Machine-readable Data Release Policy

Facilitating Access, Accountability and Automation

Dr Lesley Arnold
Research Fellow, Curtin University
Cooperative Research Centre for Spatial Information, Australia

Data Release

## Simplifying the Process

Open Data

Machine-readable rules

Easy to implement

Data Release Issues

- Large amount of data still not accessible

Data still not accessible
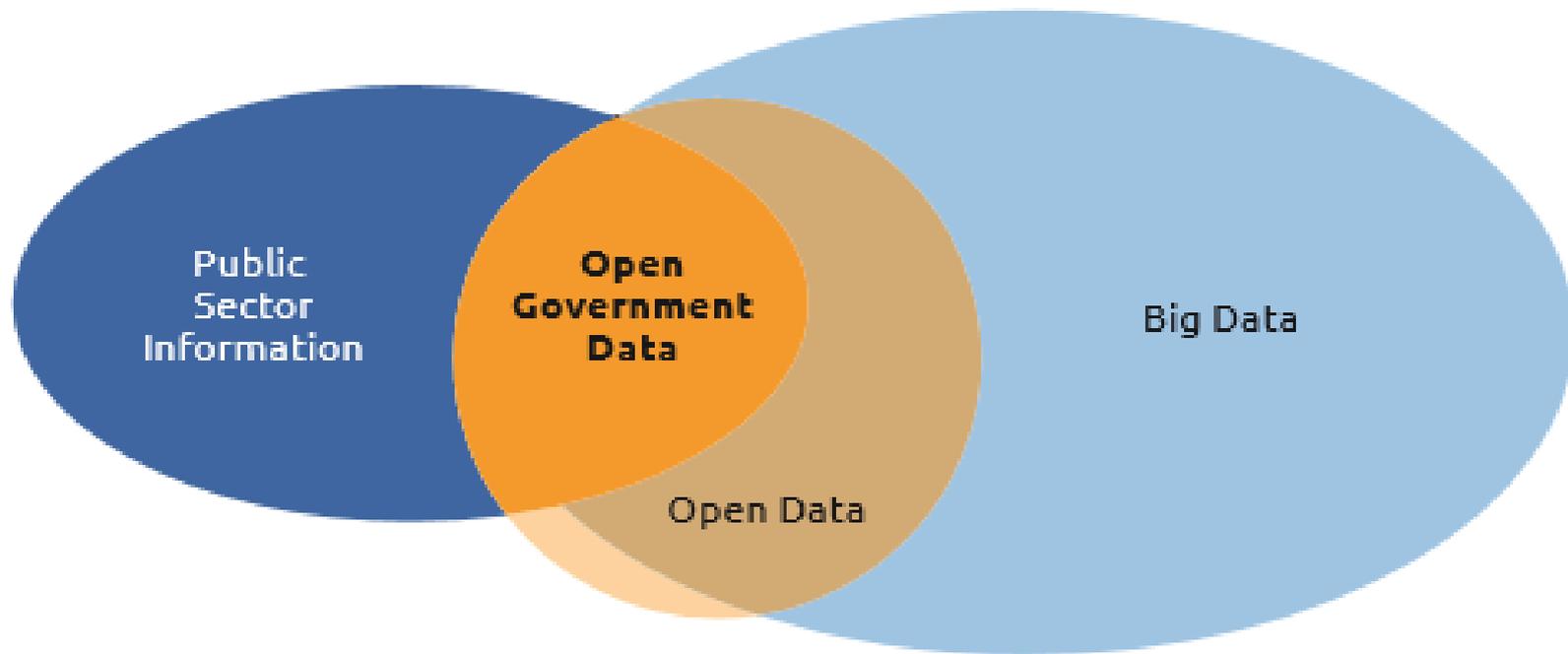
Data Release Issues

- Large amount of data still not accessible
- Liable for inadvertent release
- Fear backlash from public

