



Zoom-Austausch-Runde: "Forschung in Corona-Zeiten"

21.10.2020

Different types of data

- Janet Salmons (2016) and others
 - Enacted data
 - Elicited data
 - Extant data
- Robert Groves (2011)
 - Designed data
 - Organic data

Reactive data vs *non-reactive* data: an example

- „*Expressions*“ produced without mediation of a researcher

For example social media:

- In recent years, social media have become an important infrastructure for communication flows and thus an essential network in our social structure. Within that network participants are actively involved in expressing sentiments and perceptions
- Perceptions and sentiments are related to the behavior of the people expressing them

Data collection changes the phenomenon....

- Influence of presence of the researchers
- Influence of the data collection process
 - For example: questionnaire about attitudes changes these attitudes...
 - ...
- The phenomenon you have researched is not the same anymore after your data collection....
- Solution: unobtrusive data collection

“Researching perceptions and sentiments through a survey, for example, is inherently different from the analysis we have done on social media data. The expressions in social media data can be genuinely defined as “texts”: they are “words and images that have become recorded without the intervention of a researcher” (Silverman 2000, 825). In a survey or an interview, the researchers’ preconceptions always strongly influence the categories of topics that are revealed. When the data analyzed are “texts” as defined by Silverman, it is much more likely that original participants’ categories will be discovered.”

Torres van Grinsven & Snijkers, 2015, p. 284-286

Types of non-obtrusive data collection or data collection of non-reactive data

- Unobtrusive observation (online or not)
- Documents or archives (online or not...)
 - Movies
 - Video clips
 - Posts on social media
 - Discussions on webfora
 - Grafitti
 - Pamflets
 - Articles in newspapers
 - Diaries, narratives (....)

Types of non-obtrusive data collection or data collection of non-reactive data

Sage Research Methods:

- Video research
- Visual research
- Virtual ethnography
- (Internet research)
- Naturally-Occurring Data
- Innovative methods
- (...)

- Methodspace: <https://www.methodspace.com/>
- Online qualitative research:
<https://www.methodspace.com/online-qualitative-research-q-a/>
- ...

Reactive vs non-reactive data

- „Expressions“ or behaviors produced without the mediation or intervention of a researcher
- Reality is not changed by the data collection (!)

„E-research“ or online research / ict mediated research

- Online interviews (...)
- Online surveys
- Social media research
- (Narratives)
- ✓ (Unobtrusive) observation
- ✓ Document analysis
- ❖ Video research
- ❖ Visual research
- (...)

Other types of research without physical contact

- Telephone interviews
- Questionnaires on paper / paper surveys
- (Narratives)
- ✓ (Unobtrusive) observation
- ✓ Document analysis
- ❖ Video research
- ❖ Visual research
- (...)

COMMUNICATING ONLINE FOR RESEARCH PURPOSES

ICTs as **MEDIUM**
for data collection.



Computer-mediated communications between researcher and participant are used to investigate any aspect of the lived experience—online or in person.

ICTs as **SETTING**
for data collection.



ICTs are the electronic milieu for research-- the place where interviews, observations or other kinds of data collection can occur.

ICTs as **PHENOMENON**



Use, activities or behaviors with ICTs are part of the research phenomena the study is designed to investigate.

