Data Sharing in Psychology

Insights from My Journey in Data Sharing

Pierina Cheung, PhD
Lecturer (Research Scientist)
Centre for Research in Child Development
Office of Education Research, NIE

An Institute of
NIE NATIONAL INSTITUTE OF EDUCATION SINGAPORE
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE
70 1950-2020 Years of Inspiring Learning Transforming Teaching Advancing Research

Why did I share data?

Requested by a journal (& funding agencies)

Discussion on the Reproducibility Crisis in Psychology (importance of increasing transparency and accountability)

I had wished other people shared their data in the past

Benefits

Benefits your future self!

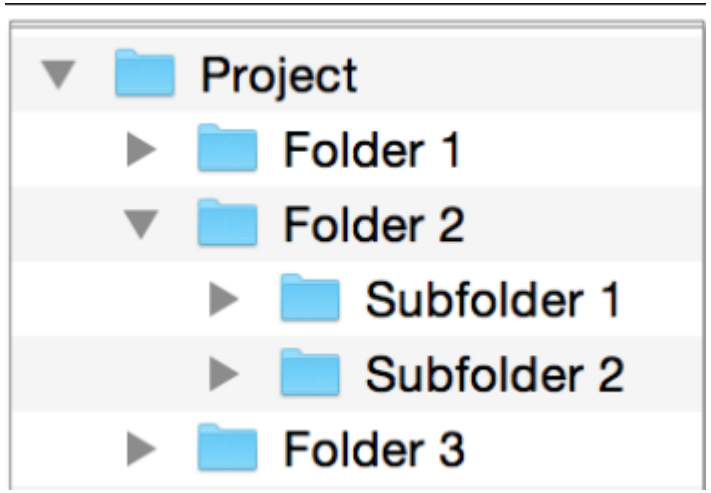


Image from UC Merced Library

3 years later.

5 years later.

10 years later.

Broken hard drive
Broken flash drive
Broken laptop / desktop

Forgot where data were saved!
Which one was the latest data file?

Benefits

Benefits your future self!

Increases impact and visibility of your research (& citations)

Provides credit to the researcher

Saves people (esp. grad students) from duplicating effort

Allows for secondary analyses, discovering patterns not previously examined

Serves as teaching resources

PLOS ONE

PUBLISH ABOUT BROWSE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

A study of the impact of data sharing on article citations using journal policies as a natural experiment

Garret Christensen, Allan Dafoe, Edward Miguel, Don A. Moore, Andrew K. Rose

Published: December 18, 2019 • https://doi.org/10.1371/journal.pone.0225883

PLOS ONE

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Sharing Detailed Research Data Is Associated with Increased Citation Rate

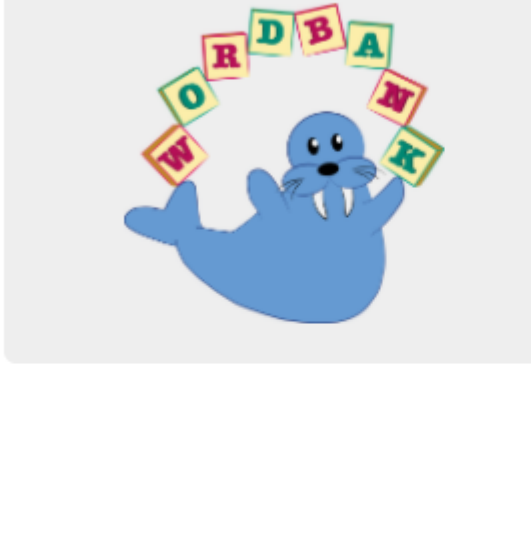
Heather A. Piwowar, Roger S. Day, Douglas B. Fridsma

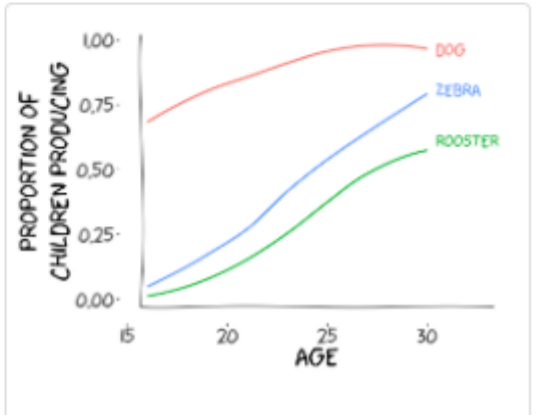
Published: March 21, 2007 • https://doi.org/10.1371/journal.pone.0000308

Wordbank | Contributors | Analyses | Population | Publications | Blog | About | FAQ

Wordbank

An open database of children's vocabulary development





Wordbank contains data from 75,144 children and 82,983 CDI administrations, across 29 languages and 56 instruments:

Joe Simmons
@jpsimmon

In my MBA class yesterday, a student asked, "Aren't scientists required to make their data available?" In response I just started screaming.

