Building Trust: facilitating Data Use and Reuse

Prof. Devika P. Madalli
Indian Statistical Institute, India
devika@drtc.isibang.ac.in
Data by itself may not convey much.

Data is used as basis for decision making.

Data exploitation, deployment and analysis largely done through automated processes.

Data Trust is crucial.
Example of the Indian Data Scenario:

data.gov.in

Open Government Data (OGD) Platform India or data.gov.in is a platform for supporting Open data initiative of Government of India. This portal is a single-point access to datasets, documents, services, tools and applications published by ministries, departments and organisations of the Government of India. The platform was customised by National Informatics Centre (NIC) in line with the National Data Sharing Accessibility Policy to develop the Data Portal India.
Other research data Websites in India CSIR, Bhuvan-ISRO, Datanet India, Human Proteome Map
The National Policy on Data Sharing and Accessibility (NPDSA) is a Government of India policy for sharing data. The Policy increases the accessibility and easier sharing of non-sensitive data amongst the registered users and their availability for scientific, economic and social developmental purposes.
Is FAIR = Trust
or
Trust subset of FAIR?
or
FAIR subset of Trust?
Discussion on data.gov.in and other data websites

Issues:

- Data Integration
- Standards for representation
- Workflows
- Sensitive Data and Privacy issues
Trust

Raw Data

QC

Privacy

Data Selection

Mediated

Automated

Checksums

Various frameworks

Curation

What needs protection?
What needs anonymization?
Open Questions

1. Metadata Indicators for Data set Quality?

2. Feedback mechanism for ensuring Privacy of data was honoured?