Data still not accessible
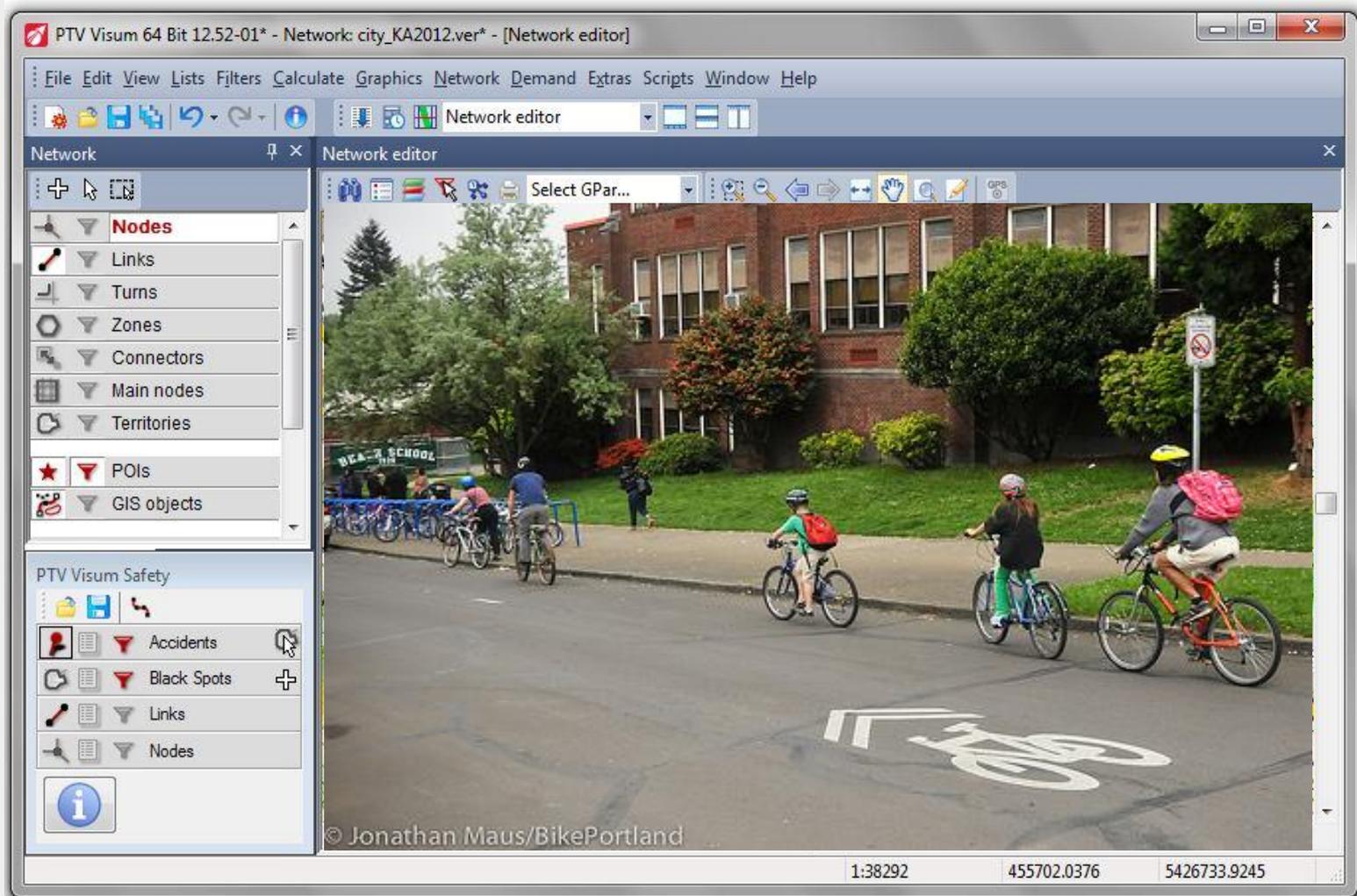
# Data Release Issues

