

# Responsible Data Management and Ethical Sharing in Context

Available resources and campus support

Renata Curty  
Research Facilitator - Social Sciences  
UCSB Library - Research Data Services

# Research Data Services

- Help to comply with funder mandates (i.e., data management planning and implementation)
- Consultations (e.g., data documentation, data de-identification, licensing)
- Extended Engagements (e.g., reproducible labs)
- Archiving/publishing and getting credit for data
- Curation (FAIR compliant datasets)
- Training (e.g., computational tools, Dryad, DMPTool, Carpentry workshops, RCR)
- Research Computing & Data website



# Research Computing & Data

rcd.ucsb.edu

University of California, Santa Barbara

UC **SANTA BARBARA** | Research Computing and Data

Home Resources Campus Units Communities Events

We offer a virtual gateway connecting affiliates to a wealth of computational and data-related resources to support academic research.

Our mission is to facilitate and advance groundbreaking scientific endeavors by providing seamless access to a myriad of tools, support services, and expertise available to our campus community.

Whether you are engaged in data-intensive projects or exploring novel computational methods, our website strives to be a one-stop destination, catering to researchers of all levels of expertise. From access to high-performance computing clusters to specialized data storage solutions and from training to personalized consultation, we have carefully curated a comprehensive suite of resources to bolster your research pursuits.

UCSB Library - Research Data Services  
5th Floor, Mountain Side  
rds@library.ucsb.edu

# Research Computing & Data

## rcd.ucsb.edu

- Identify supported data tools and available data storage options based on your research needs
- Locate relevant campus units to support your research
- Consult data-related tips, data management resources and applicable policies
- Find campus and local communities you may want to get involved in
- Stay up-to-date with relevant events and training

The underlying philosophy of **open data sharing** is that data and samples collected by scholars are part of our **global human heritage** and should be shared with other scholars and with the public.



TOP Factor - Transparency and Openness  
[topfactor.org](http://topfactor.org)



Data Sharing Requirements by Federal Agency  
[datasharing.sparcopen.org/data](http://datasharing.sparcopen.org/data)

# The carrots outweigh the sticks

Conditional Funding

Conditional Publication



Scientific Advancement

Visibility/Citations

Reputation

Promotion

New Partnerships

= Accountability!

# Reproducibility

Same data and code, different researchers, rendering same results

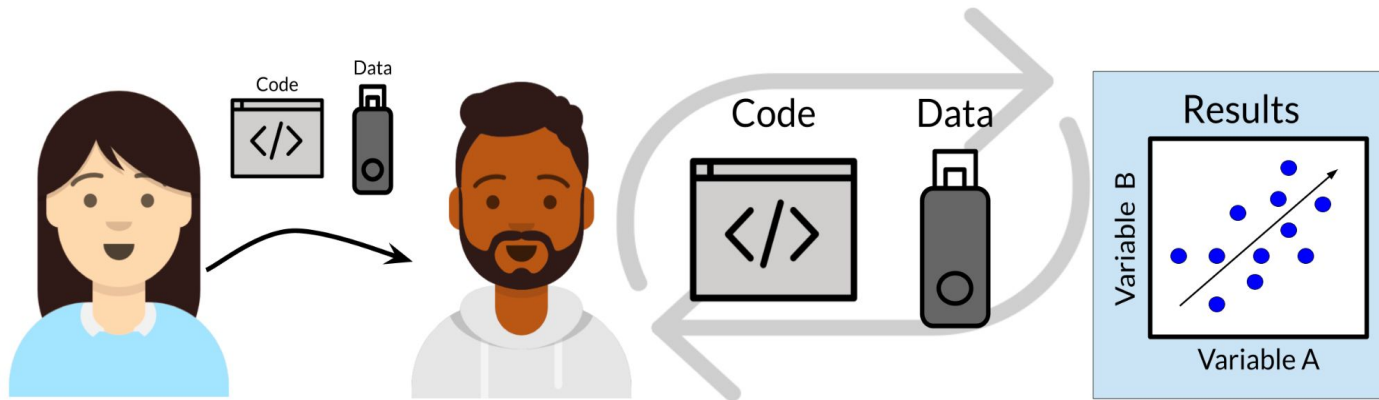


Image by: [Candace Savonen \(2022\)](#).

