

Data Management Planning

Reproducibility, Linked Open Data &
Semantic Interoperability

FORS and DaSCH

FORS – Swiss Centre of Expertise in the Social Sciences

- national infrastructure for Social Sciences' research data mainly funded by SNSF
- services: consulting / training / workshops /events for data management and archiving, SWISSUbase repository for the social sciences, mandates around the data collection and analysis, FORS Guides

DaSCH – Swiss National Data and Service Center for the Humanities

- national infrastructure for Humanities' research data mainly funded by SNSF
- services: consulting / training / workshops /events, virtual research environment, FAIR open data repository (DSP) including data publication and persistent identifiers at object level, metadata browser



Rita Gautschy (DaSCH)
rita.gautschy@dasch.swiss



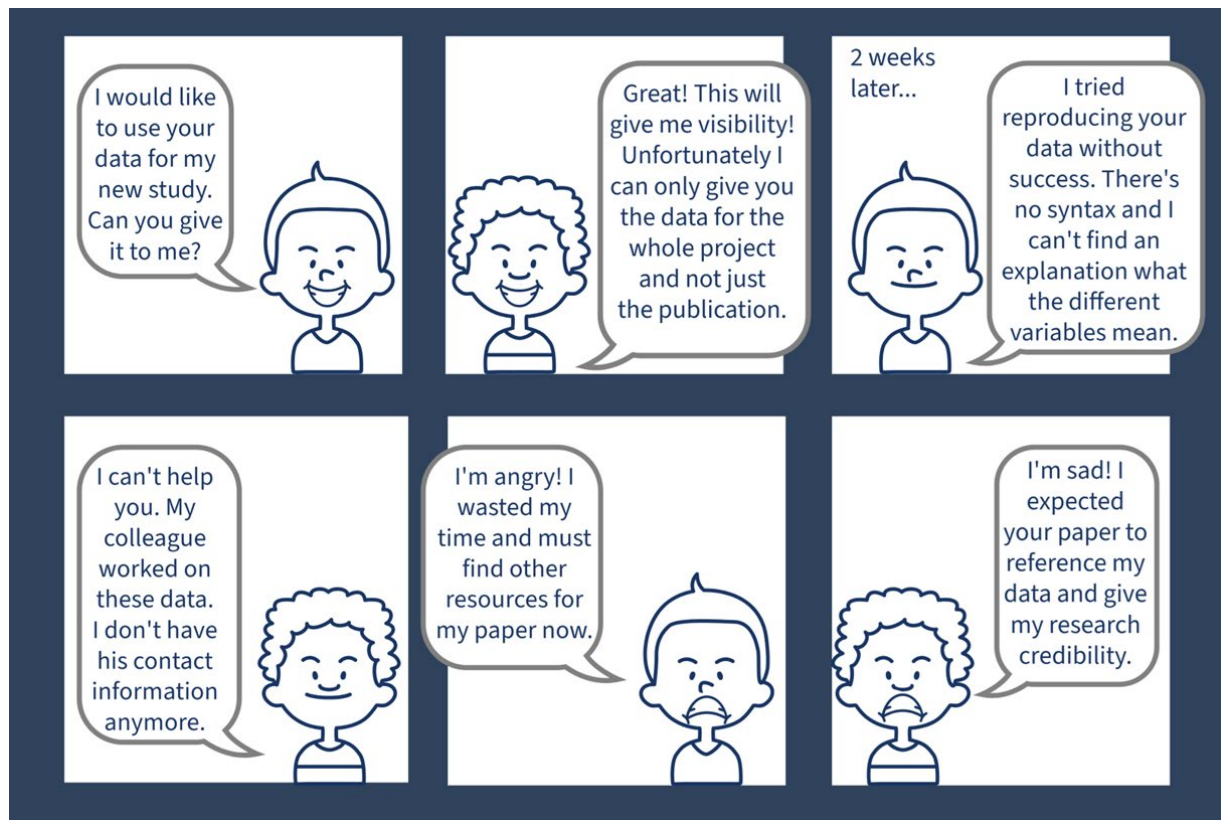
Emilie Morgan de Paula (FORS)
emilie.morgandepaula@fors.unil.ch

Programme

- 01 Reproducibility
- 02 The Replication Crisis
- 03 Practices for Reproducible Research
- 04 Linked Open Data & Semantic Interoperability
- 05 Important Standards
- 06 Q&A

Reproducibility

A familiar situation



Definitions

Mapping the Landscape of Reproducibility

	Same data	Different data
Same Methods	Reproducibility	Replicability
	Analytical or computational reproducibility	Direct replication
Different Methods	Robustness	Generalizability
	Extended or systematic replication	Conceptual replication

Inspired from: Schloss, PD. 2018. Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. *mBio* 9:10.1128/mbio.00525-18. <https://doi.org/10.1128/mbio.00525-18> ; Schmidt, S. 2009. Shall We Really Do It Again? The Powerful Concept of Replication Is Neglected in the Social Sciences. *Review of General Psychology* 13(2):90–100. <https://doi.org/10.1037/a0015108> ; Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van 't Veer, A. 2014. The Replication Recipe: What Makes for a Convincing Replication? *Journal of Experimental Social Psychology* 50:217–224. <https://doi.org/10.1016/j.jesp.2013.10.005> ; Crandall, C. S., & Sherman, J. W. 2016. On the Scientific Superiority of Conceptual Replications for Scientific Progress. *Journal of Experimental Social Psychology* 66:93–99. <https://doi.org/10.1016/j.jesp.2015.10.002> ; Derksen, M. et Morawski, J. (2022). Kinds of Replication : Examining the Meanings of "Conceptual Replication" and "Direct Replication". *Perspectives on Psychological Science*, 17(5), 1490-1505. <https://doi.org/10.1177/174569162111041116>

The Replication Crisis

A Challenge to Scientific Credibility

What is the replication crisis?