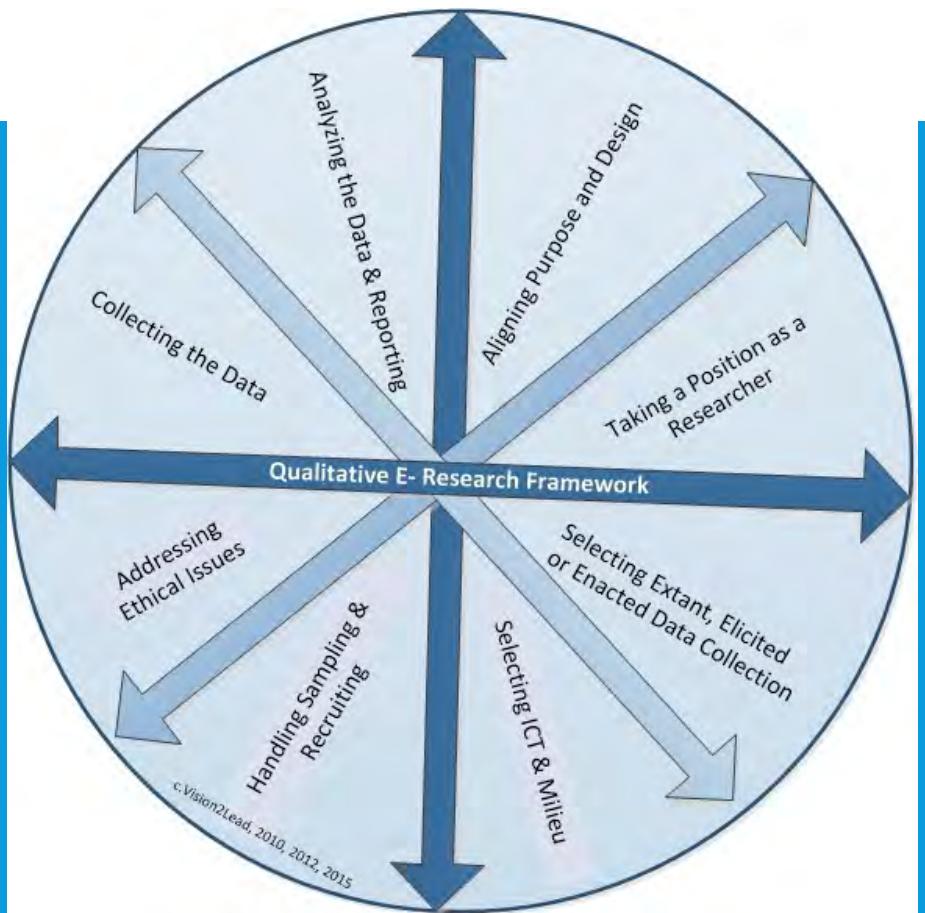


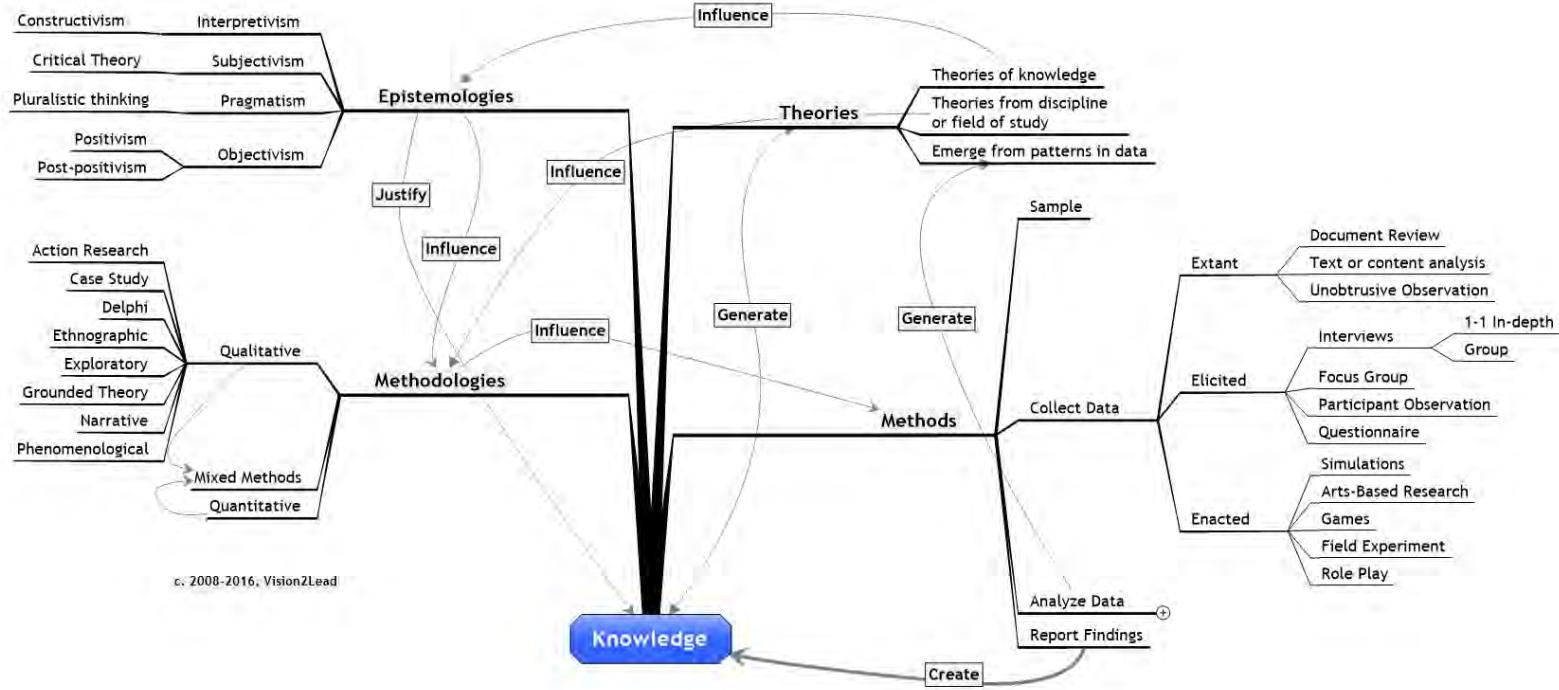
INTRODUCTION TO THE QUALITATIVE E-RESEARCH FRAMEWORK

Janet Salmons, PhD

USING THE QUALITATIVE E-RESEARCH FRAMEWORK

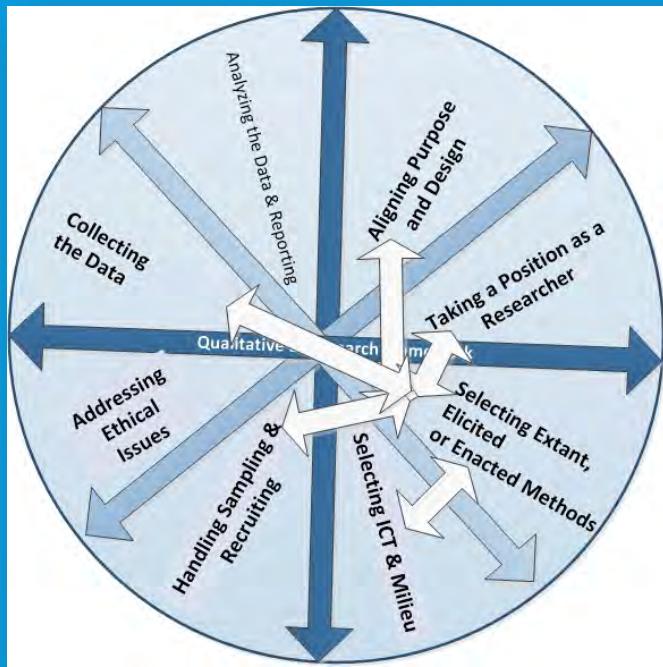
Holistic design means considering relationships





ALIGN PURPOSE AND DESIGN

COLLECTING DATA ONLINE & THE QUALITATIVE eRESEARCH FRAMEWORK



While the researcher may be deeply engaged with the activities associated with data collection, it is important to remember how this phase fits into the overall inquiry.

Links databases secondary data

- <https://www.hf.uni-koeln.de/41238?lang=2>
- Qualitative and quantitative data
- Data-repositories and datasets...

