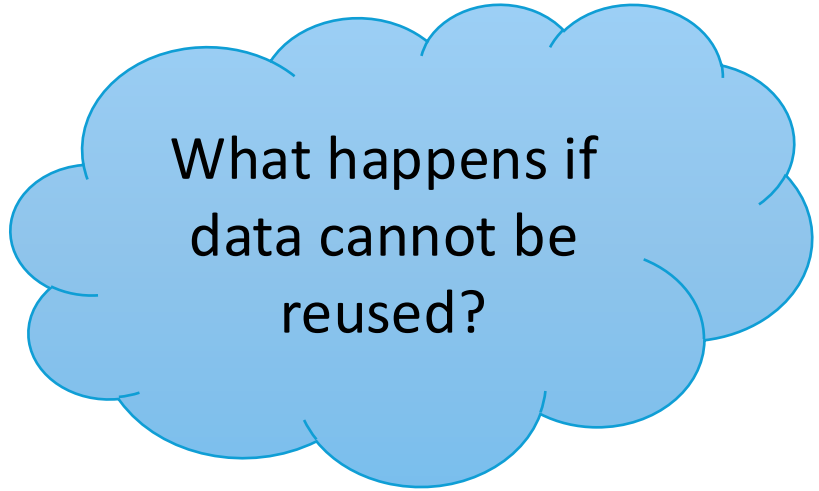
Information Systems

What You Will Learn

- Understand the concept and goals of Research Data Management (RDM)
- Learn the research data lifecycle and key RDM activities
- Identify challenges in managing heterogeneous research data
- Understand the role of metadata, standards, and data quality
- Learn key principles: FAIR and CARE
- Recognize limitations of existing RDM approaches
- Understand how "EASE” extends RDM toward exploration and interpretation

Motivation

- Research produces large and complex datasets
- Data are essential for scientific knowledge creation
- Poor data management leads to:
  - Data loss
  - Inconsistent results
  - Lack of reproducibility
- Data must be organized, stored, and accessible
- Increasing importance due to:
  - Digitalization
  - Big data
  - AI-driven research

