



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.

GOVERNMENT OF INDIA | A Digital India Initiative

Skip to navigation | Skip to main content | DataGov States/ULB

data.gov.in
Open Government Data (OGD) Platform India

Type search keyword

LOG IN | REGISTER

Call for Participation In Survey for the study on Socio Economic Value of Open Data

Analytics

- 205,527 RESOURCES
- 4,279 CATALOGS
- 111 DEPARTMENTS
- 17.07 M TIMES VIEWED
- 6.07 M TIMES DOWNLOADED
- 118 CHIEF DATA OFFICERS

CATALOG

INDICATOR DASHBOARD

Education

Demography

Economy

Industries

Transport

Labour And Employment

GOVERNMENT OF INDIA | A Digital India Initiative

Skip to navigation | Skip to main content | DataGov States/ULB

data.gov.in
Open Government Data (OGD) Platform India

Type search keyword

LOG IN | REGISTER

VISUALIZATION OF THE DAY

31 AUG 2018

TOP 10 DISTRICTS IN TERMS OF PERCENTAGE OF HOUSEHOLDS USING IMPROVED SANITATION FACILITY IN PUNJAB DURING 2015-16

Analytics

- 205,527 RESOURCES
- 4,279 CATALOGS
- 111 DEPARTMENTS
- 17.07 M TIMES VIEWED
- 6.07 M TIMES DOWNLOADED
- 118 CHIEF DATA OFFICERS
- 6,940 APIs
- 1,469 VISUALIZATIONS

CATALOG

Health Management Information System

INDICATOR DASHBOARD

Education

Demography

Economy

Industries

Transport

Labour And Employment

SECTORS

Power and Energy

POLICY and GUIDELINES

APIs

SUGGEST A DATASET

Your Name

Your Email Address*

Suggest the Dataset or App *

What do you want to do with Data ?

Source of Data

SUBMIT

Other research data Websites in India CSIR, Bhuvan- ISRO, Datanet India, Human Proteome Map

Datanet India
Empowering India-centric Socio-economic and Electoral data

Explore Insights into the socio-economic research with any of our 720 websites

Indiastat.com Districts of India Elections in India Datanet India Books Top 100 India

Datanet India is perhaps the unexampled company providing Secondary level socio-economic and political statistical information about India, its states, regions, districts and electioneering to the global research fraternity since last 17 years.

Datanet India was established as an IT-enabled private limited company in February 2000 to render its services in the socio-economic information domain.

As its core activity, Datanet collects, collates and compiles in ready to use socio-economic statistical data about India, its various sectors, regions, states and districts to make it available online.

Indiastat.com
Revealing India... Statistically

Explore the **right perspective** on India

India's most comprehensive e-resource of socio-economic data

CSIR-INFORMATICS PORTAL
Informatics related Resources Developed & Maintained by CSIR

HOME • BIOINFORMATICS • THERAPEUTICS • **RESOURCES** • SITES • CONTACT

Welcome to Home Page of Informatics Resources

Databases

CSIR (Council of Scientific and Industrial Research) is India's largest Research and Development (R&D) organization, with 37 laboratories and 39 field stations or extension centers spread across India. It has a collective staff of over 17,000. It is an autonomous body, mainly funded by the Ministry of Science and Technology, Govt. of India.

The research and development activities of CSIR includes various fields such as aerospace engineering, Structural engineering, ocean sciences, Life sciences, metallurgy, chemicals, mining, food, petroleum, leather, and environment.

CSIR laboratories have developed number of software, databases, mirror sites and web services in different fields of science and technology that includes chemoinformatics, bioinformatics, pharmacoinformatics etc. This page compile and maintain various informatics related services developed by CSIR laboratories.

News
therapeutics.

Highlights
team members.

