

# Systematic Challenges to Data Sharing (And some tentative ways forward)

SESYNC Workshop, February 27-March 1
Nic Weber, University of Washington
Sebastian Karcher, Qualitative Data Repository
https://doi.org/10.6084/m9.figshare.4700035

### **QDR** Overview

- At Syracuse University (United States)
- Online since 2014, NSF funded
- · Curated, qualitative & mixed methods data
- Small holdings (~20 datasets), growing steadily



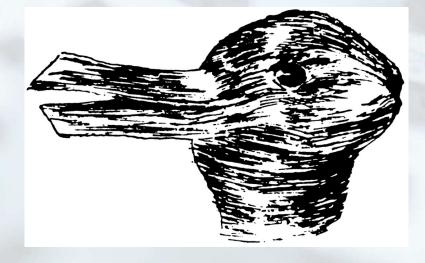
# **Epistemic / Ontological**

No objective social world

Understanding can't be shared intersubjectively

Separation/extraction of data from researcher and their

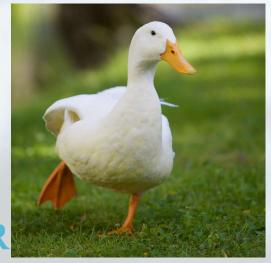
writing impossible





### Epistemic / Ontological (ways forward)

- Find language appropriate for research tradition
- Consultative approach to data sharing
- Data sharing technology that links text to data





#### Costs

- Data sharing is costly
  - increased cost in collection
  - increased cost for documentation
  - increased cost for data preparation (e.g. anonymization)



# Costs (ways forward)

Funding for data preparation

Lowering costs with good research practices

Lowering costs with technology





### **Lack of Benefits**

- Few tangible rewards for data sharing
- Data citation still very uncommon
- Risk of being scooped
- Little evidence of qualitative data re-use



### Lack of Benefits (ways forward)

- Rewards: tenure, hiring, prizes
  - Requires advocacy from faculty, support from funders, and flexibility from institutions of higher ed (ha!)
- Risk of being scooped
  - Documenting provenance of ideas and conceptual development
- Foster data citation/acknowledgment
  - Infrastructure is built, it is now largely up to communities of practice to implement and use
- Showcase & measure use
  - Compelling, and accurate use cases are badly needed



# Ethical / Legal

- Risk to participants
- Effective de-identification not possible
- No informed consent for legacy/old projects
- Broad informed consent may deter participants
- Copyright/proprietary data



# Ethical / Legal (ways forward)

- Not all data can or should be shared -but this doesn't mean it cannot be reliably archived and preserved.
  - Embargos are underexplored and often oversimplified in contemporary archiving
  - Distributed Credentialing (ICPSR)
  - Data Tags (IQSS / DataVerse)
  - Encoding privacy statements in qualitative data models (QDR + UW)



# **Image** credits

#### **Duck Rabitt**

https://commons.wikimedia.org/wiki/File:Duck-Rabbit.png PD from Fliegend Blaetter

#### Stretching by Ernst Vikne

https://commons.wikimedia.org/wiki/File:White\_domesticated\_duck,\_stretching.jpg CC-BY-SA https://creativecommons.org/licenses/by-sa/2.0/

#### Rabbit

https://commons.wikimedia.org/wiki/File:Wild\_rabbit\_us.jpg PD by Harvey Henkelmann

#### Squirrel

https://commons.wikimedia.org/wiki/File:Squirrel\_(5508296475).jpg CC-BY-SA: https://creativecommons.org/licenses/by-sa/2.0/

Nesting Season -- Pat Gaines

https://www.flickr.com/photos/patgaines/4560361619 CC-BY-NC https://creativecommons.org/licenses/by-nc/2.0/