Carsten Schwemmer Dataverse

Harvard Dataverse > Carsten Schwemmer Dataverse >

Replication Data for: Diagnosing Gender Bias in Image Recognition Systems

Version 2.0

Schwemmer, Carsten, 2019, "Replication Data for: Diagnosing Gender Bias in Image Recognition Systems", <https://doi.org/10.7910/DVN/2CEYWV>, Harvard Dataverse, V2, UNF:6:9GUYlhBEYZwglD2UOoln0w== [fileUNF]

Cite Dataset | Learn about Data Citation Standards.

Access Dataset | Contact Owner | Share

Dataset Metrics | 11 Downloads

Description: This repository contains data and replication code for the article "Diagnosing Gender Bias in Image Recognition Systems", published in the journal *Socius*. (2020-09-29)

Subject: Computer and Information Science; Social Sciences

Keyword: gender, image recognition, computational social science, bias, stereotypes

Files | Metadata | Terms | Versions

Terms of Use

Waiver: Our Community Norms as well as good scientific practices expect that proper credit is given via citation. Please use the data citation above, generated by the Dataverse.

CC0 - "Public Domain Dedication" | PUBLIC DOMAIN

https://dataverse.org/best-practices/data-citation

Dataverse Project About Community Best Practices Software Contact

Academic Credit Data Citation Dataverse Community Norms Data Management Replication Dataset Guidelines

HOME / Data Citation

Your Publication Your Data

Formal Data Citation

The diagram shows a central red double-headed arrow labeled "Formal Data Citation". To the left of the arrow is a stack of four scientific journals: "PS", "Cell", "nature", and "Turkana tools". To the right of the arrow is a screenshot of a computer screen displaying a large grid of data rows, representing a dataset.

Dataverse standardizes the citation of datasets to make it easier for researchers to publish their data and get credit as well as recognition for their work. When you create a dataset in Dataverse, the citation is generated and presented automatically. As an open source framework and research data repository, Dataverse is committed to helping researchers, journals, and organizations make scientific data accessible, reusable, and open (when possible), which includes implementing community accepted standards for data publication (Altman & Crosas 2013). For nearly 20 years, members of IQSS and its Data Science team, who work on Dataverse, have played an active part in the the work to standardize data citation (King 1995, Altman & King 2007, Altman & Crosas 2013). Illustrated in the figure below, is an example of how the data citation is formulated in Dataverse, using the [Joint Declaration of Data Citation Principles \(2014\)](#) : a synthesis of all previously existing principles and initiatives on data citation.

Principle 2 ← Author(s), Year, Dataset Title, Global Persistent Identifier → Principles 4, 5, 6 - Unique Identification, Access, Persistence

dataverse.org/best-practices/data-citation

 Dataverse Project About Community Best Practices Software Contact

Implementation." *IASSIST Quarterly* 2013;37. [Link](#).

Data Citation Standard

The citation standard defined here offers proper recognition to authors as well as permanent identification through the use of global, persistent identifiers in place of URLs, which can change frequently. Use of universal numerical fingerprints (UNFs) guarantees to the scholarly community that future researchers will be able to verify that data retrieved is identical to that used in a publication decades earlier, even if it has changed storage media, operating systems, hardware, and statistical program format.

Following are two authentic examples of replication data citations:

From International Studies Quarterly, King and Zeng, 2006, p. 209:

Gary King; Langche Zeng, 2006, "Replication data for: When Can History be Our Guide? The Pitfalls of Counterfactual Inference", Harvard Dataverse,
V2, <http://hdl.handle.net/1902.1/DXRXCFAWPUNKF.3:DaYT6QSX9r0D50ye+tXpA==>

From Political Analysis, Hanmer, Banks, and White, 2013:

Hanmer, Michael J.; Banks, Antoine J., White, Ismail K., 2013, "Replication data for: Experiments to Reduce the Over-reporting of Voting: A Pipeline to the Truth", Harvard Dataverse, V1, <http://dx.doi.org/10.7910/DVN/22893> ⓘ UNF:5:eJOVAjDU0E0jzSQ2bRCg9g==