8:12 PM · Oct 9, 2020 · Twitter Web App

8 Retweets **1** Quote Tweet **94** Likes

No Consent for Data Sharing

Parallel individuation in preschoolers

Contributors: **Pierina Cheung**, Mathieu Le Corre

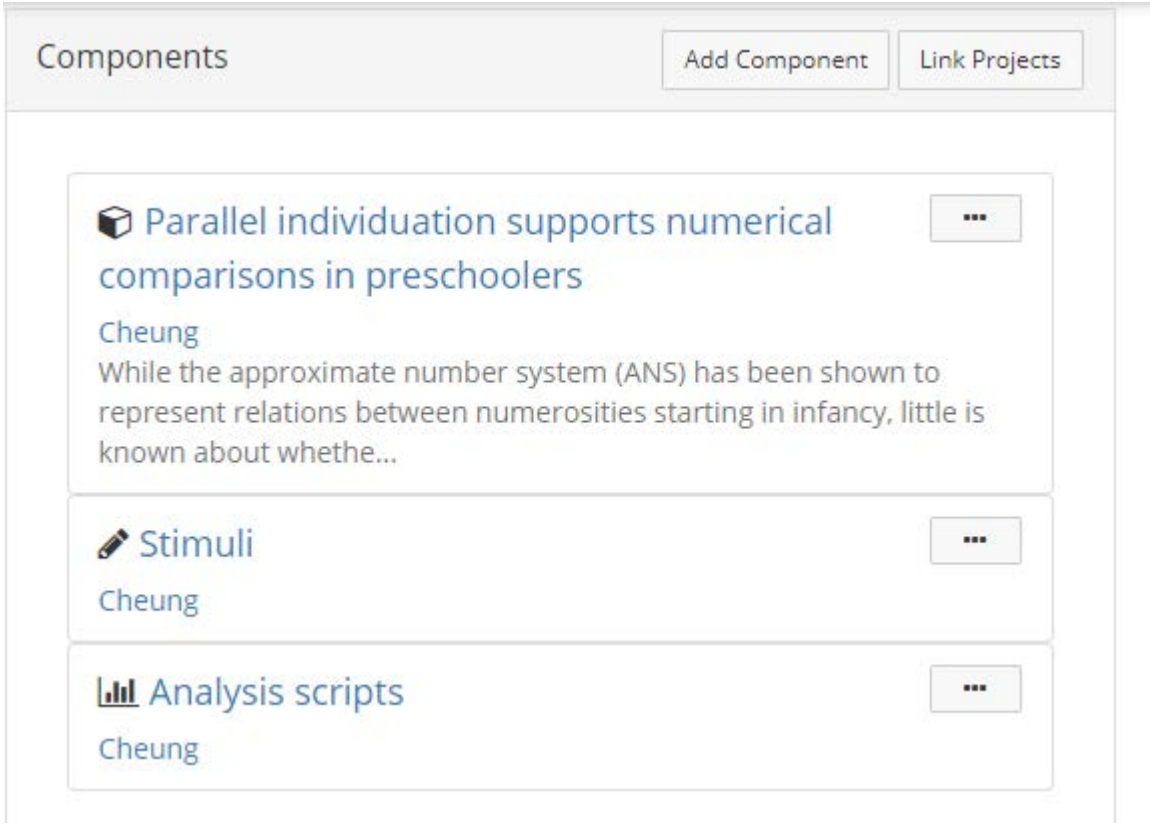
Date created: 2016-12-09 01:59 AM | Last Updated: 2017-07-25 02:27 AM

