



Fostering a Culture of Data Citation for High-Quality Research

Dr Christina Bornatici (FORS, CESSDA)
Dr Farah Karim (GESIS, CESSDA), presenter
& Tuomas J. Alaterä (FSD, CESSDA)

Based on joint contributions of the CESSDA Data Citation Working Group

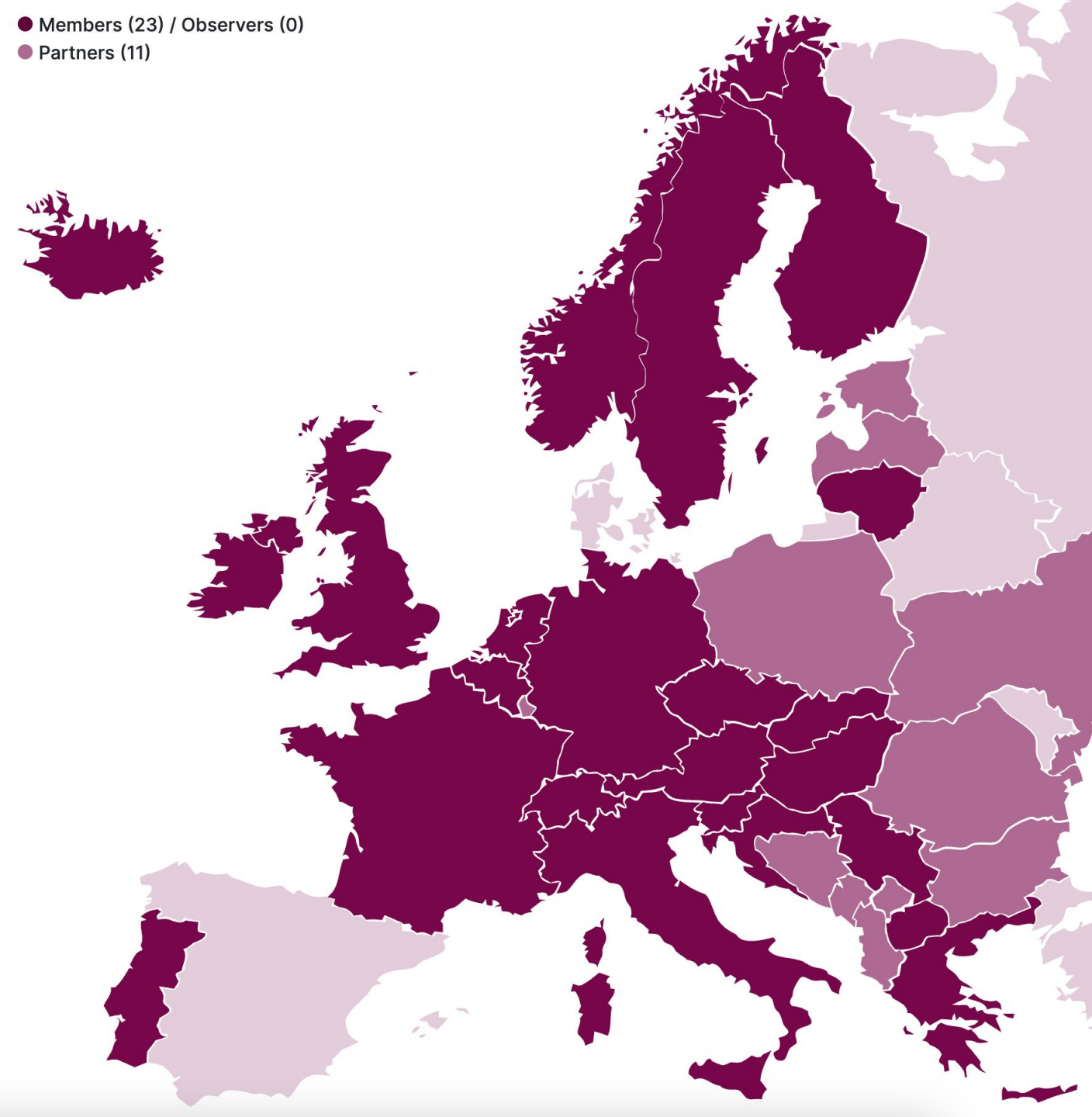
PUBMET Conference 2025, Zagreb, Croatia | 12 September 2025