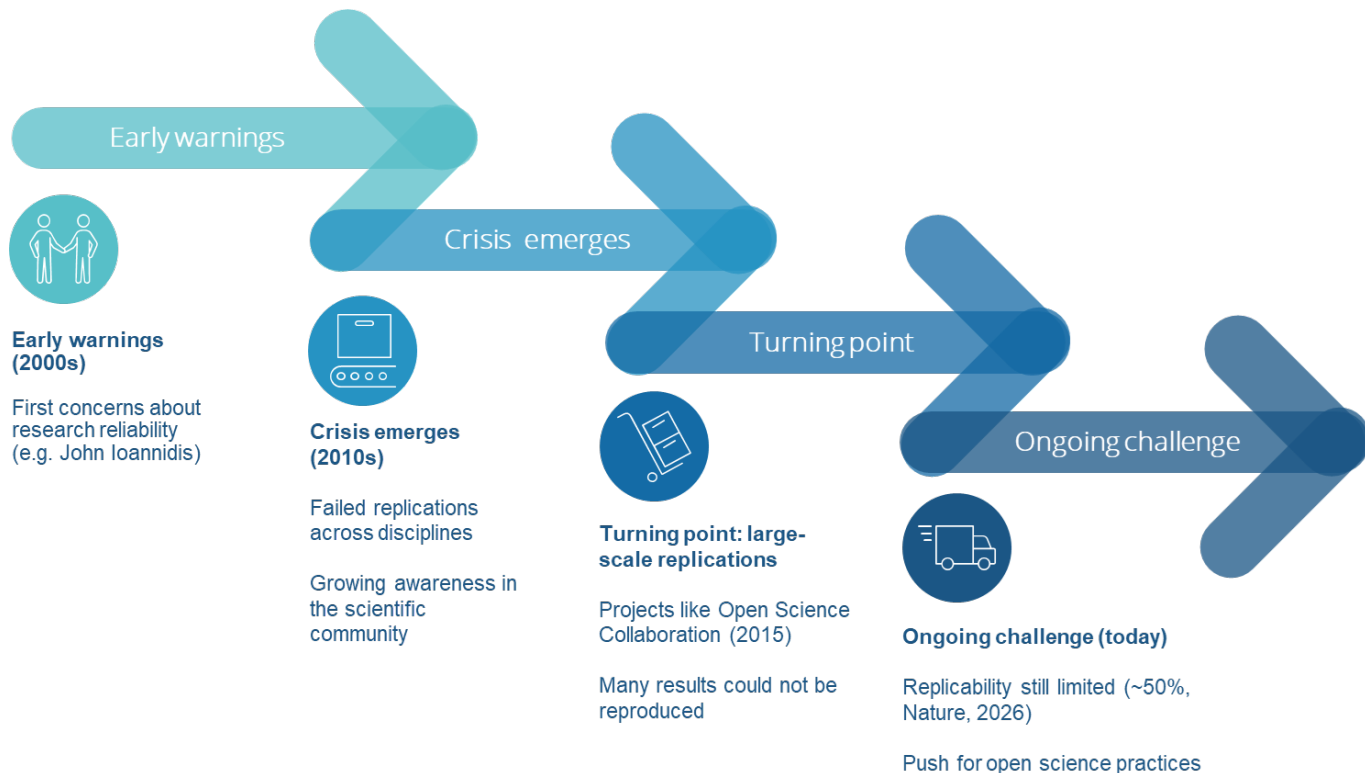


The *replication crisis* refers to a growing realization that many scientific results are **difficult or impossible to reproduce**.



Since reproducibility is **essential to the scientific method**, such failures challenge the **credibility of scientific knowledge** itself.

The Rise of the Replication Crisis




Ioannidis, J. P.. Why most published research findings are false. *PLoS medicine*, 2(8), e124. (2005). <https://doi.org/10.1371/journal.pmed.1004085>

Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* 349, aac4716(2015). <https://doi.org/10.1126/science.aac4716>

Tyner, A.H., Abatayo, A.L., Daley, M. *et al.*. Investigating the replicability of the social and behavioural sciences. *Nature* 652(8108), 143–150 (2026). <https://doi.org/10.1038/s41586-025-10078-y>

Building a culture of reproducibility

**Data citation &
Availability Statements**

 Discoverability of shared
data



Preregistration

Transparent design and reduced
bias

**Data
sharing**

Preprints

Open communication and early
feedback



**Replication & Code
Sharing**

Analytical reproducibility and
verification

Preregistration

Planning for Credibility



Definition

“Preregistration is the process of detailing research questions, variables, analysis plans, etc. before conducting research.” (Manago, 2023)

- **Time-stamped record:** captures decisions on study design, methods, and analysis.
- **Created Before Data Collection:** Established before data is collected or becomes accessible.
- **Public Availability:** Document should be accessible through a registry or repository.

Manago, B. Preregistration and Registered Reports in Sociology: Strengths, Weaknesses, and Other Considerations. *Am Soc* **54**, 193–210 (2023).