Create DOI

Category: Project

Description: Add a brief description to your project

License: Add a license



More Recent Study

Multi-digit processing and place value understanding

Contributors: **Pierina Cheung**, Daniel Ansari

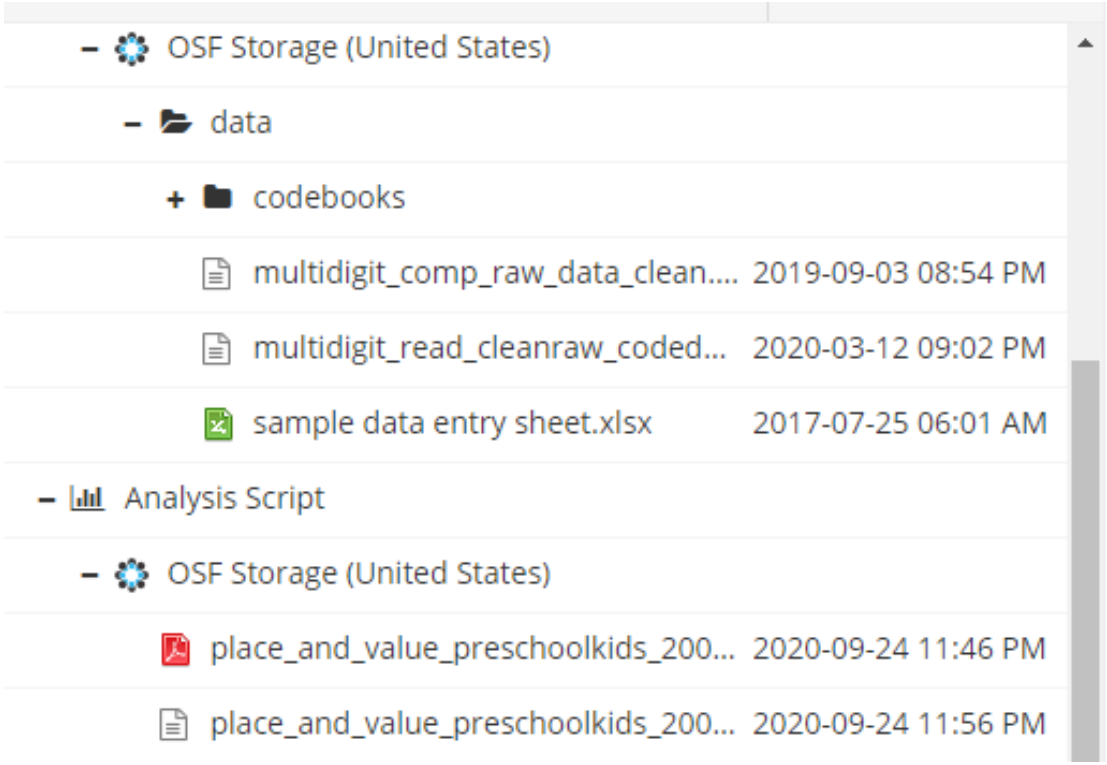
Date created: 2017-06-07 10:10 PM | Last Updated: 2020-03-12 10:26 PM