CSIR CCMB CDRI CIMAP CLRI IGB IHBT IICB IICT IITR IMTECH NBRI NIIST NCL 4PI

Bhuvan
Indian Geo-Platform of ISRO

National Remote Sensing Centre
NSIC | Search | Bhuvan Site | Bhuvan Collaborators

Visualization & Free Download

Bhuvan-3D Bhuvan-3D Open Data Archive Climate & Environment EO-derived Products

Maps & OGC services

Thematic Services Disaster Management Support Services Ocean Services Create a Map / GIS Bhuvan/ISGIS

Geospatial Central Menus

Information School Bhuvan Tail Information Ground Water Pipeline Grid Waterbodies Urban Survey Portal

Application Sectors

Collaborative applications - Platform to share your data and create geospatial applications

Agriculture Forestry Governance Water Tourism Urban Rural

Healthcare

Healthcare HealthPCC Himalay Pradesh Karnataka

Special Applications

Data Discovery Decision Support Dashboard Hydrological Products Image Gallery

International Disasters RIS Pass Quick Links Locate Satellites MANU

Multilingual Translation On Demand EO Services Online Shapefile Creation OSCAT-3D

Smart Tracking

State Portals/Applications (Click on any state)

State ANDHRA PRADESH

Asset Mapping Andhra Pradesh

Updates

Download

Visualization of MC Weather Products New

Bhuvan Training on Aug 24-25, 2018 New

2018 AWP (plan) and 2015 LISA-4 (2d) Satellite data for download New

High Resolution (HRS) Image - Year 2017 New

Ground Truth of Coastal (GT) and Wetland

Contouring of Aerial Images 3 Crores in MGNREGA

Geotagging of Aerial Images

Baseline Studies on Water Use Efficiency of Irrigation projects

g-Governance Dashboard

HUMAN PROTEOME MAP

Home Query Download FAQs Contact us

About Human Proteome Map

The Human Proteome Map (HPM) portal is an interactive resource to the scientific community by integrating the massive peptide sequencing result from the draft map of the human proteome project. The project was based on LC-MS/MS by utilizing of high resolution and high accuracy Fourier transform mass spectrometry. All mass spectrometry data including precursors and HCD-derived fragments were acquired on the Orbitrap mass analyzers in the high-high mode. Currently, the HPM contains direct evidence of translation of a number of protein products derived from over 17,000 human genes covering >84% of the annotated protein-coding genes in humans based on >290,000 non-redundant peptide identifications of multiple organs/tissues and cell types from individuals with clinically defined healthy tissues. This includes 17 adult tissues, 6 primary hematopoietic cells and 7 fetal tissues. The HPM portal provides an interactive web resource by reorganizing the label-free quantitative proteomic data set in a simple graphical view. In addition, the portal provides selected reaction monitoring (SRM) information for all peptides identified.

Statistics

Organs/cell types	30
Genes identified	17,294
Proteins identified	30,057
Peptide sequences	293,700
N-terminal peptides	4,297
Splice junctional peptides	66,947
Samples	85
Adult tissues	17
Fetal tissues	7
Cell types	6