<https://doi.org/10.1007/s12108-023-09563-6>

A solution to the crisis



TRANSPARENCY



RIGOR



CREDIBILITY



INTEGRITY



PLANIFICATION



PUBLICATION BIAS

What and when?



What:

A timestamped document that contains :

- The hypothesis
- The methods
- The planned analysis and outcome
- Information about the participants



When:

Before the data collection (or in some case before accessing or analyzing research data)

How? Research workflow

- Standard



- With preregistration



Inspired by Henderson, E. L. (2022). A guide to preregistration and Registered Reports. <https://doi.org/10.31222/osf.io/x7aqr>

Where to preregister?

- Online dedicated platform providing a DOI
 - Wide range of research disciplines:



- Specific research Fields:



Preprint

Opening Research to Early Verification

Definitions

« **Preprints** are scholarly manuscripts posted by the authors to openly accessible platforms, usually before or in parallel with submission to a journal, that have not yet undergone peer review. » (Bourne et al., 2017)

- **Early version of a scholarly article:** shared before peer review.
- **Revisions:** Document may be updated based on feedbacks of the community.
- **Publication:** May or may not be submitted for journal publication.
- **Accessibility:** Shared on preprint servers, often open access and free to use.

Source: Bourne, P.E., Polka, J.K., Vale, R.D., & Kiley, R. *Ten simple rules to consider regarding preprint submission*. PLOS Biol 15, e2003995 (2017).

<https://doi.org/10.1371/journal.pcbi.1005473>

A solution to the crisis



TRANSPARENCY



OPENNESS



FEEDBACK



ACCESSIBILITY



PUBLICATION BIAS

What and when?



What:

An early version of an article

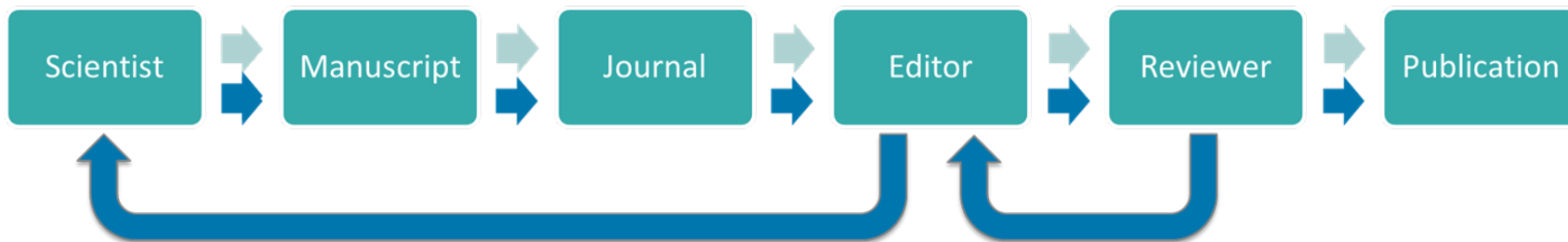
- That can be updated
- That does not necessarily undergo peer review
- That is open access
- That is on free of charge servers



When: Before an academic journal publication

How? Publication workflow

PUBLICATION WITH PEER REVIEW



PREPRINT



Where to submit a preprint?



Replication & analytical Reproducibility

Testing the Strength of Evidence

Definition (Recap)

	<u>Same data</u>	<u>Different data</u>
<u>Same Methods</u>	<u>Reproducibility</u> <u>Analytical or computational reproducibility</u>	<u>Replicability</u> <u>Direct replication</u>

A solution to the crisis



TRANSPARENCY



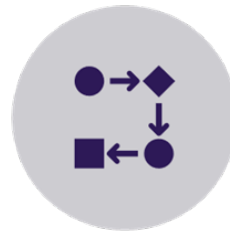
CREDIBILITY



VALIDATION



CODING INFORMATION



METHODOLOGICAL
PROGRESS

How to share the code? Good Practices

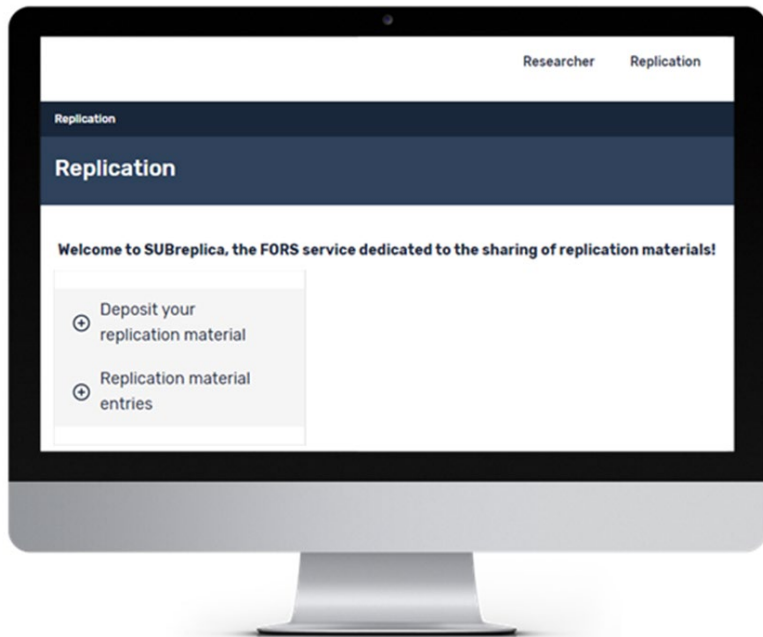


Where to share?