Plan

1. **Who we are**
 - CESSDA and the Data Citation Working Group
2. **Why data citation matters**
 - Importance of data and data citation for high-quality research
3. **What we recommend**
 - Core elements of data citation
 - Guidance for key stakeholders



● Members (23) / Observers (0)
● Partners (11)



CESSDA (Consortium of European Social Science Data Archives)

A European Research Infrastructure Consortium (ERIC) bringing together national social science data archives across Europe

Its role is to:

- Ensure that data are preserved, curated, and reusable, fostering trust in research
- Facilitate cross-border access to high quality social science research data
- Provide integrated and sustainable data services for the social sciences
- Foster collaboration among data infrastructures, researchers, and decision-makers

CESSDA Data Citation Working Group



- Brings together open and FAIR research data experts from across the CESSDA consortium
- Aims to **advance data citation practices** in the social sciences
- Develops practical guidance and resources to support effective implementation and promotion of research data citation
- Participating data archives and repositories:

FORS - Switzerland (lead)

ADP - Slovenia

CSDA - Czech Republic

FSD - Finland



GESIS - Germany

Sikt - Norway

SND - Sweden

So.Da.Net - Greece

UKDS - United Kingdom



Data: Essential but Undervalued

- **Core role:** Data constitute the foundation of analysis, the basis for generating findings, and the prerequisite for reproducibility. Yet, they remain undervalued as independent and legitimate research outputs
- **Invisible labour:**
 - Creating and sharing data requires careful collection, management, and documentation by researchers
 - Data archives and repositories with curation processes then ensure data reliability and long-term accessibility
 - Still these efforts to produce and publish high-quality data receive little academic recognition
- **Consequences:** Without recognition and citation, data risk being underused and duplicated, creators lose credit, data sharing is discouraged, and research loses transparency and reproducibility.

Challenges in Data Citation



- **Low awareness**

Many researchers, and even academic publishers, are unfamiliar with how to cite data correctly

- **Publishing barriers**

Citation requirements are often absent or applied inconsistently across academic publishers



- **Metadata gaps**

Missing components – e.g., version of the data or persistent identifiers – limit discoverability and reuse

- **Weak incentives**

Academic assessments prioritise publications over data outputs

Why Data Citation Matters



- **Recognition**

Data citation gives **visibility** to data and their authors, treats data as legitimate research outputs, and **attributes credit** to data authors

- **Transparency**

Data citation facilitates verification, replication, and traceability of findings, while reinforcing trust and **scientific integrity**

- **Reuse**

Data citation is a **FAIR-compliant** practice, ensuring data are findable, accessible, interoperable, and reusable in new research

- **Impact**

Data citation values data collection and sharing, enables **tracking of data use**, and supports new indicators for Open Science and research assessment, with careful attention to disciplinary and data-type differences, and qualitative assessments (e.g., CoARA)



Driving Change in Data Citation Practices

- Achieving **high-quality research** is not just about robust methods and transparent reporting
- It also requires a **cultural shift**: *"if you use it, you cite it"*
- Recognition for data authors = recognition for **intellectual contribution**
- **Change** happens when **policy** and **culture** overlap, driving new research practices



Rules, mandates,
infrastructure, formal
requirements

Norms, values,
attitudes, behaviours



CESSDA Data Citation Guide



CESSDA Data Citation Guide

Why Cite Data?

The Anatomy of a Data Citation

Recommendations for Data Repositories

Recommendations for Authors

Recommendations for Journals and Publishers

Recommendations for Research Performing Organisations (RPOs)

Recommendations for High-Level Research Policy and Strategy Development Entities

Recommendations for Ethics Committees

References

CESSDA Data Citation Working Group

Go

Welcome to the CESSDA Data Citation Guide!

Citing data is essential to ensuring transparency, reproducibility, and recognition in research. Just as scholarly articles are cited, datasets must be acknowledged to give proper credit to data authors and to support the discovery, access, and reuse of research data.

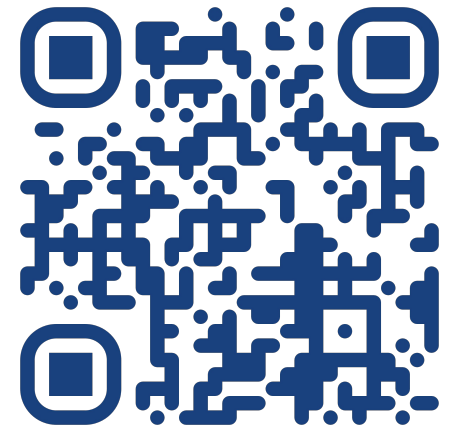
At [CESSDA](#), we advocate for **best practices in data citation** across the social sciences. This guide is designed to **support the entire research ecosystem**—from individual researchers to policymakers—in fostering a more coherent, FAIR-aligned, and sustainable data citation culture.