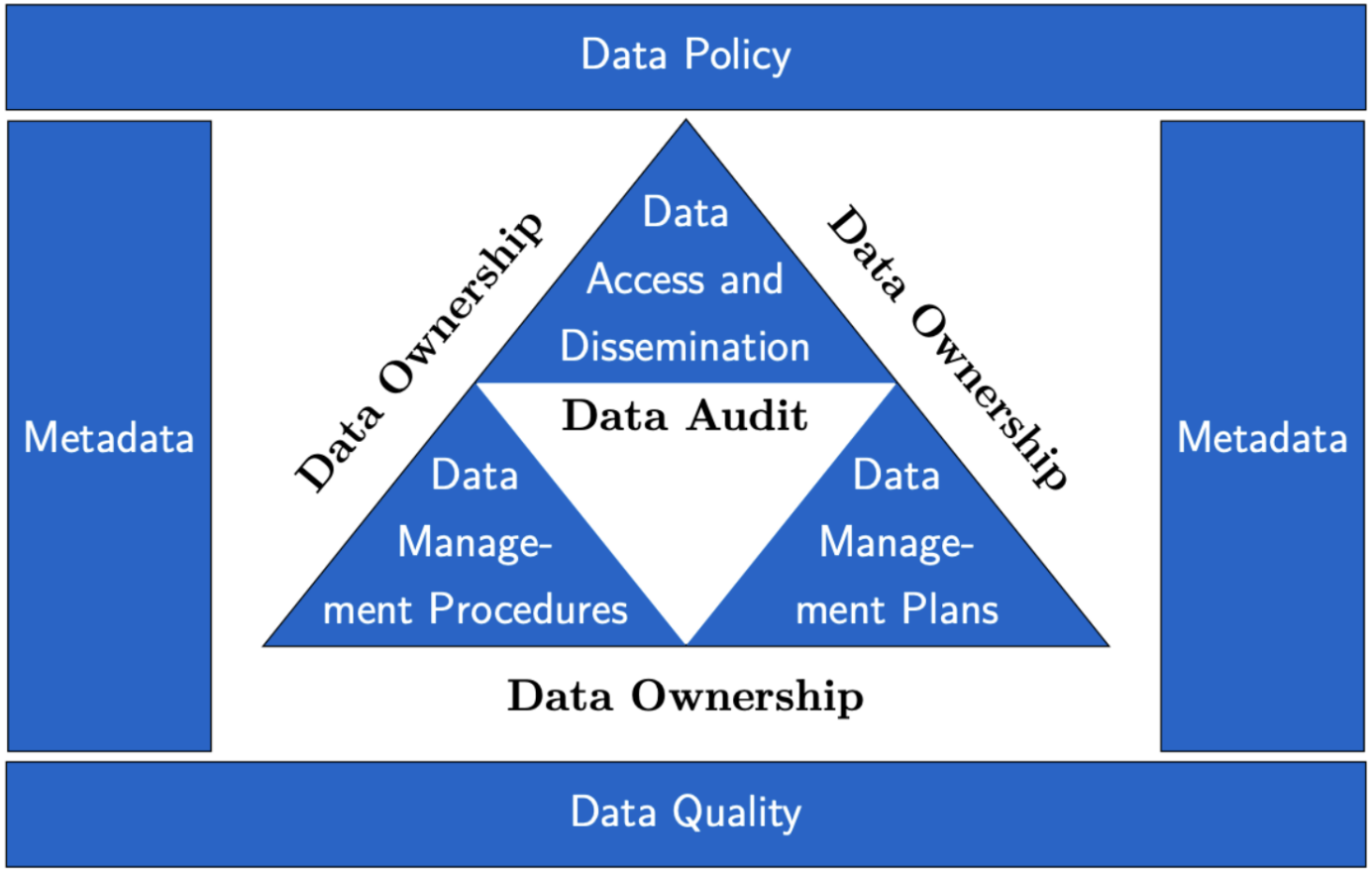


Data Management

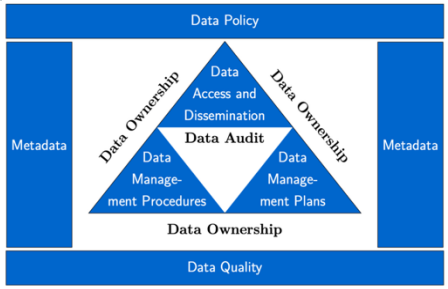
- Origin of data management dates back to the 1960s Association of Data Processing Service Organizations (ADAPSO)
- Covers all types of data:
  - Paper-based data
  - Database content
  - Multimedia and scientific data
- Definition: Data management refers to the targeted **handling of data within an organization or institution**. It includes the planning, organization, control and monitoring of all processes required for the **collection, storage, processing**, **analysis, and use of data**. The main aim is to **ensure the quality and availability of data** and thus **enable efficient and effective use**. Synonymous terms for data management are, for example, data administration, data maintenance, or data handling.

IPH - Institut für Integrierte Produktion Hannover gGmbH. Datenmanagement - neben maschinellem Lernen ein wichtiger Bereich von Data Science. https://www.iph-hannover.de/de/dienstleistungen/data-science/datenmanagement/, 2023. Accessed: October 6, 2023. [36]

The Key Principles of Data Management



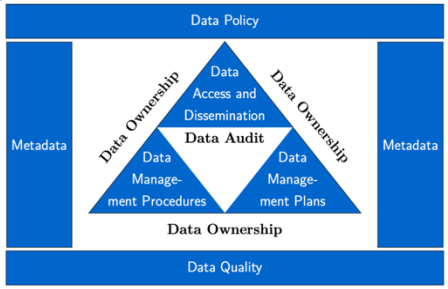
The Key Principles of Data Management



- Data management is guided by core principles:
  - Data policy → defines rules and responsibilities
  - Data ownership → clarifies rights and control
- Central components:
  - Data access and dissemination
  - Data audit and monitoring
  - Data management procedures
  - Data management plans
- Supporting elements:
  - Metadata → describes and structures data
  - Data quality → ensures reliability

Who owns data in research projects?

Data Access and Dissemination



- Data access must follow:
  - Legal regulations
  - Institutional policies
  - Ethical guidelines
- Important aspects:
  - FAIR principles
  - CARE principles
- Goal:
  - Enable controlled and responsible data sharing
  - Balance openness and protection

Research data in the humanities

- Publications, measurement data, laboratory results, audio and video data, texts, collection items, interviews, software, etc.
- Research data:

Palm leaves with text in Singhalese from Sri Lanka (ca. mid-20th century). Private collection.

Challenges:

- Paper-based processes limit the reusability of data
- Diverse data sources, collaboration with partners, and file formats

FAIR Principles

- Findable:
  - Data must have identifiers (e.g., DOI)
  - Metadata must be searchable
- Accessible:
  - Data retrievable via standardized protocols
  - Access may be restricted
- Interoperable:
  - Use of shared formats and standards
  - Integration across systems
- Reusable:
  - Clear licenses
  - Well-described data

Project sponsors are calling for the implementation of the FAIR principles

https://www.dfg.de/resource/blob/174052/1a235cb138c77e353789263b8730b1df/kodex-gwp-en-data.pdf

FAIR" Principles: the Origin

www.nature.com/scientificdata

SCIENTIFIC DATA

Amended: Addendum

OPEN

SUBJECT CATEGORIES
» Research data
» Publication characteristics

Comment: The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson *et al.*#

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.