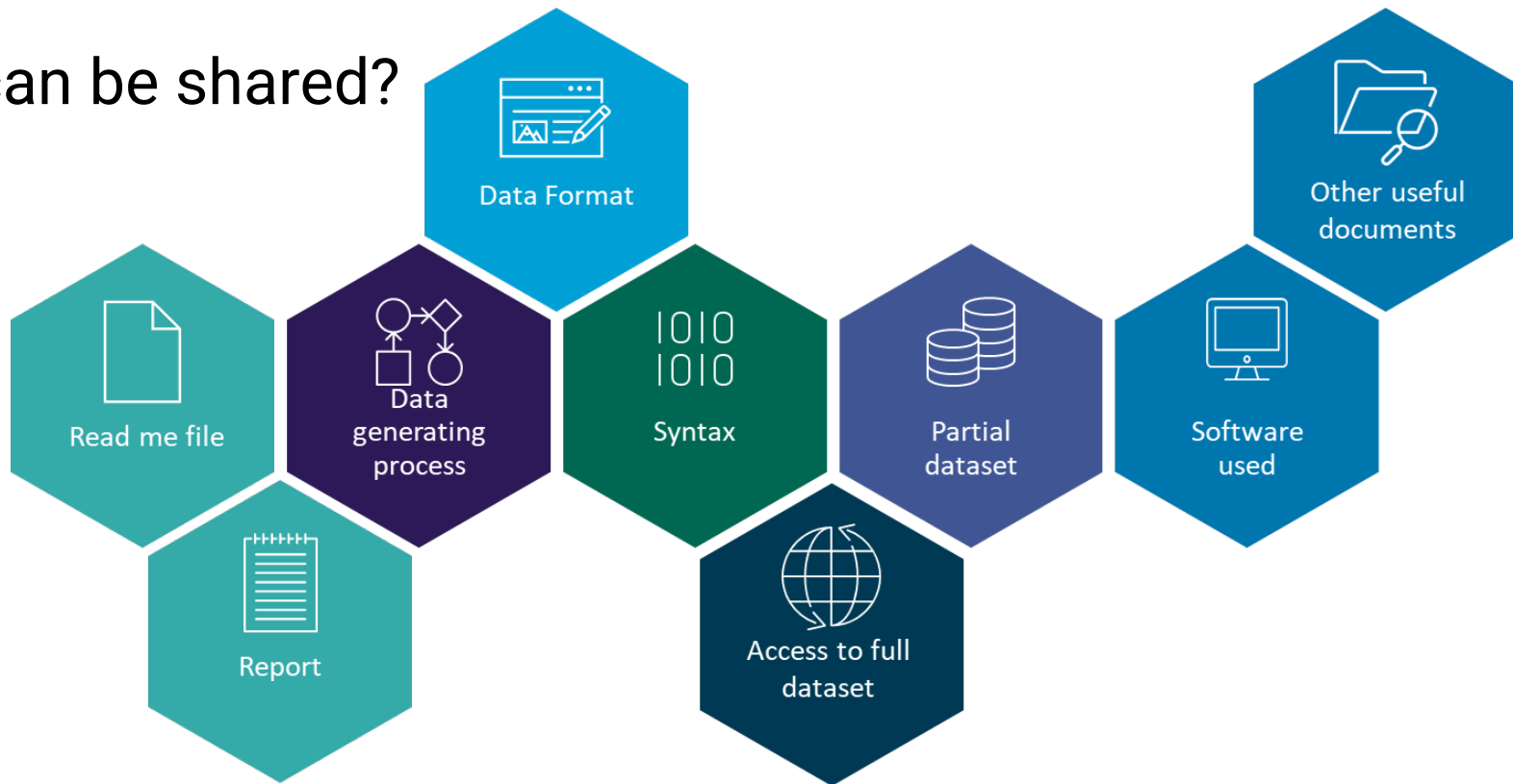
FORS Replication Service

What is FORS Replication Service for?



- It is a tool for the dissemination of replication materials related to publications.
- It enables uploading replication materials and obtaining a DOI.
- It provides access to scientific journals or any other interested people to replication materials.

What can be shared?

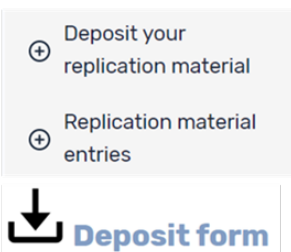


How? The deposit of replication materials in 4 easy steps



Download

Click on «Deposit your replication material» and download the deposit form



Complete

Download the deposit form and complete the metadata

Choose

Choose the creative commons license you want your replication material to be under



Send

Send the completed deposit form and all the replication materials to dataservice@fors.unil.ch

Data Citation & Data availability statements

Recognizing, Tracking and Reusing Data

What?



Data availability statement:

A short note explaining how and under what conditions the data underlying a publication can be accessed.



Data citation:

A standard reference similar to a regular bibliographic that points specifically to a dataset.