Adult tissues

Spinal cord Frontal cortex Spleen Esophagus Pancreas Colon Rectum

Lung Heart Liver Gall bladder Adrenal gland Kidney Ovary

Fetal tissues

Testis Liver Heart Brain Placenta

Ovary Gut

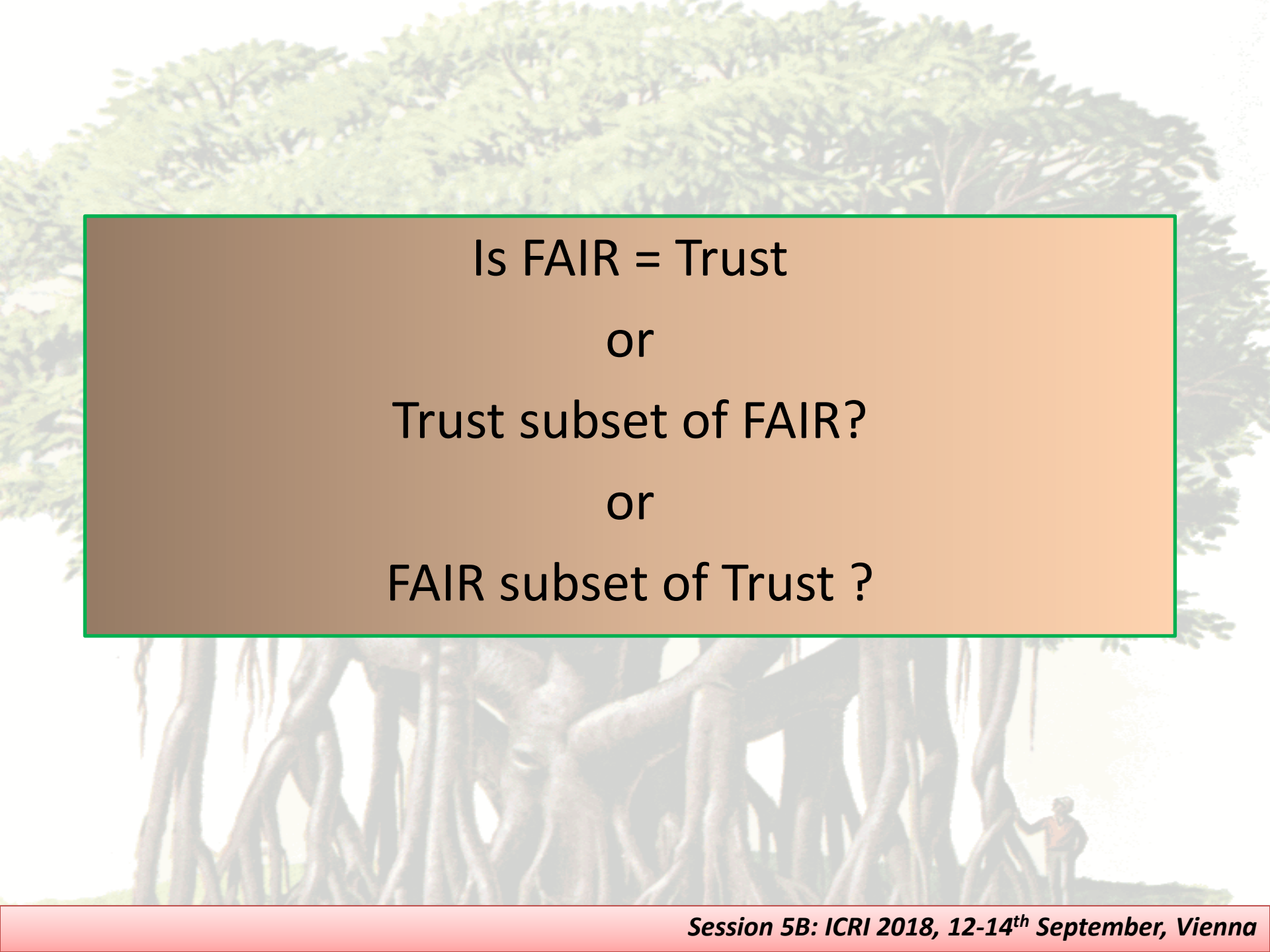
Hematopoietic cells

Common myeloid progenitor Common lymphoid progenitor

Platelets Monocytes CD4+ T cells CD8+ T cells NK cells B cells

NPDSA policy

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

The background of the slide features a large, ancient tree with thick, gnarled roots and dense green foliage. A small figure of a person stands near the base of the tree on the right side, providing a sense of scale. The text is centered within a semi-transparent brown rectangular box.

Is FAIR = Trust
or
Trust subset of FAIR?
or
FAIR subset of Trust ?