Quantitative data

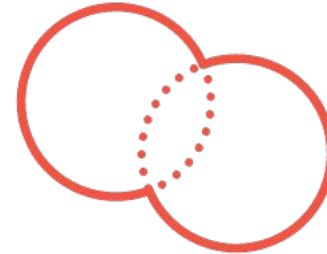
# Beyond Reproducibility

## Multiple “Reuses”



### Re-analysis

Follows the same analytical process employed by the original study for verification purposes (refute or confirm previous findings).



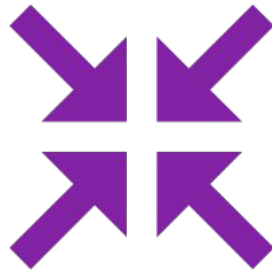
### Integration

Combines data from disparate data sources linked by one or more related variables.



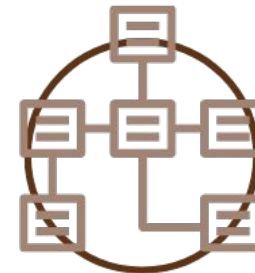
### Repurposing

Explores existing datasets in different ways to answer new research questions.



### Aggregation

Gathers data from multiple existing sources to form a unique dataset.



### Meta-analysis

Merges analyses (not raw data) from multiple independent sources.

Most data





# Not any data

## FINDABLE

Data and associated metadata should be easily discovered and located by both humans and computers



## ACCESSIBLE

Data and supporting metadata are preserved, easily available, and can be downloaded or used locally

## INTEROPERABLE

Data and metadata should work seamlessly across systems for data analysis, processing and storage



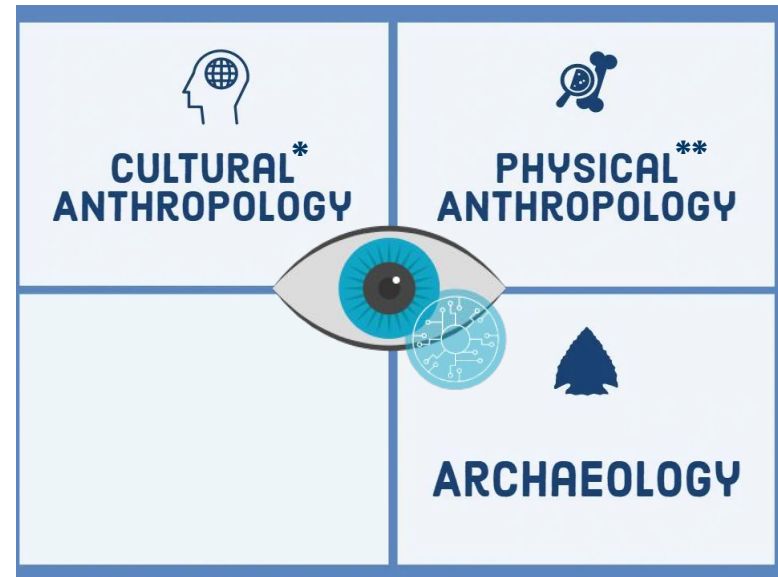
## REUSABLE

Data and metadata are well-described to allow interpretation and further reuse by replication, reanalysis or repurposing

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>

# Anthropology

“Anthropologists occupy a special position in social science: field research turns them into their own research instruments. They cannot delegate the social relationships of research to interchangeable interviewers, questionnaires or experiments.”