A solution to the crisis



TRANSPARENCY



CREDIBILITY



INTEGRITY



RIGOR



DISCOVERABILITY



REUSABILITY

How? Data Availability Statement

Display

- Place **outside the paywall**

Include

- State that data exists
- Where the data can be found
- Persistent identifiers (e.g. DOI)
- Repository or institution identifiers
- Licensing and access conditions



Example: *“The data supporting this study are available in [Repository Name] at [DOI or link], under [license/access conditions].”*

How? Data Citation

Minimum core components

- Author(s)
- Title of the data
- Publication year
- Version
- Data publisher
- Persistent identifier

Additional components

- Data number
- Resource type
- Place of publication
- Access date

**“data should be
considered legitimate,
citable products of
research”**

Force 11 Joint Declaration of
Citation Principles

Examples

APA (7th edition): Author(s) (Publication year). *Data title* (Data number; Version) [Resource type]. Publisher.

[Persistent identifier](#)

MLA (7th edition): Author(s). *Data title*. Version. Place of publication: Publisher, Publication year. Medium of publication. Date accessed. [Persistent identifier](#)

Source: Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 <https://doi.org/10.25490/a97f-egykh>

Resources on Data Citation

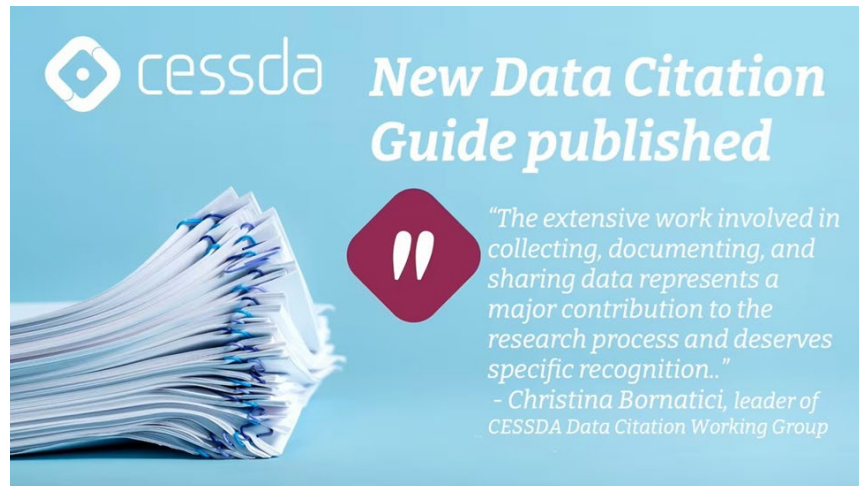
FORS⁺ GUIDES
 to survey methods
 and data management




Data Citation:
 How and Why Citing (Your Own)
 Data


Christina Bornatici¹  and Nicolas Fedrigo²

<https://doi.org/10.24449/FG-2023-00019>




cessda

New Data Citation Guide published


 "The extensive work involved in collecting, documenting, and sharing data represents a major contribution to the research process and deserves specific recognition."
 - Christina Bornatici, leader of CESSDA Data Citation Working Group

<https://datacitation.cessda.eu>

Source: Bornatici, C. & Fedrigo, N. (2023). Data Citation: How and Why Citing (Your Own) Data. FORS Guides, 19, Version 1.1, 1-15. <https://doi.org/10.24449/FG-2023-00019>

Source: Bornatici, C., Jernung, A., Alaterà, T. J., Tveit Sandberg, L., Strand, K., Štebe, J., & Trtiková, I. (2025). *CESSDA Recommendations on Data Citation: Practical Recommendations for Key Stakeholders*. <https://doi.org/10.5281/zenodo.15043854>

Why?

Advantages for the public



- Greater visibility of research data, during and after peer review
- Stable links between articles and data
- Support reproducibility, reuse and reduce waste

Advantages for researchers



- Credit for data producer
- Increase in citations
- Access to datasets for reuse or further research
- Compliance with the requirements of funders and institutions

Advantages for publishers



- Promote open science and reproducible research data
- Strengthen journal and repository collaborations
- Support authors in meeting funder requirements

Linked Open Data & Semantic Interoperability

Semantic Web

- Semantic Web extends the web by enabling data to be shared and reused across systems
 - Traditional web is human-readable
 - Semantic Web enables machine-readable data
- Tim Berners-Lee (Scientific American 284.5, 2001, S. 37)
«The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation.»

Semantic Web

- Core technologies
 - Standards such as Resource Description Framework (RDF), Web Ontology Language (OWL), SPARQL query language
- Publication of structured and semi-structured data in standardised formats on the internet
- Publication of data elements and their relationships on the internet – using Uniform Resource Identifiers (URIs) or International Resource Identifiers (IRIs) to identify data elements and their relationships
- Formalisation of the intended semantics through ontologies

Ontologies

- shared understanding of a domain is necessary to overcome terminological differences
- Terminological inconsistencies are common challenges
- **Common problems/inconsistencies**
 - 2 ontologies use the same term with different meanings
Example: ‘monument’ in architecture and archaeology
 - 2 ontologies use different terms but mean the same thing
Example: ‘Ptolemaic period’ vs. ‘Greek period of Egypt’

Ontologies

- **Possible solutions**
 - Use of a common ontology (e.g. dcterms, schema, CIDOC-CRM, fbroo)
 - Define direct mappings between different ontologies & data models
- Ontologies support semantic interoperability
 - Improving search accuracy
 - Generalisation/specialisation in web searches
- Ontologies can be interpreted by both humans and machines

Authority data

- Controlled vocabulary for the standardisation of names, terms and identifiers
- Purpose: To ensure consistency and accuracy in cataloguing and data retrieval
 - Unique identification: Distinguishing between entities with similar or identical names
 - Consistency: Uniform use of names and terms across different datasets
 - Standardisation: Standard formats for names, titles, etc.
 - **Interoperability: Enables collaboration between different systems**