- Large amount of data still not accessible
- Liable for inadvertent release
- Fear backlash from public
- Shortcomings in data
- Not formatted for public consumption
- Warrantability
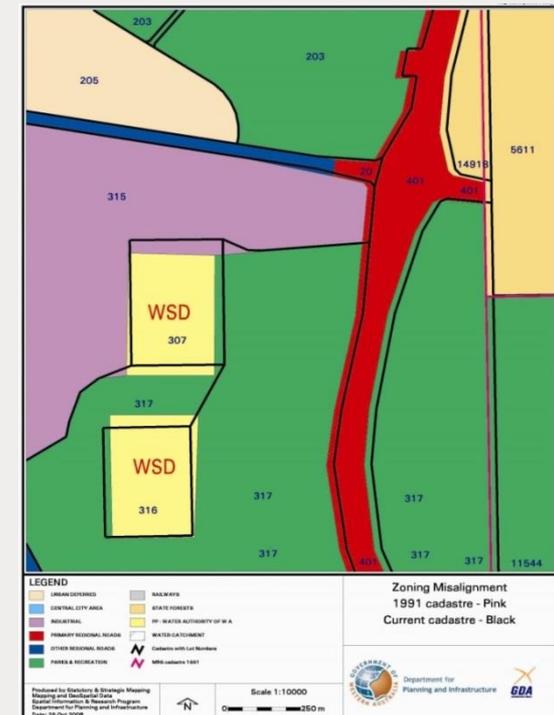
Data Release Issues

- Interpreted incorrectly
- Difficulty communicating fit for purpose
- Data misused
- Reengineered
- System hacked
- Data breach

Data Release Issues

- Complex release processes
  - Third-party Intellectual Property
  - Value-added applications
- Lack of consistent guidelines
  - Bilateral Agreements in place
 - No clear authority/accountability