Create DOI

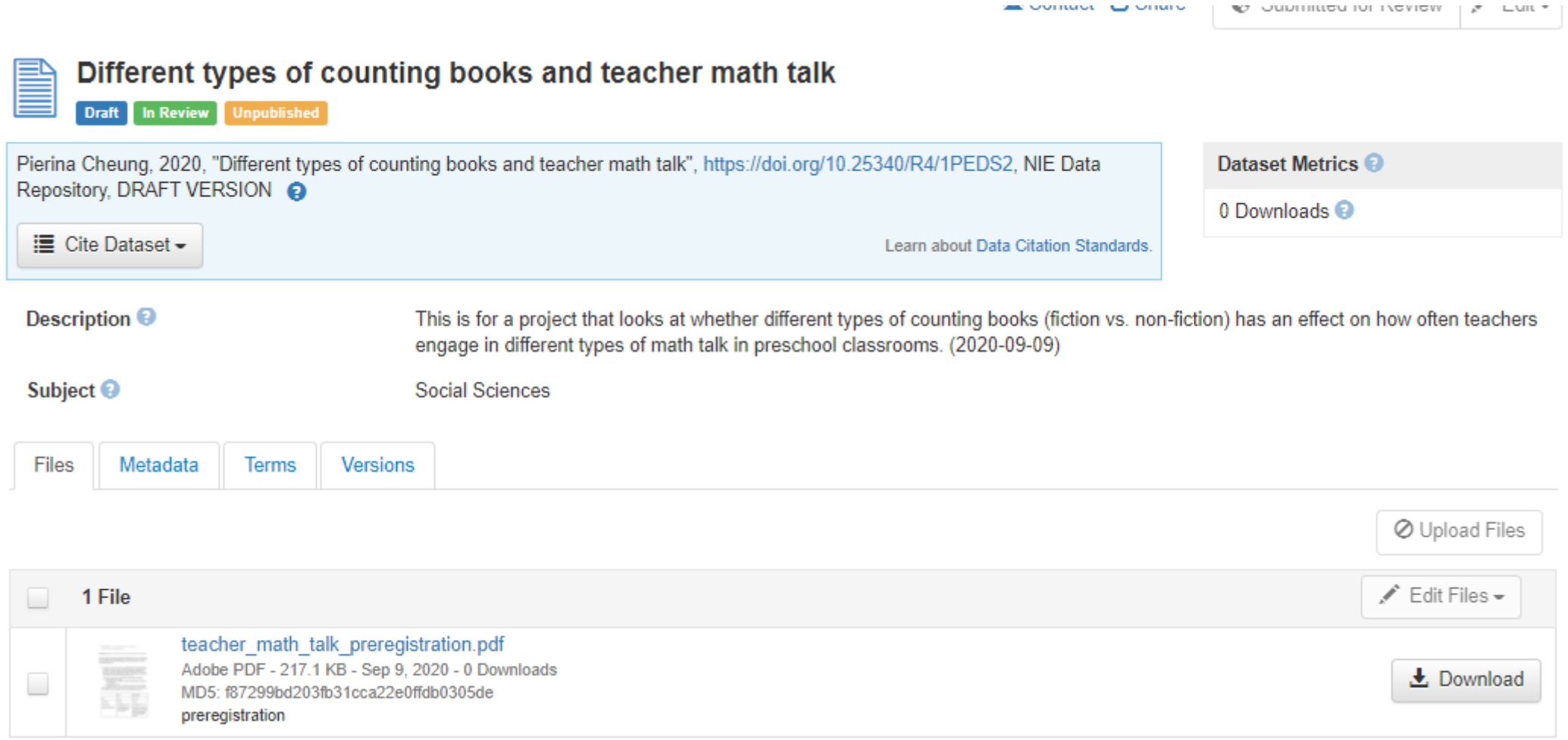
Category: Project

Description: Add a brief description to your project

License: Add a license



Latest Study (on NIE Data Repository)



Challenges

Takes time to clean up data and create codebook

(use software to help create codebooks)

Removing identifiers

(find someone in your lab to cross-check for you)

Wording on consent forms & ethics application

ICPSR

Data Management & Curation

Log In/Create Account

QUALITY PRESERVATION ACCESS CONFIDENTIALITY CITATION

Additional Resources

- ICPSR's Approach to Confidentiality
- American Statistical Association, Data Access and Personal Privacy: Appropriate Methods of Disclosure Control
- Confidentiality and Data Access Committee (CDAC) forum for staff members of Federal statistical agencies

Recommended Informed Consent Language for Data Sharing

Language to Avoid

Promises in the informed consent can appear to limit an investigator's ability to share data with the research community. In reality, investigators can inform study participants that they are scientists with an obligation to protect confidentiality and still share the study data with the broad scientific community. Many effective means exist to create public-use data files or share restricted-use data files under controlled conditions. That is, data can be modified to reduce the risk of disclosure or shared with additional safeguards while preserving their value for science.

Model Language

Here are two model statements investigators may use in informed consents to describe protection of confidentiality that also allows data sharing.

https://www.icpsr.umich.edu/web/pages/datamanagement/confidentiality/conf-language.html

Challenges

Takes time to clean up data and create codebook *(use software to help create codebooks)*

Removing identifiers *(find someone in your lab to cross-check for you)*

Wording on consent forms & ethics application

Part of a Larger project *(share some data but not all)*

Not possible to remove identifiers? *(share methods, detailed analyses)*

Final Messages

Do what you can.

Share what you can.

Keep yourself up-to-date on policies, trends, preferences of your field, etc.

(no need to try everything in one attempt)
(expect that it's a learning experience)

Thank you

Questions?

Email: pierina.cheung@nie.edu.sg