Authority data

Examples

- Gemeinsame Normdatei ([GND](#))
- Identifiants et Référentiels pour l'enseignement supérieure et de la recherche ([IdRef](#))
- Virtual International Authority File ([VIAF](#))
- Library of Congress Name Authority File ([LCNAF](#))

Authority data

Examples: Getty Vocabulary

- Art and Architecture Thesaurus ([AAT](#))
- Iconography Authority ([IA](#)) Online
- Union List of Artist Names ([ULAN](#))
- Getty Thesaurus of Geographic Names ([TGN](#))
- Cultural Objects Name Authority® Online ([CONA](#))

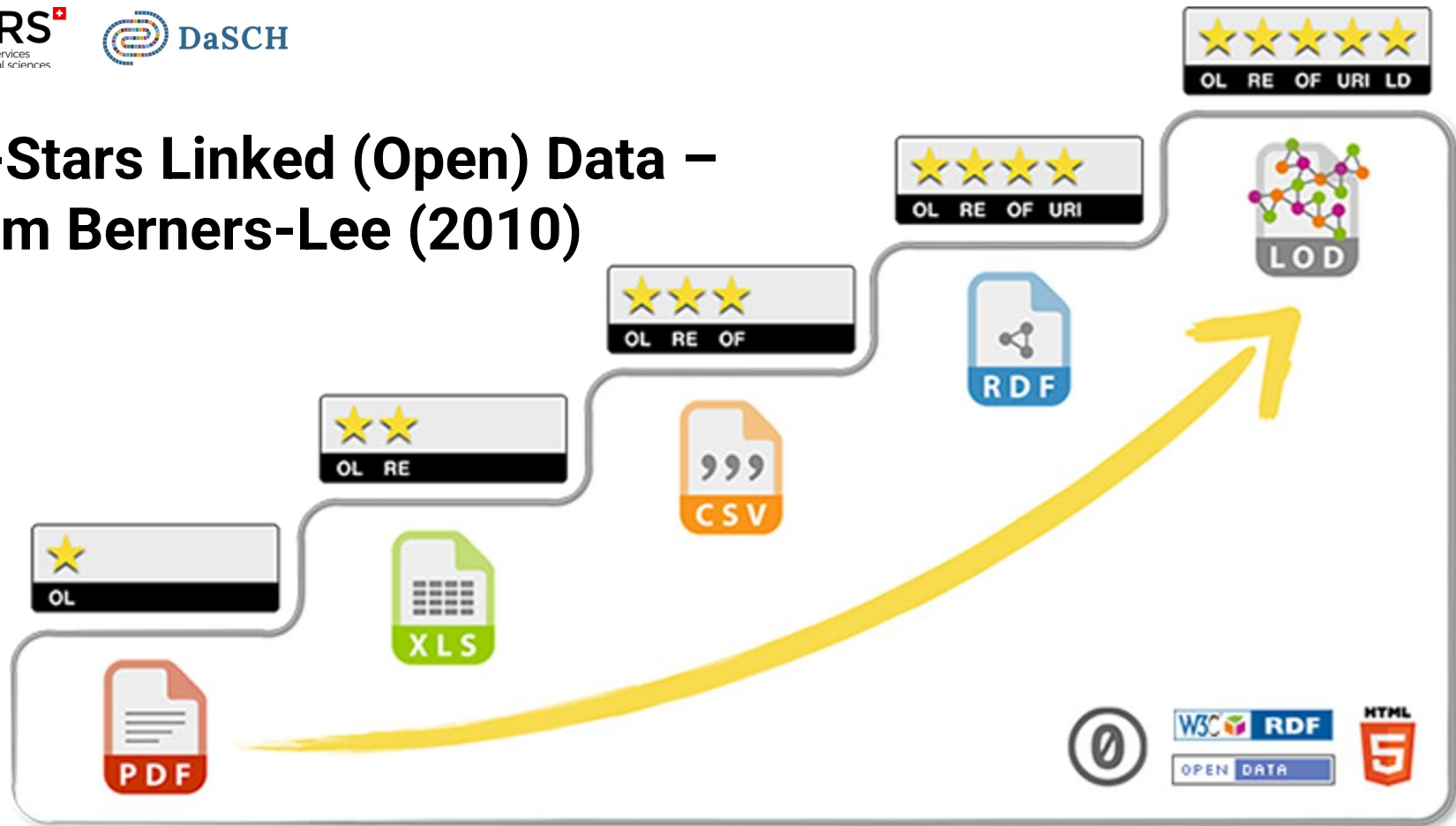
Linked (Open) Data

- **Linked data**
 - is structured data that is linked to other data
 - makes it more useful for semantic queries
- Term coined by Tim Berners-Lee (2006) in a design note (<https://www.w3.org/DesignIssues/LinkedData.html>)
- **Linked Open Data (LOD)** is structured data that is linked to other data and, furthermore, published under an open licence (CC0, CC BY)
 - Examples: [Wikidata](#), [DBpedia](#)

Linked (Open) Data

- **Key components of Linked (Open) Data**
 - URIs
 - HTTP
 - Structured data using controlled vocabularies and data model definitions, expressed in Resource Description Framework serialisation formats such as RDFa, RDF/XML, N3, Turtle or JSON-LD
 - Linked Data platform

5-Stars Linked (Open) Data – Tim Berners-Lee (2010)