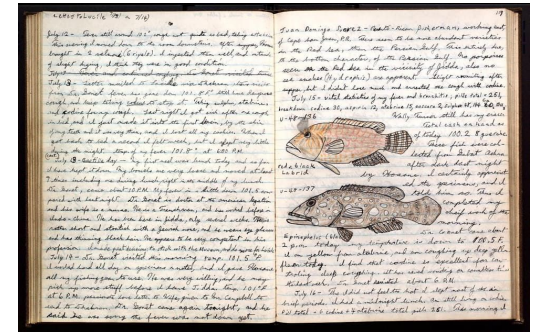
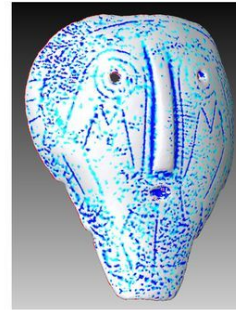
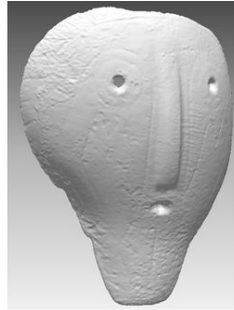
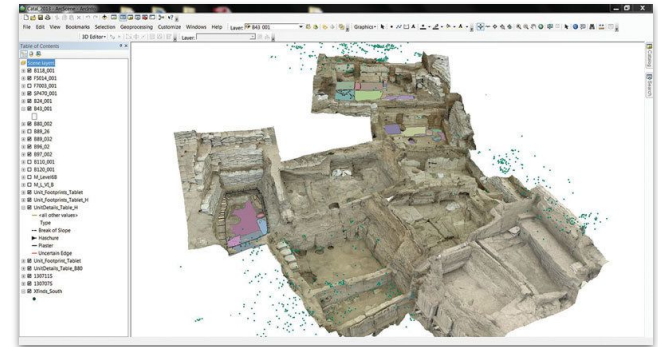
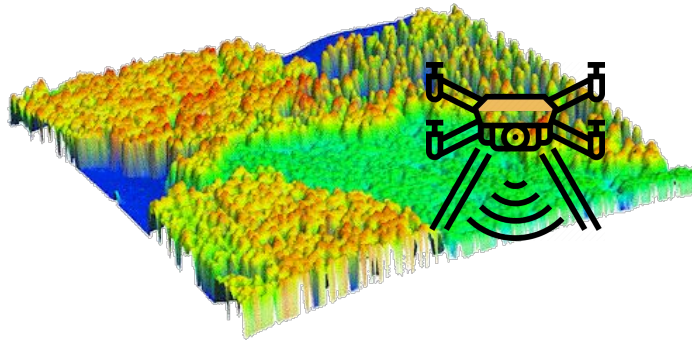
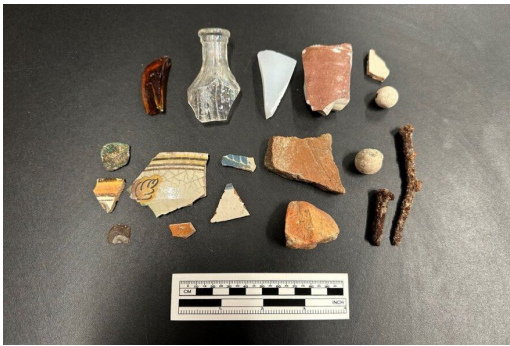
Pels, P. et al. (2018). Data management in anthropology: The next phase in ethics governance?. *Social Anthropology/Anthropologie sociale*, 26(3), 391-413. Retrieved Feb 25, 2024, from <https://doi.org/10.1111/1469-8676.12526>

# Data in Anthropology



Illustration created by the Author with Canva graphics.

- Great variety of data types
- Digital and physical
- Primary and secondary sources
- Naturalistic and some non-naturalistic data
- Increase use of mixed-method approaches
- Triangulation matters
- Inherently complex and often non-repeatable/re-obtainable



UCSB Library - Research Data Services  
5th Floor, Mountain Side  
rds@library.ucsb.edu

UC SANTA BARBARA

# Challenges for Data Management & Sharing

- Linking multiple types of artifacts (physical and digital) and their documentation
- Community often involved in the decision-making process
- Confidentiality and protection of people and research sites
- Indigenous data sovereignty
- Repatriation of samples

# Multiple Ethical Responsibilities

- With: people or communities from whom records were collected or generated from, to the discipline/field, potential reusers and society at large
- Assess the sensitiveness of the data and make informed decisions about how to handle it after consulting:
  - the individuals or groups affected (directly or indirectly) by your study
  - data archivists, data specialists and alike
  - other relevant subject experts
- Familiarity with licensing, intellectual property, legislations
- Ethical reuse of existing sources

# Hence Another Set of Principles...



## COLLECTIVE BENEFIT

Data ecosystems should be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.