What you will find in this guide

The Data Citation Guide begins by explaining [why data citation matters](#) and [what a data citation should include](#).

It then offers **targeted and practical recommendations** for different stakeholder groups:

- ◆ [Data repositories](#)
- ◆ [Authors](#) of publications mentioning and/or using data
- ◆ [Academic journals and publishers](#)
- ◆ [Research performing organisations](#) (RPOs)





Data Citation Ecosystem: Key Stakeholders and Roles



- > Provide citation elements
- > Enable interoperability
- > Raise awareness

Data Repositories

- > Cite data
- > Support reproducibility and recognition

Authors

- > Set policies and standards
- > Foster scholarly communication

Journals and Publishers

Research Performing Organisations

Research Policy and Strategy Development Entities

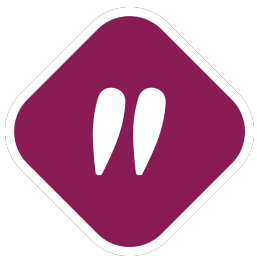
Ethics Committees

- > Promote data citation via policies and training

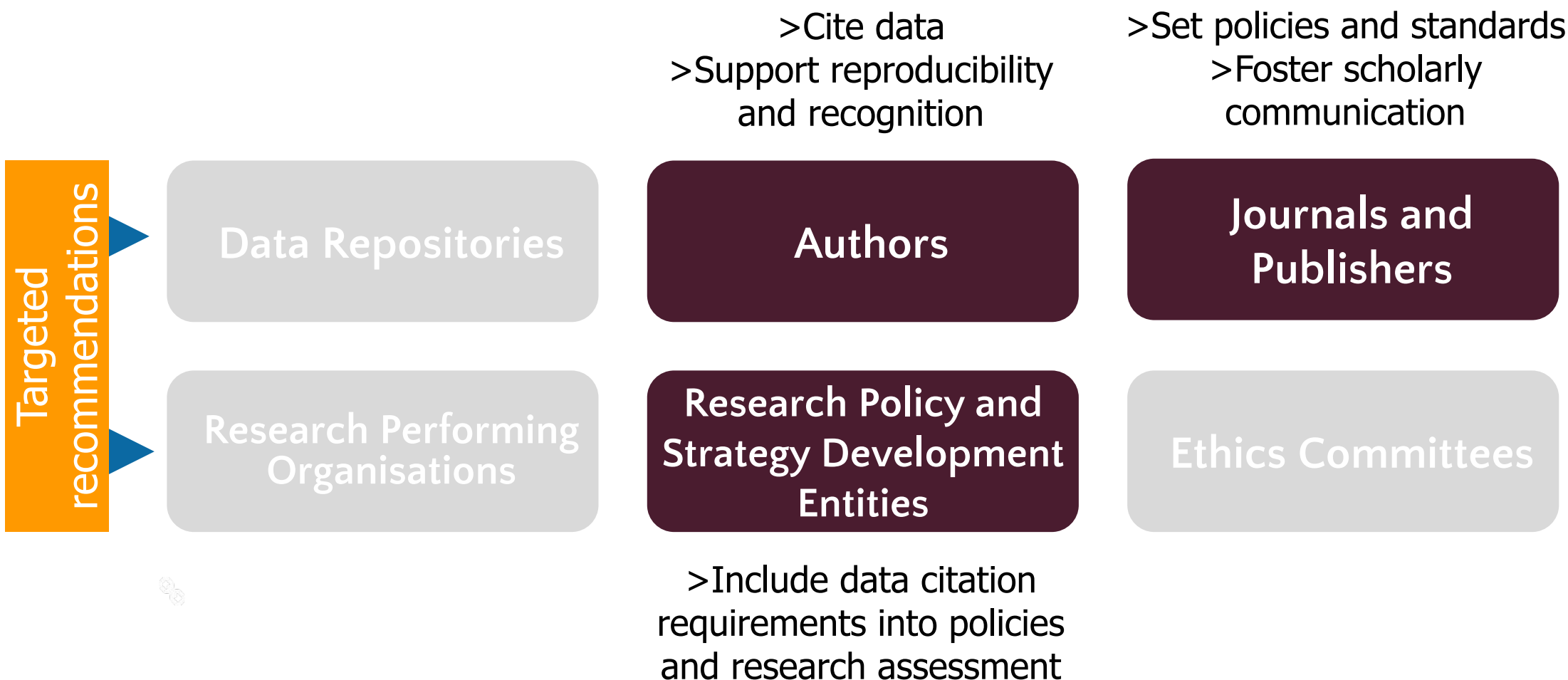
- > Include data citation requirements into policies and research assessment