Benefits

- Interoperability
- Discoverability
- Reuse
- Reproducibility

Applications

- Knowledge Graphs, Digital Humanities, Cultural Heritage

Important Standards



Use of standards: TEI

Text Encoding Initiative (TEI)

- standard for encoding texts in digital form, widely used in the humanities
- TEI is a set of guidelines for representing texts in Extensible Markup Language (XML)
- maintained by the TEI Consortium
- TEI enables generation of structured, machine-readable texts
- supports analysis, preservation, and reuse
- hierarchical and human-readable



Use of standards: TEI

Applications

- Digital editions
 - Manuscripts
 - Corpora
 - Cultural Heritage Data
-
- texts are encoded with tags that describe structure and meaning



Use of standards: TEI

Main parts

- `<teiHeader>` (metadata)
 - title
 - author
 - source
- `<text>` (content)
 - marks structure (e.g. chapters, paragraphs)
 - marks entities (e.g. names, dates, places)

```

TEI  teiHeader  fileDesc  titleStmt
1  <?xml version="1.0" encoding="UTF-8"?>
2  <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml" schematypens="http://relaxng.org/
3  <?xml-model href="http://www.tei-c.org/release/xml/tei/custom/schema/relaxng/tei_all.rng" type="application/xml"
4  schematypens="http://purl.oclc.org/dsdl/schematron"?>
5  <TEI xmlns="http://www.tei-c.org/ns/1.0">
6  <teiHeader>
7  <fileDesc>
8  <titleStmt>
9  <title>Title</title>
10 </titleStmt>
11 <publicationStmt>
12 <p>Publication Information</p>
13 </publicationStmt>
14 <sourceDesc>
15 <p>Information about the source</p>
16 </sourceDesc>
17 </fileDesc>
18 </teiHeader>
19 <text>
20 <body>
21 <p>Some text here.</p>
22 <div>
23 <p>A list:</p>
24 <list type="bulleted">
25 <item>First item</item>
26 <item>Second item</item>
27 <item>Third item</item>
28 </list>
29 </div>
30 <div>
31 <p>A table:</p>
32 <table>
33 <row>
34 <cell role="label">first column</cell>
35 <cell role="label">second column</cell>
36 </row>
37 <row>
38 <cell>value 1</cell>
39 <cell>value 2</cell>
40 </row>
41 </table>
42 </div>
43 </body>
44 </text>
45 </TEI>

```

TEI



Use of standards: IIIF

- International Image Interoperability Framework (IIIF)
- <https://iiif.io/>
- international open standards
- frequently used in the field of cultural heritage, particularly by museums, libraries, archives and research institutions
- several APIs that enable advanced image processing, presentation and data exchange



Use of standards: IIF

Image API

- provides standardised access to images hosted on web servers
- users can request images in various sizes, formats, orientations and regions

Use of standards: IIF

Examples:

- original image: <https://iiif.dasch.swiss/0803/H2Kye390dOP-FJpCoZWaF9C.jp2/full/5451,7062/0/default.jpg>
- A square image crop of 2000x2000px, rotated by 90°:
<https://iiif.dasch.swiss/0803/H2Kye390dOP-FJpCoZWaF9C.jp2/square/2000,2000/90/default.jpg>

Dem all and die diener Daz zu waum er such sin
angebornen kind wart Dazumb ez in scote und in
fir hielte Das er von allem sinem haffysind und
sinen freunden und gomey, geyet wart



Dus wiles Do graff künich nach siner ge
wrouheit off einem geyget wart Da die
sinen luyten ein waldes forren vor den hunden
hin und zoch das gams geogt hin noch gye sich
Dazumb der vorgeant graff der geyget schnell
nortgliche uff sinem fstraden und vorgeant
und alle in Rymont gye schnell nach umb das
er sinen vettern mit vlinz in dem walds und
hief sich der walds von Columbier und er wart
spot am tag Das se by dem monesteyn Keten
in dem walds und alle dem geygte nach Das



Dus wiles Do graff künich nach siner ge
wrouheit off einem geyget wart Da die
sinen luyten ein waldes forren vor den hunden
hin und zoch das gams geogt hin noch gye sich
Dazumb der vorgeant graff der geyget schnell
nortgliche uff sinem fstraden und vorgeant
und alle in Rymont gye schnell nach umb das
er sinen vettern mit vlinz in dem walds und
hief sich der walds von Columbier und er wart
spot am tag Das se by dem monesteyn Keten
in dem walds und alle dem geygte nach Das

Fir hielte Das er von allem sinem haffysind und
sinen freunden und gomey, geyet wart



Use of standards: IIF

Presentation API

- describes the structure and metadata of digital objects such as images, books or manuscripts using JSON
- enables the creation of complex, multi-page objects



Use of standards: IIF

Images

- detailed zooming, comparisons, structure and annotations are possible
- <https://ark.dasch.swiss/ark:/72163/1/0838/rS2oEITHTbi8HTWwJz3QFA1>
- <https://ark.dasch.swiss/ark:/72163/1/0844/=Ws=IKGvT5ShPfiWTjK71wV>

Use of standards: IIF

Audios and videos

- several film reels, together with subtitles, transcripts, translations and notes
- Cookbook: <https://iiif.io/api/cookbook/> (IIF Manifest)
 - <https://iiif.io/api/cookbook/recipe/0009-book-1/>
 - <https://iiif.io/api/cookbook/recipe/0017-transcription-av/>
 - <https://iiif.io/api/cookbook/recipe/0219-using-caption-file/>