Received: 10 December 2015
Accepted: 12 February 2016
Published: 15 March 2016

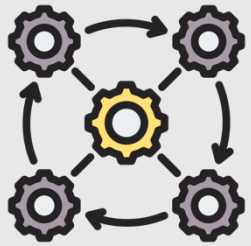
FAIR Principles

Findable

F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

Accessible



Interoperable

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

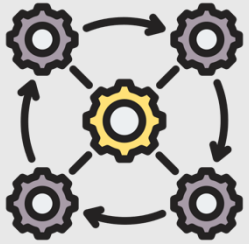
Reusable

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

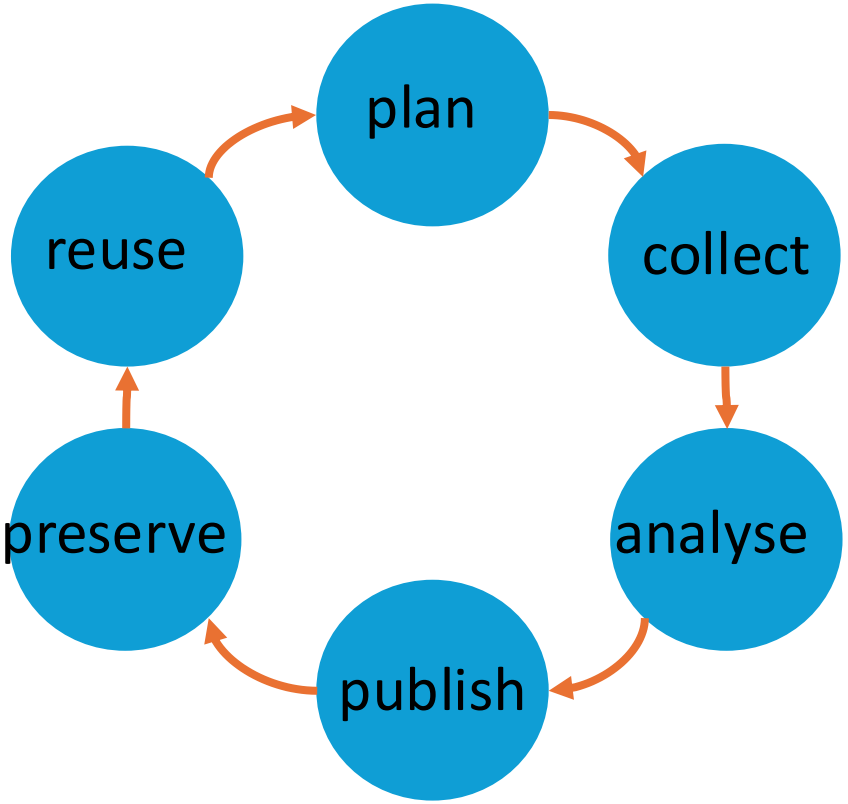
Unlocking icons created by Mihimihi, Interoperability icons created by Uniconlabs, Reusable icons created by Lumi - Flaticon

FAIR Data Management

Findable



Interoperable



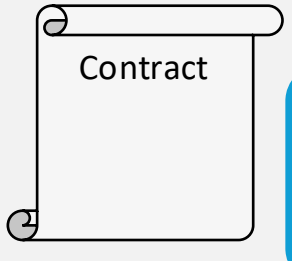
https://labfolder.com/guide-research-data-management/

Unlocking icons created by Mihimihi, Interoperability icons created by Uniconlabs, Reusable icons created by Lumi - Flaticon

Accessible

Reusable

FAIR Data Management



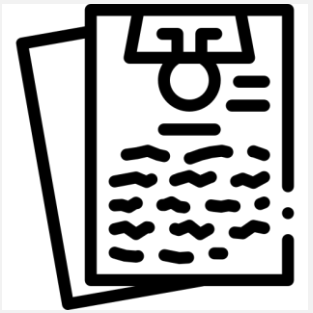
FAIR Data ≠ Public data

§

How can you find data?

publish

Ethics, and law icons created by Freepik - Flaticon



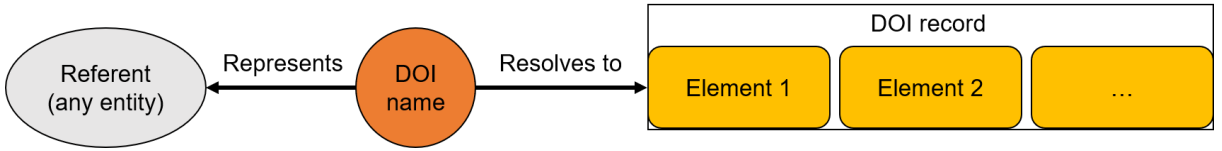
Images from Flaticon.com

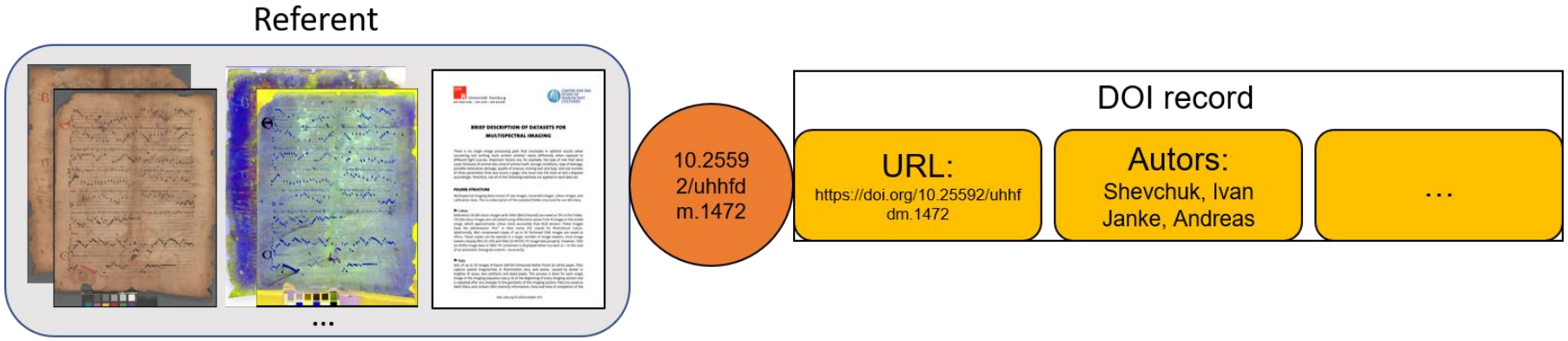
Palm leaves with text in Singhalese from Sri Lanka (ca. mid-20th century). Private collection.