This citation has seven components. Five are human readable: the author(s), title, year, data repository (or distributor), and version number. Two components are machine-readable:

1. Of the machine-readable components to these citations, the unique global identifier begins with either "hdl" (this refers to the international [HANDLE.NET](#) system) or "doi" (this refers to a [Digital Object Identifier \(DOI\)](#) system). This identifier is designed to persist even if URLs-or the web itself-are replaced with something else. When the citation appears online, the identifier is hot-linked to the URL that references the identifier, which works in browsers available today. In print, the URL is also included in the citation.
2. The universal numerical fingerprint begins with "UNF". Four features make the UNF especially useful: The UNF algorithm's cryptographic technology ensures that the alphanumeric identifier will change when any portion of the data set changes. Not only does this assure future researchers that they can use the same data set referenced in a years-old journal article, it enables the data set's owner to track each iteration of the owner's research. When an original data set is updated or incorporated into a new, related data set, the algorithm generates a unique UNF each time. The UNF is determined by the content of the data, not the format in which it is stored. For example, you create a data set in SPSS, Stata or R, and five years later, you need to look at your data set again, but the data was converted to the next big thing (NBT). You can use NBT, recompute the UNF, and verify for certain that the data set you're downloading is the same one.

- <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/2CEYWV>
- <https://dataverse.org/best-practices/data-citation>
- <https://www.force11.org/datacitationprinciples>

Learn more:

- Micah Altman and Gary King. (2007). “A Proposed Standard for the Scholarly Citation of Quantitative Data,” D-Lib Magazine, Vol. 13, No. 3/4 (March).

<http://datascience.iq.harvard.edu/publications/proposed-standard-scholarly-citation-quantitative-data>

- Paul E. Uhlir, R., Board on Research Data, Information, Policy, Global Affairs, & National Research Council. (2012). For attribution – developing data attribution and citation practices and standards: Summary of an international workshop. The National Academies Press.

http://www.nap.edu/openbook.php?record_id=13564



JOURNAL AUTHORS/EDITORS/REVIEWERS
JOURNAL AUTHOR GATEWAY
How to Get Published Resources >>
SAGE Author Services >>
Supplemental Material – Guidelines for authors
Advance: a SAGE preprints community
On Acceptance and Publication
Help Readers Find Your Article
Promote Your Article
Research Data Sharing Policies
ORCID
Author Resource: SAGE Consent Form
Journal Author FAQs
Useful Links/Resources
Manuscript Submission Guidelines
Open Access Options
Top Reasons to Publish with SAGE
JOURNAL EDITOR GATEWAY
JOURNAL REVIEWER GATEWAY
ETHICS & RESPONSIBILITY
PUBLISHING POLICIES

Research Data Sharing Policies

At SAGE we are committed to facilitating openness, transparency and reproducibility of research. We support and encourage research data to be shared, discoverable, citable and recognised as an intellectual product of value.

The following guidelines outline SAGE's general policy for handling the publication of research data alongside your published research, but please always refer to individual journal author guidelines to check for any bespoke policies.

- [What is 'research data'?](#)
- [Why is open research data important?](#)
- [What are SAGE's research data sharing policies?](#)
- [How do I share my research data in a repository](#)
- [SAGE partnership with the Figshare Data Repository](#)
 - [How do I use this service as a SAGE journal author?](#)
 - [What file formats can be accepted?](#)
- [How do I license my research data?](#)
- [What is a Data Accessibility Statement?](#)
- [What are data citation principles?](#)
- [Additional data sharing policies](#)
 - [Open Practice Badges](#)
 - [Data Replication](#)
 - [Pre-Registration of Studies](#)
 - [Code Ocean Trial](#)

What is 'research data'?

'Research data' refers to units of information collected through the course of a research project, which the researcher then analyses to produce knowledge. Research data varies widely in format across disciplines, and can be anything from spreadsheets of quantifiable information, to sound recordings of an interview.