Use of standards: IIF

Other IIF APIs

- [Content Search API](#) – to search annotations
- [Authorization Flow API](#) – to control access
- [Change Discovery API](#) – for aggregators to retrieve IIF data
- [Content State API](#) – improvement to the method of transferring content from one viewer to another



Use of standards: IIIF

Extensions

- [Maps Extension](#) – to link a geographical location to an IIIF object (georeferencing)

Examples:

- <https://iiif.io/api/cookbook/recipe/0240-navPlace-on-canvas/>
- <https://iiif.io/api/cookbook/recipe/0154-geo-extension/>

Example of use: Getty Museum Collection

<https://www.getty.edu/art/collection/object/103VQF#full-artwork-details>

Respective IIF Manifest:

<https://media.getty.edu/iiif/manifest/ee9d8817-4a38-4d9a-94bd-08c028d32361>

The file does not need to be stored locally for IIF-compatible viewers (e.g. Mirador, Universal Viewer) – the URL of the IIF Manifest is sufficient!

<https://universalviewer.io/uv.html?manifest=>

e.g.

<https://universalviewer.io/uv.html?manifest=https://media.getty.edu/iiif/manifest/ee9d8817-4a38-4d9a-94bd-08c028d32361>

Finger Ring with the Ambush of Achilles

550–500 B.C.

Unknown artist/maker

Not currently on view

[View full record details](#)

An intricate scene decorates the bezel of this Etruscan ring. Two men approach a fountain where water gushes into a vessel from a lion's head spout. Behind the fountain, a man squats as if hiding, holding a sword. These details identify the scene as a standard depiction of the ambush of Troilos, prince of Troy, by the Greek hero Achilles during the Trojan War. On this ring, however, a strange dog-headed creature, who is not part of the Troilos myth, sits atop the fountain. The creature may actually be jackal-headed and thus meant to recall the Egyptian god Anubis.

All Greek and Etruscan metal rings with engraved bezels ultimately derive from Egyptian and Phoenician cartouche-



[Public Domain](#)

3 images



Front



Impression



Profile



APIs & other identifiers

The reference information listed below is intended for those who work with the Getty's data.

API URL:	object/4885977c-5bd4-49a9-a8b7-a658f93f1352 ↗	🔗
	The entity identifier in the Museum Collection API	
TMS ID:	13001	🔗
	The identifier in the Museum's internal collection management system	
DOR ID:	10997	🔗
	The identifier in the Museum's legacy internal digital object repository	
IIIF:	iiif Manifest URL (v2) ↗	🔗
	iiif Manifest URL (v3) ↗	🔗
	Open in comparison viewer ↗	
	The International Image Interoperability Framework, or IIIF, is an open standard for delivering high-quality, attributed digital objects online at scale. Visit iiif.io to learn more.	
SPARQL API:	Open in Query Builder ↗	
	The SPARQL API provides an endpoint and interface for constructing graph queries of the Museum Collection Linked Open Data	
Related IDs:	Internal Identifier Relationships ↗	
	The ID management system records relationships between identifiers across Getty's systems and resources	



Example of use: IIF 3D

- IIF standard for 3D is still under development
- Image and Presentation API 4.0 (release June 2026) include the first implementations

Example: IIF Manifest

<https://gautschr.github.io/3D-modelling/iif-manifests/mitreo/mitreo.json>

View via Universal Viewer:

<https://uv-v4.netlify.app/#?manifest=https://gautschr.github.io/3D-modelling/iif-manifests/mitreo/mitreo.json&c=&m=&cv=&xywh=>



Example of use: IIF 3D

- Instructions on how to deploy a 3D model using GitHub Pages, create IIF manifests, and display both: <https://gautschr.github.io/3D-modelling/>
- Test how it looks in different 3D viewers:
<https://morphosource.github.io/manifestival/>
- Example IIF 3D Manifests:
 - https://gautschr.github.io/3D-modelling/iiif-manifests/St._Gallus_Kirche/St._Gallus_Kirche.json
 - <https://gautschr.github.io/3D-modelling/iiif-manifests/pharos/pharos.json>
 - <https://gautschr.github.io/3D-modelling/iiif-manifests/zytglogge/zytglogge.json>

Viewer: Aleph + UV (MorphoSource) ▾

Mithraeum in the Baths of Mithras in Ostia



ATTRIBUTION



Provided by Rita Gautschy



Load Manifest Into Viewer

```
items : [
  {
    "id": "https://github.com/gautschr/3D-modelling/edit/main/iif_manif
    "type": "Scene",
    "items": [
      {
        "id": "https://github.com/gautschr/3D-modelling/edit/main/iif_ma
        "type": "AnnotationPage",
        "items": [
          {
            "id": "https://github.com/gautschr/3D-modelling/edit/main/iif_n
            "type": "Annotation",
            "motivation": ["painting"],
            "body": {
              "id": "https://raw.githubusercontent.com/gautschr/3D-modellir
              "type": "Model",
              "format": "model/gltf-binary"
            }
          }
        ]
      }
    ]
  }
]
```

Select an Example Manifest ▾

<https://gautschr.github.io/3D-modelling/iif-mar>

Load

Save the Date!

Eighth Webinar

Topic **Copyright, Authorship & Licenses**

When **September 29, 2026 at 2pm**

Where **Online**



Questions?