Digital Object Identifier (DOI)

- A DOI is a persistent identifier used to uniquely identify various objects
- Specification in ISO 26324:2012
- Example:

- Concept:





DOI Foundation. DOI HANDBOOK. DOI Foundation, info@doi.org, April 2023;
Wikipedia contributors. Digital Object Identifier
https://en.wikipedia.org/wiki/Digital_object_identifier, 2024

Metadaten-Standards

- DataCite Metadata Schema
- Dublin Core
- Machine-Readable Cataloging
- Metadata Object Description Schema
- Metadata Encoding and Transmission Standard
- …

Table 1: DataCite Mandatory Properties

ID	Property	Obligation
1	Identifier (with mandatory type sub-property)	M
2	Creator (with optional given name, family name, name identifier and affiliation sub-properties)	M
3	Title (with optional type sub-properties)	M
4	Publisher	M
5	PublicationYear	M
10	ResourceType (with mandatory general type description sub-property)	M

https://datacite.org/

Metadaten-Standards

- DataCite Metadata Schema
- Dublin Core
- Machine-Readable Cataloging
- Metadata Object Description Schema
- Metadata Encoding and Transmission Standard
- …

Core elements [edit]

The Dublin Core vocabulary that was published in 1999 consisted of 15 terms:

• contributor	• date	• identifier	• relation	• subject
• coverage	• description	• language	• rights	• title
• creator	• format	• publisher	• source	• type

The vocabulary was commonly expressed in HTML 'meta' tagging in the "<head>" section of an HTML-encoded page.[16]

```
<head>
   <meta name="DC.title" content="Services to Government" >
   <meta name="DC.date" content="1997-07" >
 </head>
```

https://en.wikipedia.org/wiki/Dublin_Core

Metadaten-Standards

- DataCite Metadata Schema
- Dublin Core
- Machine-Readable Cataloging
- Metadata Object Description Schema
- Metadata Encoding and Transmission Standard
- …

MARC Field	Description
001	Control number
024	Other Standard Identifier
100	Main Entry (Author)
245	Title Statement
260	Publication Information
300	Physical Description
856	Electronic Location and Access

Metadaten-Standards

- DataCite Metadata Schema
- Dublin Core
- Machine-Readable Cataloging
- Metadata Object Description Schema
- Metadata Encoding and Transmission Standard
- ...

Element	Description
TitleInfo	Information about the title of the resource.
Name	Information about the creator or contributor.
TypeOfResource	The type of resource (e.g., text, image, sound).
Genre	The genre of the resource.
OriginInfo	Information about the origin of the resource, such as date of creation or publication.
Language	Language of the resource.
PhysicalDescription	Information about the physical characteristics of the resource.
Abstract	A summary or abstract of the content.
Subject	The subject of the resource.
RelatedItem	Information about related resources.

Metadaten-Standards

- DataCite Metadata Schema
- Dublin Core
- Machine-Readable Cataloging
- Metadata Object Description Schema
- Metadata Encoding and Transmission Standard
- ...

```
1  <mets:behaviorSec ID="behavior1">
2      <mets:behavior ID="nav_behavior" STRUCTID="dmd1" LABEL="Page
           Navigation" TYPE="navigation">
3          <mets:interfaceDef ID="nav_interface" LABEL="Simple Page Viewer">
4              <mets:mechanism xlink:href="http://example.com/viewer" xlink:title
                   ="Viewer Application"/>
5          </mets:interfaceDef>
6          <mets:behaviorMechanism ID="mechanism1" LABEL="Page Navigation
               Script">
7              <mets:mechanism xlink:href="http://example.com/script.js" xlink:
                   title="Page Navigation Script"/>
8          </mets:behaviorMechanism>
9      </mets:behavior>
10 </mets:behaviorSec>
```

Listing 4.5: METS Behavior Section.

FAIR Data Management at UHH/UWA

RDR@UHH Research Data Repository (https://www.fdr.uni-hamburg.de/)

- Objective: Storing research data in a repository
- Application of standards
- Compliance with data protection regulations
- Supporting researchers in locating, identifying, and citing research data

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

ZENTRUM
FÜR NACHHALTIGES
FORSCHUNGSDATENMANAGEMENT

Recent open datasets

August 6, 2024 (v1) | Dataset | Open Access

View

AO 8295 (X-Ray Tomography 3D data of an Enveloped Clay Tablet, Louvre Museum, Paris)

Michel, Cécile; Schroer, Christian; Olbrich, Stephan; Ehteram, Samaneh; Beckert, Andreas

Written Artefact Metadata Object type: Enveloped clay tablet Material: Clay Writing: Cuneiform writing Language: Cappadocian Nature of the text: Debt contract Provenience: Kanesh (mod. Kültepe) Period/Date: Old Assyrian (ca. 1950-1850 BC) (1st half of 2nd millennium) Place of discovery:...