The Data Release Project

Systematic data release process

UNGGIM **+** CRCSI **+** Geoscience Australia

Global Recognised Policy

Machine Readable Rules

Software Proof of Concept
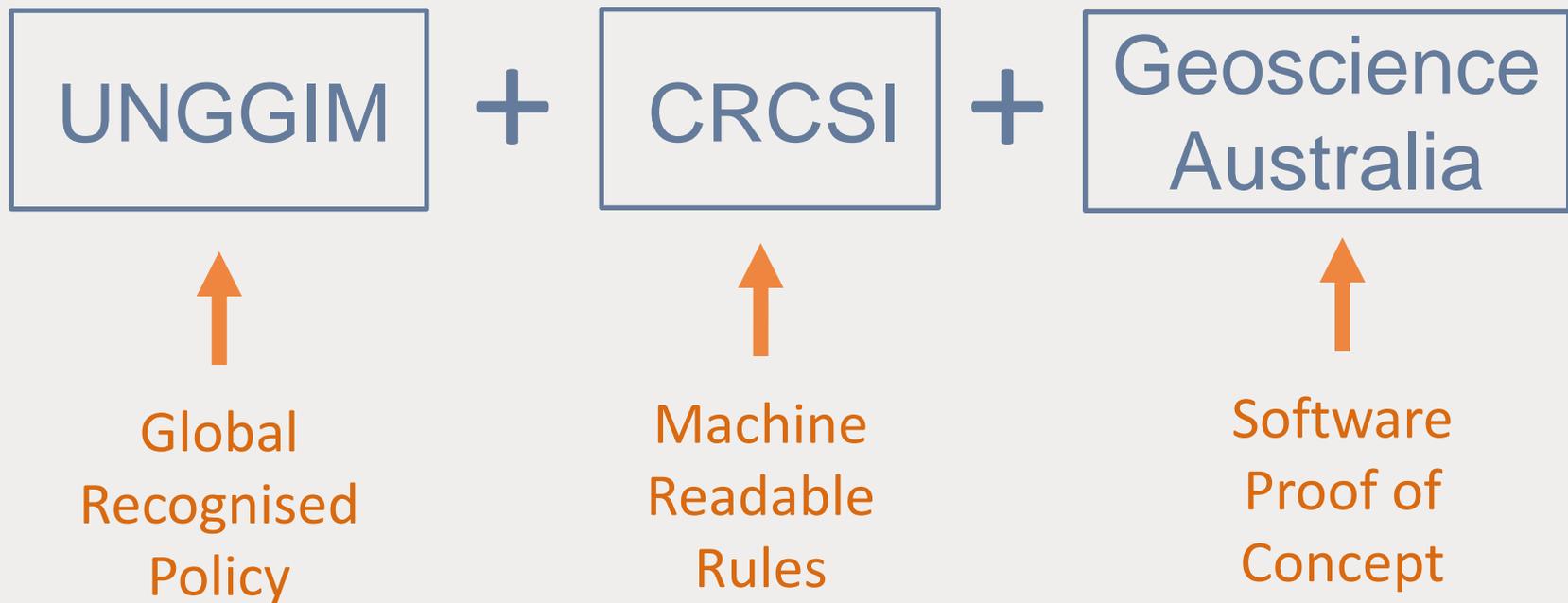
# Methodology

1. Policy documents

Data Sharing
Data Access
Open Data
Privacy Act
Freedom of Information

2. Expert knowledge

Team Procedures
Checklists
Flowchart
Data Release Forms
Data Licence

3. Ontology
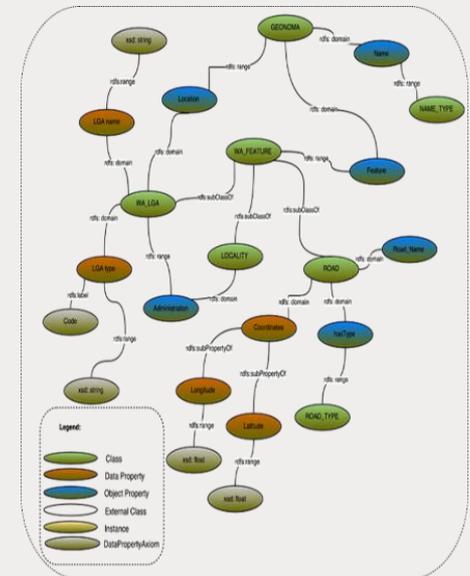
Identify Concepts
Create Relationships
Rule development

# **Methodology**

4. Develop Software Code that is shareable

Ontology

Map Portal

Access Services

Catalogue

Metering

View/download

Data Release

Security

Licensing

Ecommerce
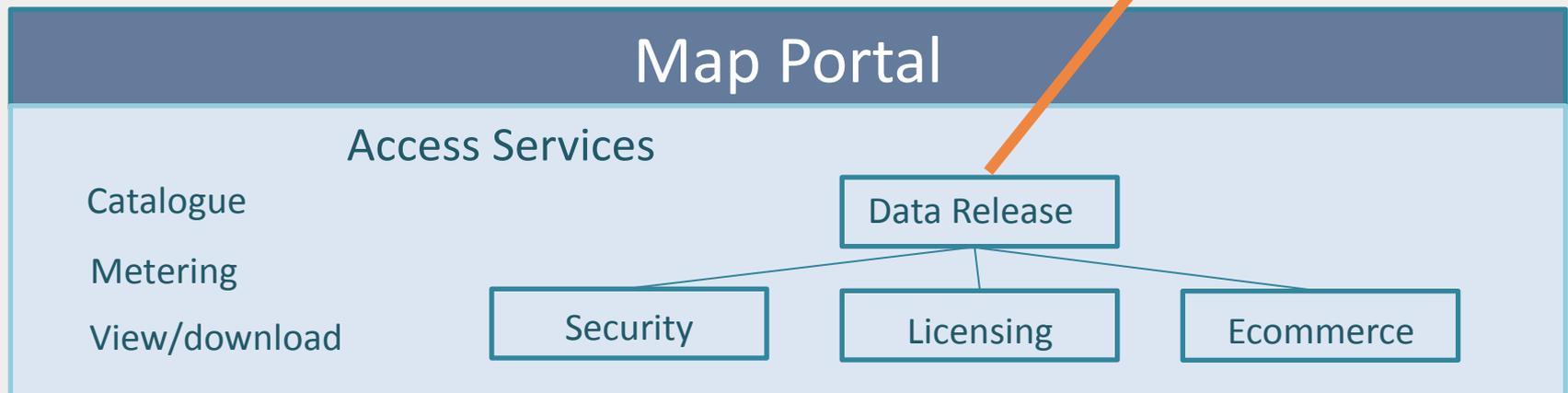
Software Code
- Captures commonality
- Individual differences

Data Release Ontology
Access Category

Dataset — *can be accessed by* — User

Dataset
- Access Category
  - Open Access
  - Restricted
  - Confidential
  - Not to be Released
- Value
  - Price
  - Free
  - Full-commercial
  - Cost Recovery
  - Freemium
- Licence
  - Public Domain
  - Commercial
  - Non-Commercial (copyrighted)