- > Ensure compliance with ethics and citation standards

Targeted recommendations



Data Citation Ecosystem: Key Stakeholders and Roles





General Recommendation for All Stakeholders: Definition and Anatomy of a Data Citation



What is a data citation?

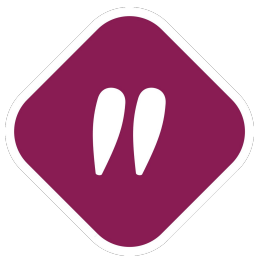
- A reference to the data themselves
- Placed in-text (author-date system) or as a footnote (notes system) whenever data are mentioned
- Listed in the reference list, like any other scholarly source

Six core components

1. Data author(s)
2. Title of the data
3. Year of publication
4. Version
5. Data publisher
6. Persistent identifier (e.g. DOI)

Additional components

- Accession or dataset number
- Resource type (e.g. [Dataset])
- Place of publication



Data Citation in Action



Full reference following [APA 7th](#) bibliographic style:

Data author(s) (Publication year). *Data title* (Data number; Version)
[Resource type]. Data publisher. Persistent identifier

Examples:

ISSP Research Group (2025). *International Social Survey Programme: Family and Changing Gender Roles V - ISSP 2022* (ZA10000; Version 1.0.0) [Dataset]. GESIS.
<https://doi.org/10.4232/5.ZA10000.1.0.0>

Mach, A., Ginalski, S., Davoine, E., Leimgruber, M., Strebel, M. A., Piguet, S., Araujo, P., & Benz, P. (2025). *OBELIS Swiss Local Elites: Basel, Geneva and Zurich* (2808; Version 1.0) [Dataset]. FORS. <https://doi.org/10.48573/b9p3-rz53>





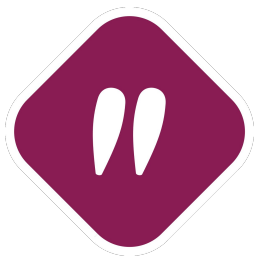
Authors



1. **Cite the data themselves**, citing associated report or article is not sufficient, nor declaration of availability statement
2. Include all **six core components** in the reference list
3. Cite **primary and secondary data** as well as **replication materials**



Explicitly cite all data used



Journals and Publishers



1. Allow data citations in the **standard list of references**
2. **Require** authors to cite data specifically
3. Include all **six core components** in data citations and provide citation examples
4. Encourage authors to **publish data** and replication materials as independent publications
5. Encourage authors to provide **PIDs** for published data
6. Allow **time for quality assurance** and data curation by data repositories



Allow proper citing of data as an independent output



High-Level Research Policy and Strategy Development Entities



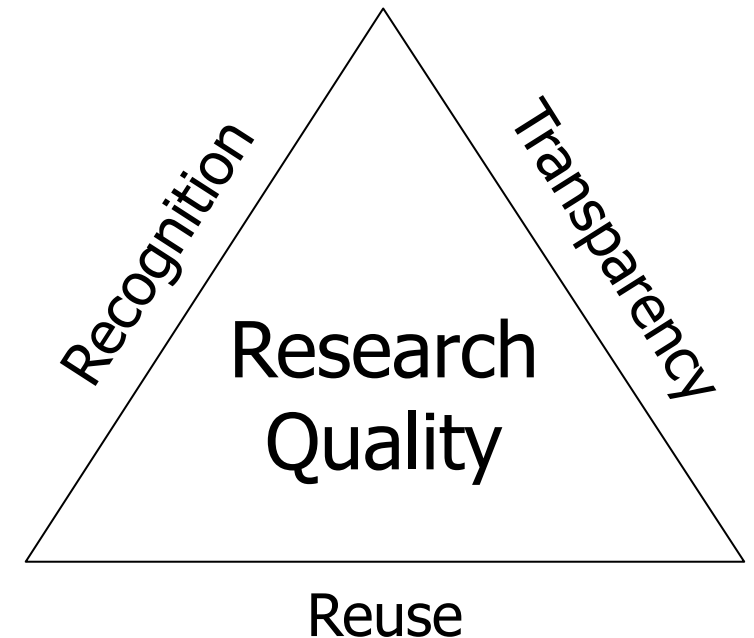
1. **Require** data to be cited
2. Include all six core components in guiding policies
3. Make the **benefits** of data citation clear in policies
4. Require **active reporting** of data publication outputs through research infrastructures