Uploaded on August 6, 2024

June 19, 2024 (v2) | Dataset | Open Access

View

FTIR spectra of 11 palm-leaf manuscripts

Voges, Lucas F.; Horn, Nils; Colini, Claudia; Seifert, Stephan

This dataset was created in the context of the Research Field F (RFF02), Research Field A, and the Palm-Leaf Manuscript Profiling Initiative (PLMPI) at the Centre for the Study of Manuscript Cultures (CSMC) and Cluster of Excellence 'Understanding Written Artefacts' (UWA). Manuscripts Eleven...

Uploaded on August 6, 2024

1 more version(s) exist for this record

August 1, 2024 (v1) | Dataset | Open Access

View

Behavioral and Eye-tracking Data for "Individual differences in belief updating and phasic arousal are related to psychosis proneness"

Recent activity of the FDM-Center

- DFG-Checklist Research Data Management
- Broschure Research Data Management
- New UHH Data Notes
- Manual for the Research Data Repository
- Research Data Repository FDR@UHH started

Read more about our latest activity.

Get Help

- read the user manual for the ZFDM Repository
- consult the ZFDM Repository terms of service
- contact repository.fdm@uni-hamburg.de for support.

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

ZENTRUM
FÜR NACHHALTIGES
FORSCHUNGSDATENMANAGEMENT

DOI record
Element 1 | Element 2 | …

HOME UPLOAD SYLVIA.MELZER@UNI-HAMBURG.DE

August 6, 2024

Dataset Open Access

AO 8295 (X-Ray Tomography 3D data of an Enveloped Clay Tablet, Louvre Museum, Paris)

Michel, Cécile; Schroer, Christian; Olbrich, Stephan; Ehteram, Samaneh; Beckert, Andreas; Schropp, Andreas; Pätzold, Philipp; Wiljes, Patrik; Bohn, Matthias; Döhrmann, Ralph; Pataï, Véronique

Written Artefact Metadata

Object type: Enveloped clay tablet

Files (24.3 GB)

Referent (any entity)

Name	Size	
000_AO8295_anim5b_small.gif md5:328663ad424e46cc96779d2c89e42bd6	4.0 MB	Preview Download
001_AO8295_anim5b.gif md5:2112aa16d99ab3a194910a5ac2abe782	69.5 MB	Preview Download
002_AO8295_2_4_0.5_2_crossBilateralTukey_auto_1_3_mean_0.0625_auto_1_3_auto_1.25_0.015_32_8_-0.5_i2 Envelope.glb md5:09266678f14ba83672ff0a6afe12e05b	172.6 MB	Download
003_AO8295_2_4_0.5_2_crossBilateralTukey_auto_1_3_mean_0.0625_auto_1_3_auto_1.25_0.015_32_8_-0.5_i4 Letter.glb	43.0 MB	Download

Publication date:
August 6, 2024

DOI:
DOI 10.25592/uhhfdm.14776

Keyword(s):
Written Artefact | UWA | Cluster of Excellence | CSMC | DESY | RFA09 | Artefact Profiling | 3D | X-Ray | Tomography | ENCI | Cuneiform | Clay tablet | Assyriology | Informatics | Physics | Louvre | Museum | Paris

Communities:
Centre for the Study of Manuscript Cultures
UHH

Versions

Version 1 10.25592/uhhfdm.14776 Aug 6, 2024

RDR Export Formats

April 14, 2026

Dataset | Open Access

Dataset for "Laser-dressed partial density of states"

Tatiana Bezriadina; Daria Popova-Gorelova

MARC21 XML Export

```
<?xml version='1.0' encoding='UTF-8'?>
<record xmlns="http://www.loc.gov/MARC21/slim">
  <leader>00000nmm##2200000uu#4500</leader>
  <datafield tag="653" ind1=" " ind2=" ">
    <subfield code="a">Theoretical physics</subfield>
  </datafield>
  <controlfield tag="005">20260414072838.0</controlfield>
  <datafield tag="024" ind1=" " ind2=" ">
    <subfield code="a">10.25592/uhhfdm.18551</subfield>
    <subfield code="2">doi</subfield>
  </datafield>
  <datafield tag="909" ind1="C" ind2="O">
    <subfield code="o">oai:fdr.uni-hamburg.de:18551</subfield>
    <subfield code="p">user-uhh</subfield>
  </datafield>
  <datafield tag="700" ind1=" " ind2=" ">
    <subfield code="a">Daria Popova-Gorelova</subfield>
    <subfield code="u">Brandenburg University of Technology Cottbus Senftenberg</subfield>
    <subfield code="0">(orcid)0000-0002-3036-0467</subfield>
  </datafield>
  <datafield tag="773" ind1=" " ind2=" ">
    <subfield code="a">10.25592/uhhfdm.18550</subfield>
    <subfield code="i">isVersionOf</subfield>
    <subfield code="n">doi</subfield>
  </datafield>
  <datafield tag="540" ind1=" " ind2=" ">
    <subfield code="u">https://creativecommons.org/licenses/by/4.0/legalcode</subfield>
    <subfield code="a">Creative Commons Attribution 4.0 International</subfield>
  </datafield>
  <datafield tag="100" ind1=" " ind2=" ">
    <subfield code="a">Tatiana Bezriadina</subfield>
    <subfield code="u">University of Hamburg</subfield>
  </datafield>
  <datafield tag="520" ind1=" " ind2=" ">
    <subfield code="a">&lt;p&gt;File names with the data for Figure 1b:&lt;/p&gt;

&lt;p&gt;O_p_states.txt&lt;/p&gt;

&lt;p&gt;O_s_states.txt&lt;/p&gt;

&lt;p&gt;Zn_d_states.txt&lt;/p&gt;
```

Publication date:
April 14, 2026

DOI:
DOI 10.25592/uhhfdm.18551

Keyword(s):
Theoretical physics

Communities:
UHH

License (for files):
Creative Commons Attribution 4.0 International

Versions

Version 1 10.25592/uhhfdm.18551 Apr 14, 2026

Cite all versions? You can cite all versions by using the DOI 10.25592/uhhfdm.18550. This DOI represents all versions, and will always resolve to the latest one.