User
- Action
  - View only
  - Query
  - Download
- Type
  - Citizen
  - Academia
  - Government
  - Business

Rules:
Can only have one access category
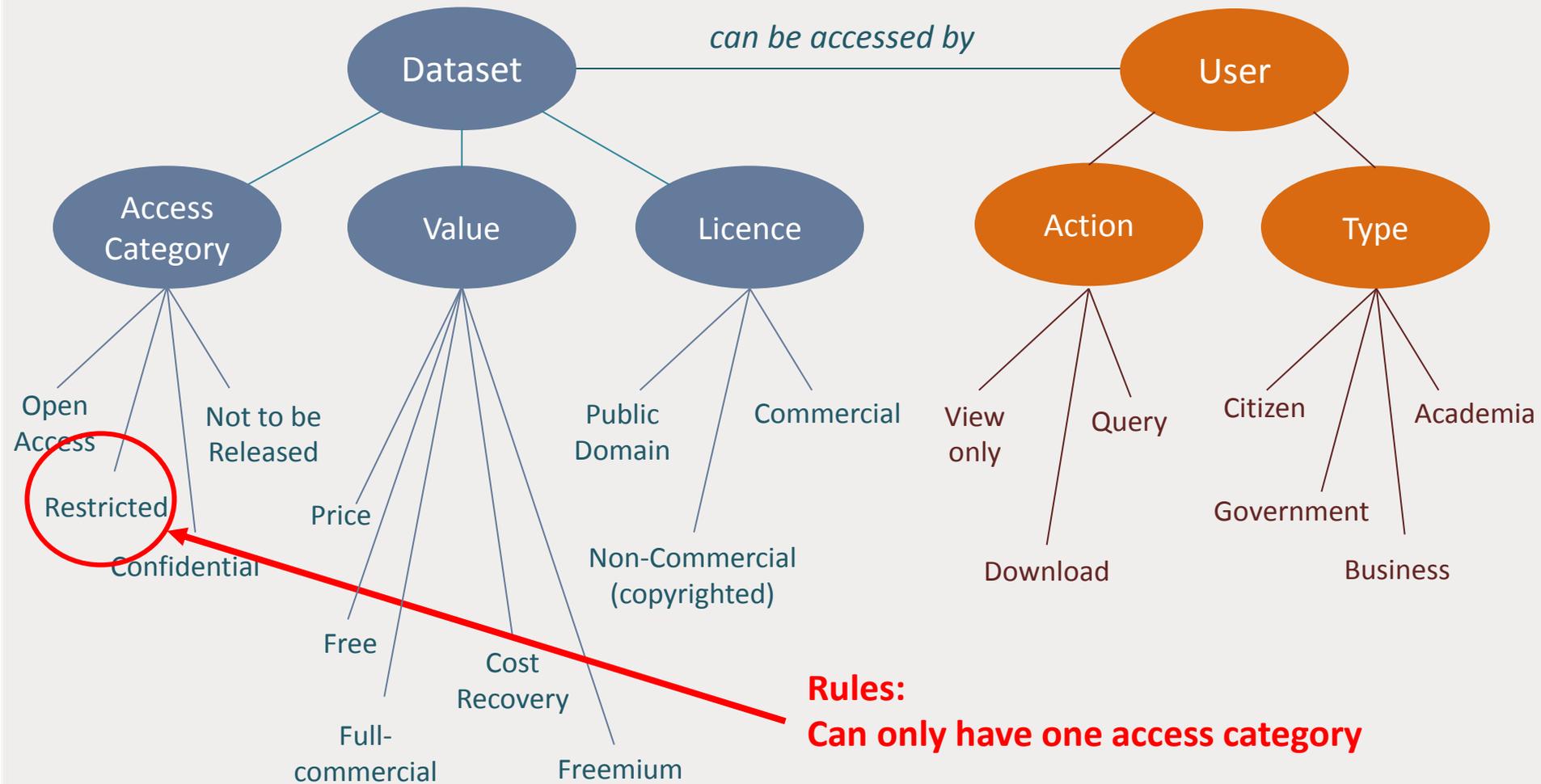
Data Release Ontology Aggregated Data