Require citing and reporting of data publications

Conclusion: Data Citation Enhances Research Quality

- **Recognition**
 - Credit for data authors
 - Value data as legitimate research outputs
- **Transparency**
 - Strengthen reproducibility and openness
- **Reuse**
 - Enable data use beyond the original study, maximising impact



Together, we can foster a culture of data citation!



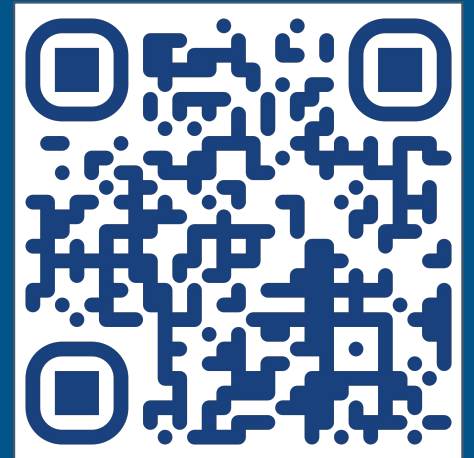
CESSDA Data Citation Guide

datacitation.CESSDA.eu

This guide is based on the publication:

Bornatici, C., Jernung, A., Alaterä, T. J., Tveit Sandberg, L., Strand, K., Štebe, J., & Trtíková, I. (2025). *CESSDA Recommendations on Data Citation: Practical Recommendations for Key Stakeholders*.

<https://doi.org/10.5281/zenodo.15043854>





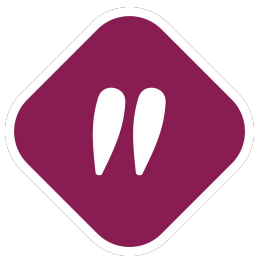
Thank you for your attention!

Farah.Karim@geis.org

Christina.Bornatici@fors.unil.ch

www.cessda.eu

Illustrations CC0 pxhere.com



Data Repositories



1. **Request data to be cited**
2. Provide all six core data citation components
3. Provide data **versioning and access information**
4. Provide **citation suggestions** including the core components
5. Facilitate data citation through export and download of metadata for reference management software
6. Make citation information and other publication metadata accessible through generic **structured metadata** files
7. Make data publication and **relations** to other entities machine-actionable through rich metadata and **PIDs**
8. Raise **awareness** about data citation



Provide core components and foster interoperability



Research Performing Organisations



1. Encourage data citation practices
2. Encourage authors to **report data publications** as research outputs
3. Include data publication **metadata in indexing** services
4. Include and leverage the **added value of metadata** describing data publications in administrative systems
5. Take data publications into account when summarising outputs



Index and take advantage of data publication metadata

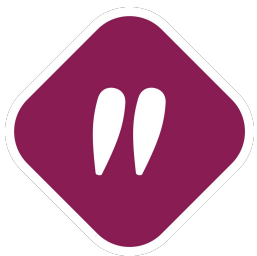


Ethics Committees

1. All research should follow **good citation practice**
2. Even **restricted** data should be **citable**
3. Address data citation in data management plans



Citing the work of others as a research ethical value



Main Goals of the Guide



Not just “data citation is important”, but how to make it work in scholarly communication

Our focus:

- **Enable and facilitate citation** - clear practices for authors, publishers, repositories
- **Core requirements** - metadata completeness, persistent identifiers, versioning
- **Integration into publishing workflows** - policies, submission systems, evaluation criteria
- **Interoperability and FAIRness** - make data citations machine-actionable and linkable across systems
- **Culture & awareness** - engage journals, funders, and institutions in promoting recognition of data as a scholarly output