Cite record as

Tatiana Bezriadina, & Daria Popova-Gorelova. (2026). Dataset for "Laser-dressed partial density of states" [Data set]. http://doi.org/10.25592/uhhfdm.18551

Start typing a citation style...

Export

BibTeX CSL DataCite Dublin Core JSON JSON-LD MARCXML Mendeley

Compliance with the FAIR principles

F1. (meta)data are assigned a globally unique and persistent identifier
F2. data are described with rich metadata (defined by R1 below)
F3. metadata clearly and explicitly include the identifier of the data it describes
F4. (meta)data are registered or indexed in a searchable resource

A1. (meta)data are retrievable by their identifier using a standardized communications protocol
A1.1 the protocol is open, free, and universally implementable
A1.2 the protocol allows for an authentication and authorization procedure, where necessary
A2. metadata are accessible, even when the data are no longer available

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles
I3. (meta)data include qualified references to other (meta)data

R1. meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1. (meta)data are released with a clear and accessible data usage license
R1.2. (meta)data are associated with detailed provenance
R1.3. (meta)data meet domain-relevant community standards

- FA: Supported by RDR
- I: For humans → Viewer
- I: For machines → CSV, JSON, TEI,…
- R: User-specific tools?

Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

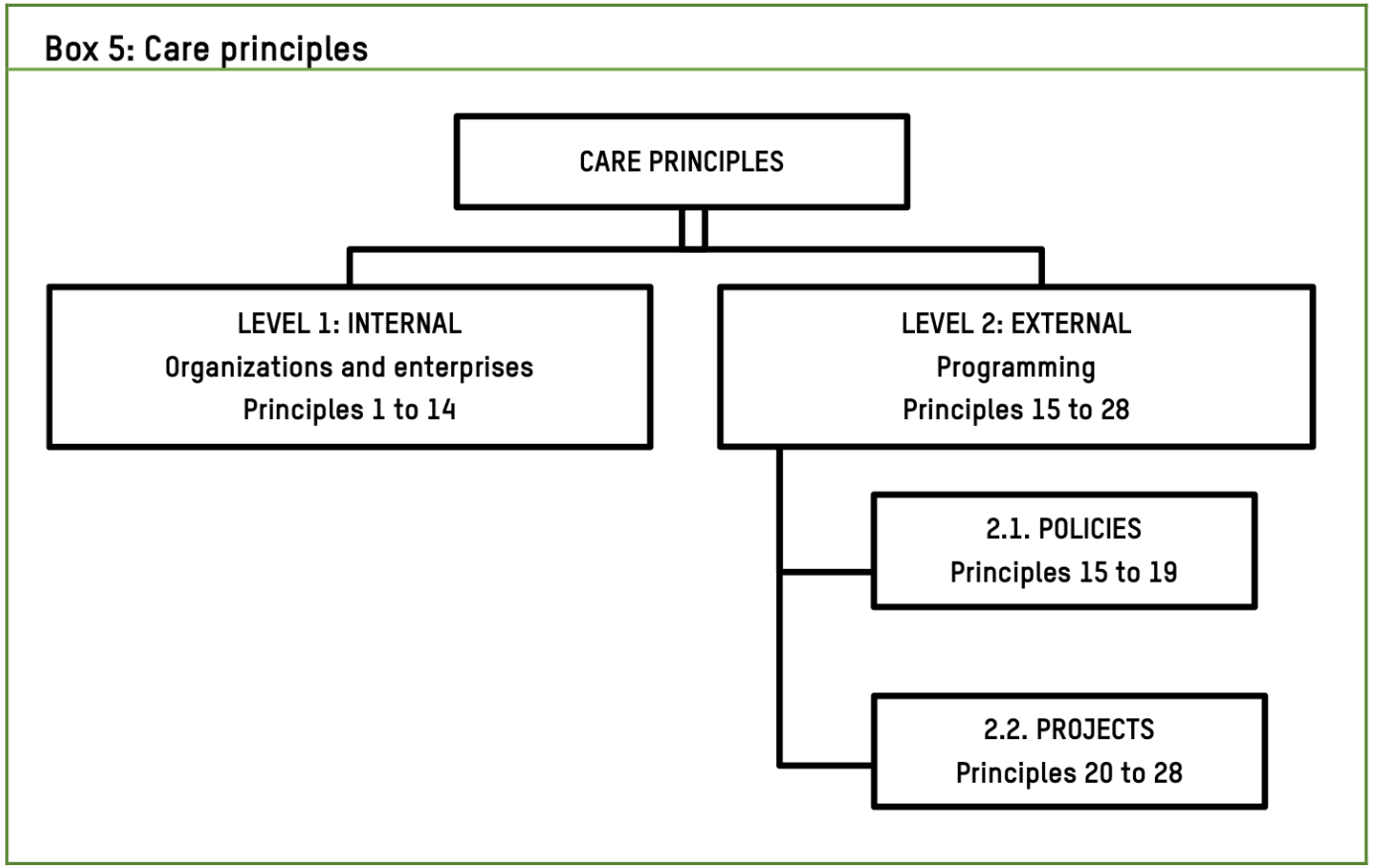
Data Set Availability and Citation in the Digital Age