## AUTHORITY CONTROL

Indigenous Peoples' rights and interests should be recognized and their authority to control such data be empowered.



## RESPONSIBILITY

Researchers are accountable for how data supports Indigenous Peoples' self-determination and collective benefit.



## ETHICS

Indigenous Peoples' rights and wellbeing should be the primary concern at all stages of the data life cycle and across the data ecosystem.



Carroll, S.R., Herczog, E., Hudson, M. *et al.* Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Sci Data* 8, 108 (2021). <https://doi.org/10.1038/s41597-021-00892-0>

Potential values of sharing data need to be balanced against the risks of harms associated with the identification of participants, communities and research sites.





# Open vs. Restricted

**HARVARD**  
Dataverse

ADD DATA - SEARCH - ABOUT - USER GUIDE - SUPPORT - SIGN UP - LOG IN

Biocultural Anthropology Lab at Washington State University | Lab Website  
(Washington State University)

Harvard Dataverse > Biocultural Anthropology Lab at Washington State University >

## Replication Data for: Breastfeeding patterns are associated with human milk microbiome composition: the Mother-Infant Microbiomes, Behavior, and Ecology Study (MIMBES)

Version 1.0

Holdsworth, Elizabeth A.; Williams, Janet E.; Gartstein, Maria; McGuire, Mark A.; McGuire, Michelle K.; Meehan, Courtney L., 2023, "Replication Data for: Breastfeeding patterns are associated with human milk microbiome composition: the Mother-Infant Microbiomes, Behavior, and Ecology Study (MIMBES)", <https://doi.org/10.7910/DVN/PSE1NQ>, Harvard Dataverse, V1, UNF:9:P-h7TBV4-vvVCCkRkVLYg== [BIU:INF]

Access Dataset -  
Contact Owner | Share

Dataset Metrics  
14 Downloads

**Description**  
Replication data for publication "Breastfeeding patterns are associated with human milk microbiome composition: the Mother-Infant Microbiomes, Behavior, and Ecology Study (MIMBES)" in PLOS One. This dataset includes all variables used in the publication analyses. (2023-06-19)

**Subject**  
Medicine, Health and Life Sciences; Social Sciences

**License/Data Use Agreement**  
CC BY-NC 4.0

Files | Metadata | Terms | Versions

Search this dataset...

Filter by: File Type: All | Access: All | Sort

1 to 7 of 7 Files

Profile Unknown - 26 B Published Jul 11, 2023 2 Downloads MDS: 57f_05e	Download
activate.R 19 Bytes - 28.0 KB Published Jul 11, 2023 2 Downloads MDS: 406_307	Download
Codebook.pdf Adobe PDF - 21.8 KB Published Jul 11, 2023 2 Downloads MDS: 443_394	Download
genusdata.Rdata Rzip Archive - 7.9 KB Published Jul 11, 2023 2 Downloads MDS: 82f_0d8	Download
metadata.tab Tabular Data - 6.5 KB Published Jul 11, 2023 2 Downloads 18 Variables, 40 Observations UNF:9:P-h7TBV4-vvVCCkRkVLYg== Sample data/metadata	Download
renv.lock Unknown - 66.4 KB Published Jul 11, 2023 2 Downloads MDS: 54c_0ef	Download
Replication Code Rmd R Notebook - 26.4 KB Published Jul 11, 2023 2 Downloads MDS: 57f_05e	Download

Replication code to be run in R. Use renv package to load appropriate versions of packages.

<https://doi.org/10.7910/DVN/PSE1NQ>,

**tDAR** The Digital Archaeological Record  
A SERVICE OF DIGITAL ANTIQUITY

LOG IN OR SIGN UP

Search...

ABOUT SEARCH USING TDAR UPLOAD NEWS SHA SAA GIVE MY ACCOUNT

## Tribally Approved American Indian Ethnographic Analysis of the Proposed Escalante Valley Solar Energy Zone

Part of the BLM Utah Project Metadata project

**Author(s):** Richard W. Stoffle; Kathleen A. Van Vlack; Hannah Z. Johnson; Phillip T. Dukes; Stephanie C. De Sola; Kristen Simmons

**Year:** 2011

### Summary

Ethnography and ethnographic synthesis for the Solar Programmatic Environmental Impact Statement and Solar Energy Study Areas in Portions of Arizona, California, Nevada, and Utah. The proposed Escalante Valley solar energy zone (SEZ) is located in Iron County, Utah.