Why is open research data important?

Sage Publishing:

- What are data citation principles?
- Where datasets are hosted in public repositories that provide datasets with Digital Object Identifiers (DOIs), we encourage these datasets to be formally cited in reference lists. Citations of datasets, when they appear in the reference list, should include the minimum information recommended by [DataCite](#) and follow journal style.
- DataCite recommended format for data citation is as follows:
- Creator (PublicationYear). Title. Publisher. Identifier (if available)
- <https://datacite.org/cite-your-data.html>

en [M](#) [S](#) [S](#) [S](#) [S](#) [S](#) [S](#) [S](#) [S](#) [F](#) [20](#) [A](#) [S](#) [M](#) [M](#) [S](#) [S](#) [G](#) [w](#) [+](#) [S](#) [O](#) [M](#) [✓](#) [20](#) [N](#) [Q](#) [Q](#) [S](#) [W](#) [G](#) [S](#) [@](#) [F](#) [20](#) [W](#) [M](#) [S](#) [x](#) [+](#) [-](#) [o](#) [d](#) [e](#) [c](#) [h](#) [a](#) [b](#) [g](#) [i](#) [j](#) [k](#) [l](#) [m](#) [n](#) [o](#) [p](#) [q](#) [r](#) [s](#) [t](#) [u](#) [v](#) [w](#) [x](#)

 DataCite

About us [▼](#) Services [▼](#) Resources [▼](#) Community [▼](#) Become a member [✉](#) [RSS](#) [Twitter](#)

DATACITE - CITE YOUR DATA

WHY IS IT SO IMPORTANT TO CITE DATA?

Books and journal articles have long benefited from an infrastructure that makes them easy to cite, a key element in the process of research and academic discourse. We believe that you should cite data in just the same way that you can cite other sources of information, such as articles and books.

DataCite DOIs help further research and assures reliable, predictable, and unambiguous access to research data in order to:

- support proper attribution and credit
- support collaboration and reuse of data
- enable reproducibility of findings
- foster faster and more efficient research progress, and
- provide the means to share data with future researchers

DataCite also looks to community practices that provide data citation guidance. The Joint Declaration of Data Citation Principles is a set of guiding principles for data within scholarly literature, another dataset, or any other research object (Data Citation Synthesis Group 2014). The FAIR Guiding Principles provide a guideline for the those that want to enhance reuse of their data (Wilkinson 2016).

Data Citation Examples

We recognise that the challenges associated with data publication vary across disciplines, and we encourage research communities to develop citation systems that work well for them. Our recommended format for data citation is as follows:

Creator (PublicationYear). Title. Publisher. Identifier

It may also be desirable to include information about two optional properties, Version and ResourceType (as appropriate). If so, the recommended form is as follows:

Creator (PublicationYear). Title. Version. ResourceType. Publisher. Identifier

References

¹ Data Citation Synthesis Group (2014). Joint Declaration of Data Citation Principles. Martone M. (ed) San Diego CA: [https://doi.org/10.5281/zenodo.35660](#).



14:38
20.10.2020 

■ References

Data Citation Synthesis Group (2014). Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11 <https://www.force11.org/group/joint-declaration-data-citation-principles-final>

- Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Bourne, P. E. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

Joint Declaration of Data Citation Principles

1. Importance

Data should be considered legitimate, citable products of research. Data citations should be accorded the same importance in the scholarly record as citations of other research objects, such as publications[[1](#)].

2. Credit and Attribution

Data citations should facilitate giving scholarly credit and normative and legal attribution to all contributors to the data, recognizing that a single style or mechanism of attribution may not be applicable to all data[[2](#)].

3. Evidence

In scholarly literature, whenever and wherever a claim relies upon data, the corresponding data should be cited[[3](#)].

■ 4. Unique Identification

■ A data citation should include a persistent method for identification that is machine actionable, globally unique.