- Mandatory for DFG-funded research: **Adherence to the FAIR principles**
- Ensuring the discoverability, accessibility, interoperability, and reusability of digital assets
- **Common misconception:** Simply building a database and providing a browser-based GUI is sufficient.
- Access via computational processes and human inspection (an interoperability issue)
- Databases can nevertheless be useful in the creation of datasets

CARE Principles

- The GIDA (Global Indigenous Data Alliance) developed the CARE principles for Indigenous Data Governance. These principles are a valuable addition to the FAIR principles. They provide important guidance on utilizing research data.
- **Collective benefit:** Data ecosystems should be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.
- **Authority to control:** Indigenous Peoples? rights and interests in Indigenous data must be recognized and their authority to control such data be empowered.
- **Responsibility:** Those working with Indigenous data have a responsibility to share how data is used to support Indigenous Peoples? self-determination and collective benefit.
- **Ethics:** Indigenous People rights and well-being should be the primary concern at all stages of the data life cycle and across the data ecosystem. → z. B. Ethics Guidelines CSMC

CARE Principles



https://oxfamilibrary.openrepository.com/server/api/core/bitstreams/8258ba32-34b4-491c-ab73-768cb56a7eff/content

INTERNAL CARE PRINCIPLES FOR ORGANIZATIONS AND ENTERPRISES

Realign institutional policies to recognize, reduce and redistribute unpaid care work

Develop strategies that promote and reward decent work practices

Create opportunities for representation and social dialogue

CARE PRINCIPLES FOR EXTERNAL PROGRAMMING

Influence sector policies to recognize, reduce and redistribute unpaid and underpaid/paid care work.

Develop strategies to incentivize care-centric reforms within sectors and in public finance management systems

Initiate dialogue on recognizing, reforming and regularizing informal workers and workers in non-standard employment

Create opportunities for representation by strengthening civil society and trade union participation in public policy and budget discussions

Ethical and Responsible Research

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CENTRE FOR THE STUDY OF MANUSCRIPT CULTURES

Ethical and Responsible Research at CSMC

Cécile Michel, Michael Friedrich, Jost Gippert, Konrad Hirschler, Patrick Huber, Davidson MacLaren, Sylvia Melzer, Ralf Möller, Maria Luisa Russo, Doreen Schröter and Stefan Thiemann

'Treat others as you would like others to treat you.
Treat others' objects as you would like others to treat yours.'

Preamble[1]

The *Centre for the Study of Manuscript Cultures* (CSMC) and its Cluster of Excellence *Understanding Written Artefacts* (UWA) are a large international community of scholars from the humanities and natural and computer sciences who collaborate closely in studying written artefacts from the beginning of writing to the present day. Throughout their work, researchers interact with people (citizens, colleagues, and officials) as well as institutions. Such large-scale, global, and cross-disciplinary collaborations should be carried out in accordance with ethical and responsible research practices. The purpose of this document is to offer advice to researchers working on written artefacts and to provide a brief overview of common issues concerning the ethical treatment of researchers, artefacts, and data.

In all our disciplines, senior researchers are strongly encouraged to train students and early career researchers in the best practices and make sure that ethics is part of their education. Good scientific practice should be seen as a part of ethical behaviour.

https://doi.org/10.25592/uhhfdm.14757

Key questions in the context of research data and ethics

- How to receive consent6 (to collect data, to take images and other research data)?
- How to collect data?
- How to use and cite data generated through social media and the internet?
- How to handle data?
- How to publish data (access rights, ‘as open as possible, as closed as necessary’)?
- How to protect personal data (to ensure that no personal or organisational rights are violated)?

https://doi.org/10.25592/uhhfdm.14757

EASE: The Extension of FAIR and CARE

FAIR principles ensure:

- Data is findable and accessible
- Data can be shared and reused

BUT in practice:

- Data is often difficult to understand
- Context and meaning are not fully captured
- Complex relationships remain hidden

EASE: The Extension of FAIR and CARE

Especially in humanities research:

- Data is heterogeneous and context-dependent
- Interpretation plays a central role
- Research is iterative and exploratory