The Escalante Valley SEZ American Indian study area was traditionally occupied, used, aboriginally owned, and historically related to the Numic-speaking peoples of the Great Basin and western Colorado Plateau. The SEZ American Indian study area extends beyond the boundaries of the SEZ because of the existence of cultural resources in the surrounding landscape.

Participating tribes include the Confederated Tribes of the Goshute Reservation and the Paiute Indian Tribe of Utah.

### Cite this Record


Tribally Approved American Indian Ethnographic Analysis of the Proposed Escalante Valley Solar Energy Zone.  
Richard V. Stoffle  
Simmons, K.

Name	Size	Creation Date	Date Uploaded	Access
2011_StoffleEtAl_EthnoEscalanteValleySolarEnergyZone.pdf	30.50mb		May 17, 2023 10:35:26 AM	Confidential

**This Resource**  
Bureau of Land Management

**Accessing Restricted Files**  
At least one of the files for this resource is restricted from public view. For more information regarding access to these files, please reference the contact information below

**Contact(s):** Bureau of Land Management, Utah State Office



**Downloads** 1

Restricted Access Some or all of this resource's attached file(s) are not publicly accessible.

2011\_StoffleEtAl\_EthnoEscalanteV (30.50mb)

Request Access, Submit Correction, Comment (requires login)

Tweet

Like

Email

<https://doi.org/10.48512/XCV8475290>

# Sensitive Data

<b>Tag Type</b>	<b>Description</b>	<b>Security Features</b>	<b>Access Credentials</b>
Blue	Public	Clear storage, Clear transmit	Open
Green	Controlled public	Clear storage, Clear transmit	Email- or OAuth Verified Registration
Yellow	Accountable	Clear storage, Encrypted transmit	Password, Registered, Approval, Click-through DUA
Orange	More accountable	Encrypted storage, Encrypted transmit	Password, Registered, Approval, Signed DUA
Red	Fully accountable	Encrypted storage, Encrypted transmit	Two-factor authentication, Approval, Signed DUA
Crimson	Maximally restricted	Multi-encrypted storage, Encrypted transmit	Two-factor authentication, Approval, Signed DUA

Dataverse DataTags - <https://privacytools.seas.harvard.edu/datatags>

# Anthropological Ethics



AMERICAN  
ANTHROPOLOGICAL  
ASSOCIATION

- Do no harm
- Obtain informed consent and necessary permissions
- Be open and honest regarding your work
- Weigh competing ethical obligations due collaborators and affected parties
- Make your results accessible
- Protect and preserve your records
- Maintain respectful and ethical professional relationships

<https://americananthro.org/about/anthropological-ethics>

RCR Training!

# Obtain informed consent & necessary permissions

IRB | Laws (National and International) | Community & Disciplinary Norms

- Navigate options and make more informed decisions in the planning stage
- Craft better language for informed consents and account for shareability when applicable
- Navigate licenses for reusing pre-existing data
- Define strategies to implement promises made to research participants and in data management plans (e.g., de-identification, data expunge)

# Common Mistakes in Informed Consents

- Stating that the data will only be accessed by the research team
- Stating that the data will only be used for a particular project
- Not mentioning where the data will live after the project is completed
- Putting an expiration date on the data
- Not differentiating between/defining administrative study records versus research data
- Not adequately explaining how personal identifiers will be removed/masked in plain language
- Misusing the word anonymous

# Example



“The data we collect will only be used for [*Project Name*] and only accessible to the research team.”



“Any personal information that could identify you will be removed or changed before study files are shared with other researchers or results are made public.”

