



Australian Research Data Commons

# Data Quality IG

PRESENTED BY

Mingfang Wu

Australian Research Data Commons

ANZLIC/ICSM Metadata Working Group Meeting

23 March, 2020

# Data quality

Data is considered to have sufficient quality if it is “fit for intended uses”

# FAIR data principles

## Findable

- F1.** (Meta)data are assigned a globally unique and persistent identifier
- F2.** Data are described with rich metadata (defined by R1 below)
- F3.** Metadata clearly and explicitly include the identifier of the data they describe
- F4.** (Meta)data are registered or indexed in a searchable resource

## Interoperable

- I1.** (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2.** (Meta)data use vocabularies that follow FAIR principles
- I3.** (Meta)data include qualified references to other (meta)data

## Accessible

- A1.** (Meta)data are retrievable by their identifier using a standardised communications protocol
  - A1.1** The protocol is open, free, and universally implementable
  - A1.2** The protocol allows for an authentication and authorisation procedure, where necessary
- A2.** Metadata are accessible, even when the data are no longer available

## Reusable

- R1.** Meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1.** (Meta)data are released with a clear and accessible data usage license
  - R1.2.** (Meta)data are associated with detailed provenance
  - R1.3.** (Meta)data meet domain-relevant community standards

# Four key dimensions of data quality

- intrinsic data quality (accuracy, objectivity, believability, reputation)
- contextual quality (value-added, relevancy, timeliness, completeness, appropriate amount of data)
- representational quality (interpretability, ease of understanding, representational consistency, concise representation)
- accessibility quality (accessibility and access security)

Wang, R. Y. and Strong, D. M. (2013). *Beyond accuracy: What data quality means to data consumers*. Journal of Management Information Systems, Vol. 12, No. 4  
<http://www.jstor.org/stable/40398176>

# Australian Data Quality Interest Group

- Purpose: IG provides a forum for Australian data providers, repository operators and data consumers to discuss challenges and strategies for meeting data quality standards and procedures.
- Formed in October 2019, meets on the first Thursday (1-2pm AEST) every two months via Zoom
- The group is in close collaboration with the ESIP Information Quality Cluster
- Resource page: <https://ardc.edu.au/resources/communities-of-practice/>