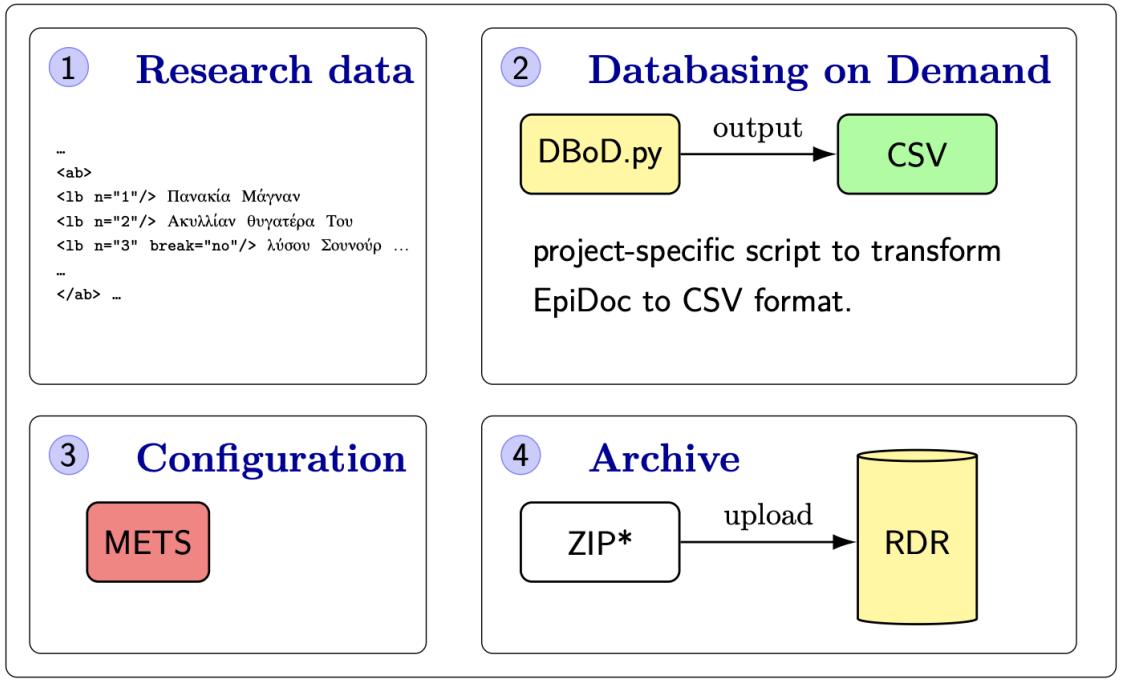
Key problem:

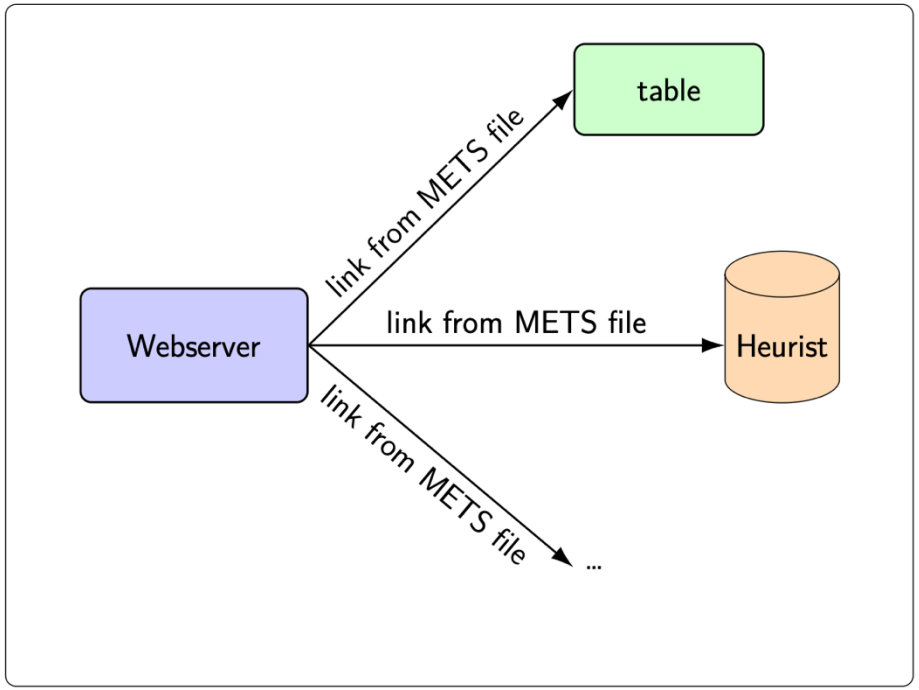
- FAIR focuses on data availability
- Not on data usability in research practice

Therefore, data must be EASE-oriented:

- Research data should be **E**xplorable, **A**ccessible in terms of data quality, **S**eeable in operation in new contexts, and **E**asily checkable for reuse.

EASE: Example





How EASE Extends RDM?

Traditional RDM focuses on storing, documenting, and providing access to data

This ensures availability but does not guarantee understanding or usability

In research practice, data must be explored, interpreted, and contextualized

EASE extends RDM by making data explorable and navigable

It enables researchers to see data in context and understand relationships

It also ensures that data quality can be assessed and results can be verified

As a result, data is transformed into usable knowledge for research

Conclusion

- Research Data Management ensures that data is organized, accessible, and reusable
- The FAIR principles provide a strong foundation for data sharing and interoperability
- FAIR mainly focuses on data availability rather than its usability in research practice
- In many research contexts, especially in the humanities, data is complex, heterogeneous, and strongly context-dependent
- Researchers need to explore, interpret, and contextualize data to generate knowledge and be CARE
- The EASE approach extends RDM by enabling exploration, interpretation, and validation of data
- It transforms data into usable knowledge and supports the research process
- The key takeaway is a shift from data management to data understanding and from data access to knowledge generation