Discussion on data.gov.in and other data websites

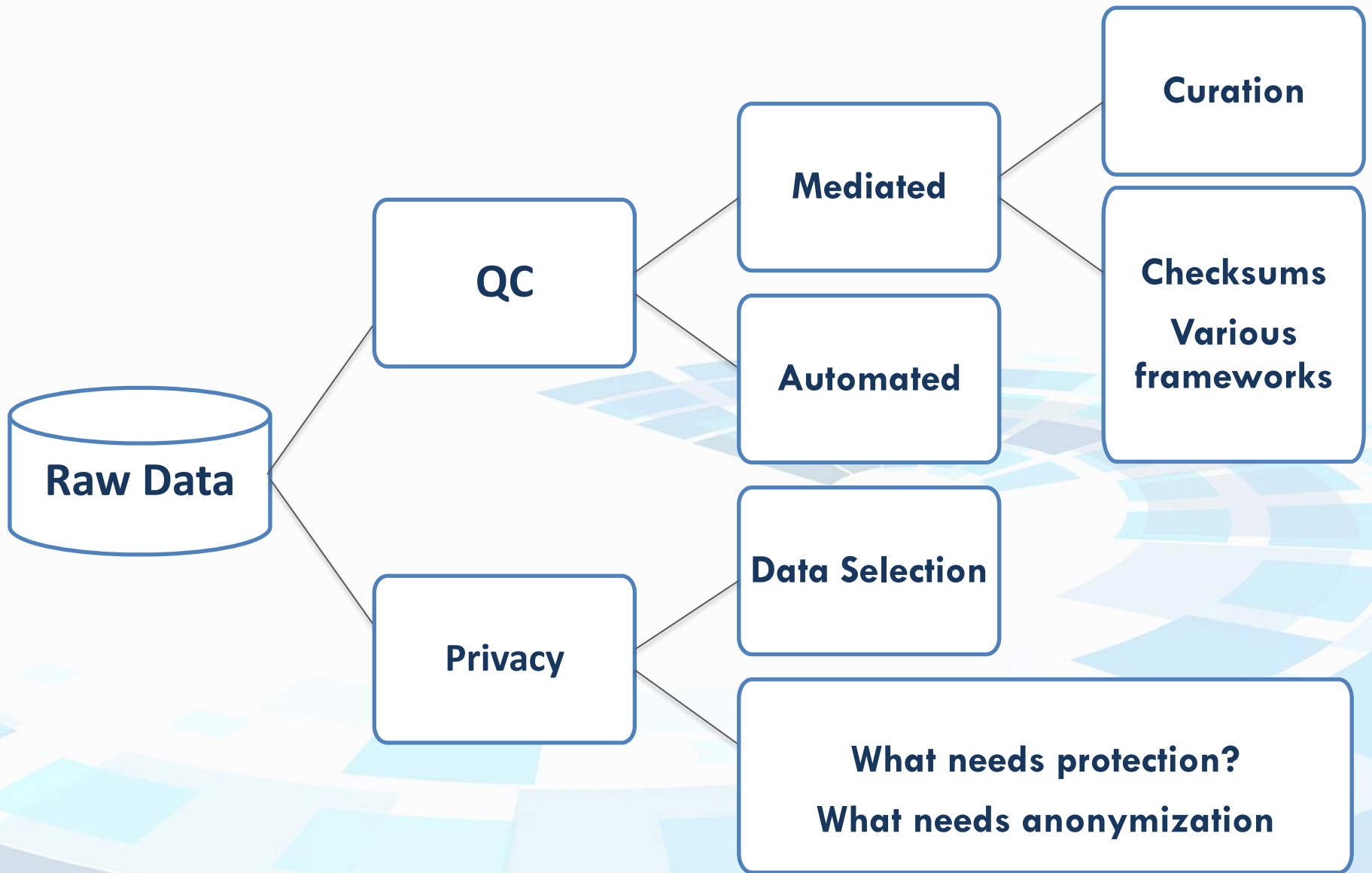
Issues:

Data Integration

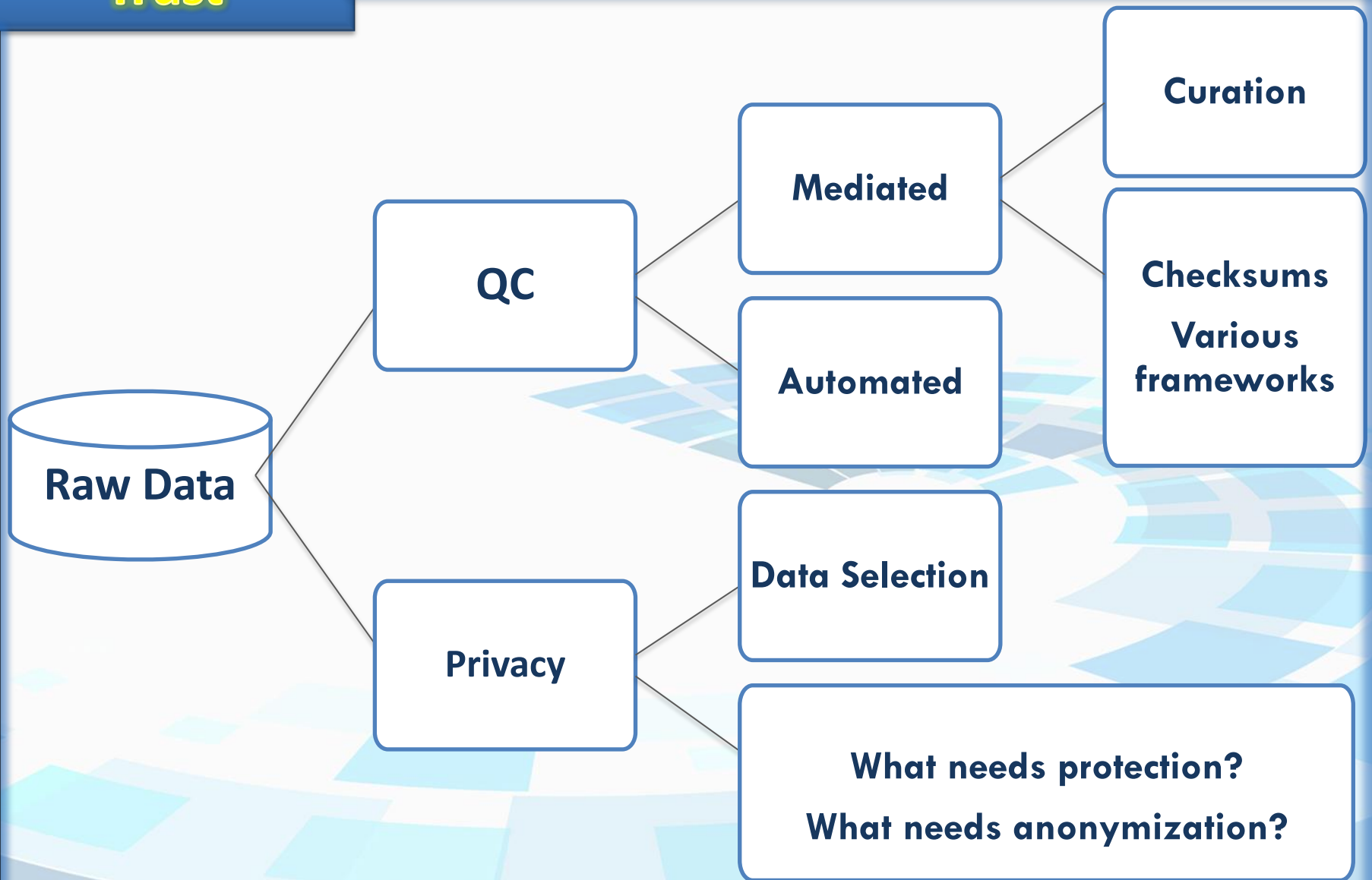
Standards for representation

Workflows

Sensitive Data and Privacy issues



Trust



Open Questions

Can we incorporate

1.

Metadata Indicators for
Data set Quality?

2.

Feedback mechanism for
ensuring Privacy of data
was honoured?

Thank you!