Suggested guide: [Representing Data Sharing in Informed Consent](#) (2022)

# Be Open Regarding your Work

- Define strategies and updated your plan as your progress
- Produce robust documentation to allow further interpretation and inspectability (e.g., codebooks, data dictionaries, READMEs) and follow metadata standards
- Choose open tools for data analysis (e.g., OpenRefine, Taguette, QGIS, SQL, R, Python)
- Stick to open formats

# Metadata & Controlled Vocabulary

Metadata element	Controlled Vocabulary
<location>	Thesaurus of Geographic Names (TGN) (Getty)
<culture>	Outline of World Cultures (OWC)
<language>	ISO 639 standard
<subject>	<ul style="list-style-type: none"><li>● General -Library of Congress Subject Headings (LCSH)</li><li>● AFS Ethnographic Thesaurus</li><li>● Outline of Cultural Materials (OCM)</li></ul>

Metadata Standards: Dublin Core, Darwin Core, Genome Metadata, EAD, ISO Geospatial  
<https://rdamsc.bath.ac.uk>



# Make your Results Accessible

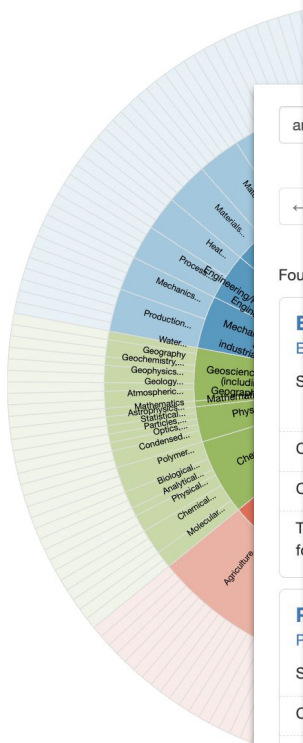
- Get familiar with data de-identification techniques and packages/software to help automate the process
- Choose a data repository to house shareable data and other research deliverables (e.g., Dryad, tDAR, ICPSR, QDR, MorphoSource, NCBI, GBIF)
- Determine restrictions if needed
  - Embargo
  - Access Control
  - DUA
  - Metadata only
- Include Data Availability Statement (DSA) in your associated publications (cross-referencing)

# Data Repository Selection

- Are there any repositories dedicated to your discipline/area of study?
- Have been recommended by your institution, publishers and/or funding agencies?
- Assign persistent and unique identifiers to datasets?
- Offer clear terms and conditions that meet legal requirements (e.g. data protection, embargo, and restrictions)?
- Offer curation services and review data and documentation prior to data publication?
- Any costs involved?

And more...





Repository details

## MorphoSource

General Institutions Terms Standards

Name of repository **MorphoSource**

Repository URL <https://www.morphosource.org/>

Subject(s)
 

- Anatomy
- Life Sciences
- Basic Biological and Medical Research
- Biology
- Systematics and Morphology
- Zoology
- Humanities and Social Sciences
- Classical Archaeology
- Art History
- Plant Sciences
- Ancient Cultures
- Humanities
- Fine Arts, Music, Theatre and Media Studies

Description
 

MorphoSource is a data repository specialized for 3D representing physical objects used in research in education (e.g., from museum or laboratory collections). It allows researchers and museum collection staff to store and organize, share, and distribute their own 3d data. Furthermore any registered user can immediately search for and download 3d morphological data sets that have been made accessible through the consent of data authors.

Contact morphosource@duke.edu

Content type(s)
 

- Images
- Audiovisual data
- Structured graphics
- other
- Raw data

Keyword(s)
 

- 3D data
- Computed Tomography (CT)
- Magnetic Resonance Imaging (MRI)
- Synchrotron imaging
- amphibia
- anthropology
- archaeology
- arts
- botany
- cultural heritage
- geology
- laser scan
- mammalia
- paleontology
- photogrammetry
- structured light

Persistent identifier(s) of the repository FAIRsharing\_doi:10.25504/FAIRsharing.7mFpMg

Repository size 186.369 Media Files; 65,004 Physical Objects Represented; 1,980 Data Projects; 20,810 Taxonomies (GBIF Linked); 27,149 Taxonomies (User Created); 1,356 Collections and Facilities; 19,353 Users; 2,073 Contributors

Repository type(s) disciplinary

Mission statement for designated community <https://www.morphosource.org/docs/about>

Research data repository language(s) English

Data and/or service provider data provider  
service provider

Search

Toggle short help

Sort by ▾

Studies and Religious Studies

disciplines. The project

Sciences

of 19 years.

# Data “Available Upon Request”

- Lack of findability
- Limited accessibility
- Low response rate to requests (<40%)
- Have become less accepted by academic publishers without reasonable justification

Research Data Services Data Literacy Series, n.02/2024

### Crafting Effective Data Availability Statements

Many scientific journals require authors to provide a data availability statement (DAS) to adhere to openness and transparency principles. The DAS specified how and where the underlying data used in a study can be accessed and the terms under which they can be obtained. We cover best practices for writing such statements in compliance with policy requirements.

**When writing a DAS:**

- Check for specific journal guidelines for acceptable data sharing methods
- Indicate if the publication generated or analyzed new data
- Make generated data available through a data repository whenever possible
- Ensure compliance with funder, publisher, and institutional policies (e.g., IRB)
- Specify where resources are located, preferably with DOIs
- Attribute and provide access to public source datasets and derivatives
- Explain any access restrictions (ethical, legal, or commercial)
- Honor data requests and clearly specify any criteria for acquiring restricted data
- Offer accessible contact points for requests and inquiries about data sources

**Examples of DAS for specific scenarios:**

<b>Theoretical Research (no data)</b>	Data sharing does not apply to this article as no datasets were generated or analyzed for this study.	<b>Shared as Supplementary File</b>	The data supporting this study's findings are available as supplementary materials (click link to file).
<b>Openly Available</b>	The data supporting this study's findings are openly available in repository named at [DOI/DOIc], reference number.	<b>Subject to Embargo Period*</b>	The data supporting the findings will be available in repository named at [DOI/DOIc link] following an 18-month embargo period from the publication date to protect the researcher's intellectual property rights.
<b>Derived from Public Sources</b>	The data that support this study's findings are available in repository named at [DOI/DOIc], reference number. Data is derived from the following repository available in the public domain: list all resources and citations.	<b>Restricted Data</b>	Due to privacy or ethical restrictions, the data supporting this study's findings are available upon reasonable request (corresponding author and IRB/Institutional Body contact info).
<b>Data User Agreement Required</b>	The data that support this study's findings are available in repository named at [DOI/DOIc], reference number. Access and reuse is conditional to a Data User Agreement (DUA).	<b>Third Party Restricted Data**</b>	The study's supporting data can be accessed and analyzed through third party vendor(s) subject to restrictions. Data requests can be made via [URL/contact info].
<b>Non-digital Data</b>	Non-digital data supporting this study are stored at [location/archive/business]. Data requests can be made via [contact info].	<b>Not Shared (Ethics/No Consent)</b>	The study participants did not give written consent for their data to be shared publicly, so supporting data is unavailable due to the research's sensitive nature. IRB/Institutional Body contact info.

**DON'T FORGET!**

Include citations for datasets you have used or produced in the reference list. Need help with that? Check out the [Data Citation & Usage](#) handout or contact us at [info@library.ucsb.edu](mailto:info@library.ucsb.edu)

UC SANTA BARBARA Library [www.library.ucsb.edu](http://www.library.ucsb.edu)

Hussey, I. (2023, May 8). Data is not available upon request. <https://doi.org/10.31234/osf.io/jbu9r>

Mulligan, C. J. et. al. (2022). Data sharing in biological anthropology. *American Journal of Biological Anthropology*, 178, 26-53. <https://doi.org/10.1002/ajpa.24499>

Tedersoo, L., et al. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data*, 8(1), 192. <https://doi.org/10.1038/s41597-021-00981-0>

UCSB Library - Research Data Services  
5th Floor, Mountain Side  
[rds@library.ucsb.edu](mailto:rds@library.ucsb.edu)

# Data “Available Upon Request”

- Lack of findability
- Limited accessibility
- Low response rate to requests (<40%)
- Have become less accepted by academic publishers without reasonable justification



## Contact us:

[rds@library.ucsb.edu](mailto:rds@library.ucsb.edu)

## Join our Carpentry Workshops:

[ucsbcarpentry.github.io](https://ucsbcarpentry.github.io)

## Subscribe to our Data Literacy Series!

[rcd.ucsb.edu/data-literacy-series](https://rcd.ucsb.edu/data-literacy-series)