Rules:
Average price of all State datasets
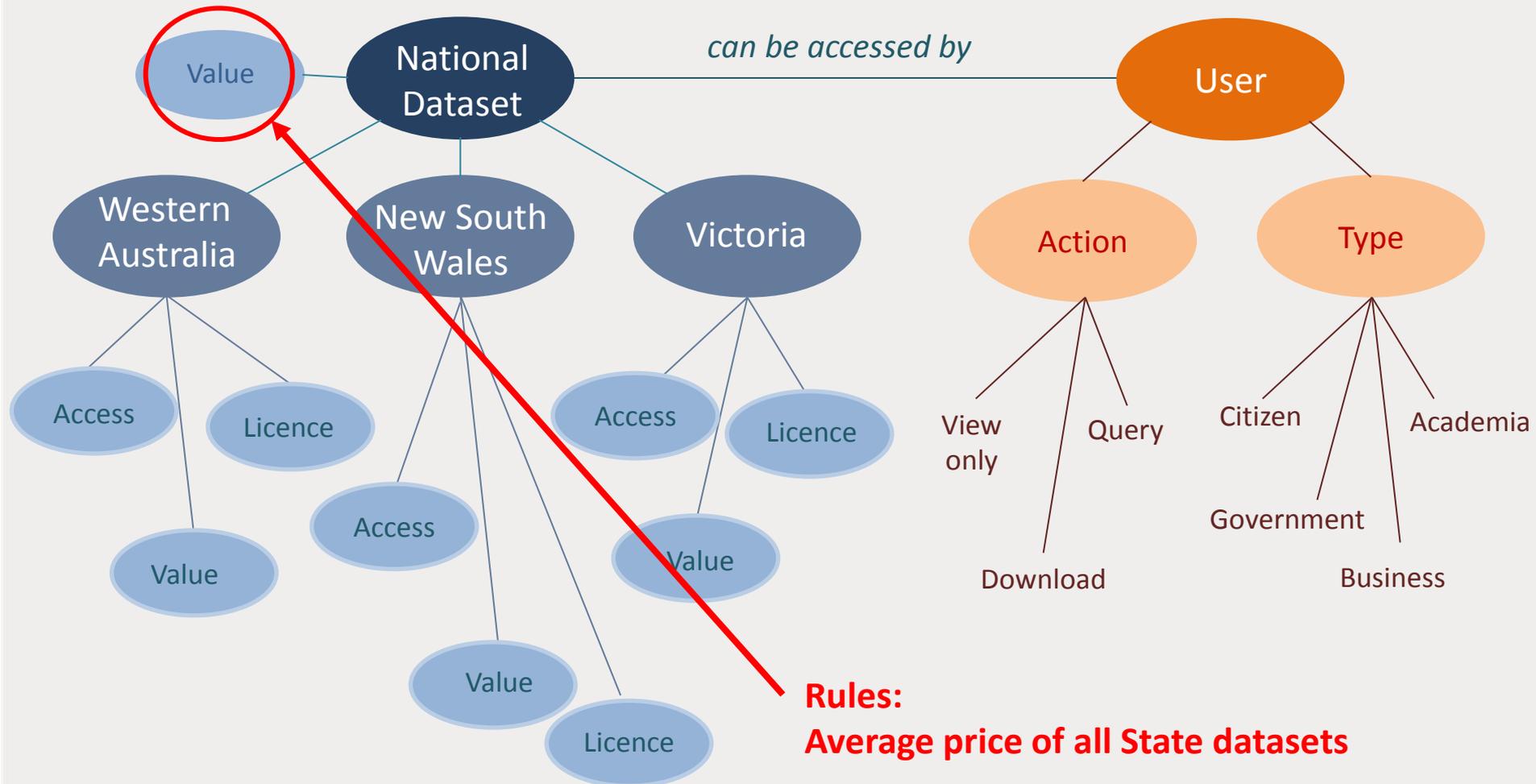
Semantic Web

3 stages of Evolution

Web 1.0 — One-way interaction

Web 2.0 — Two-way participation

Web 3.0 — Linked Data (Semantic Web)
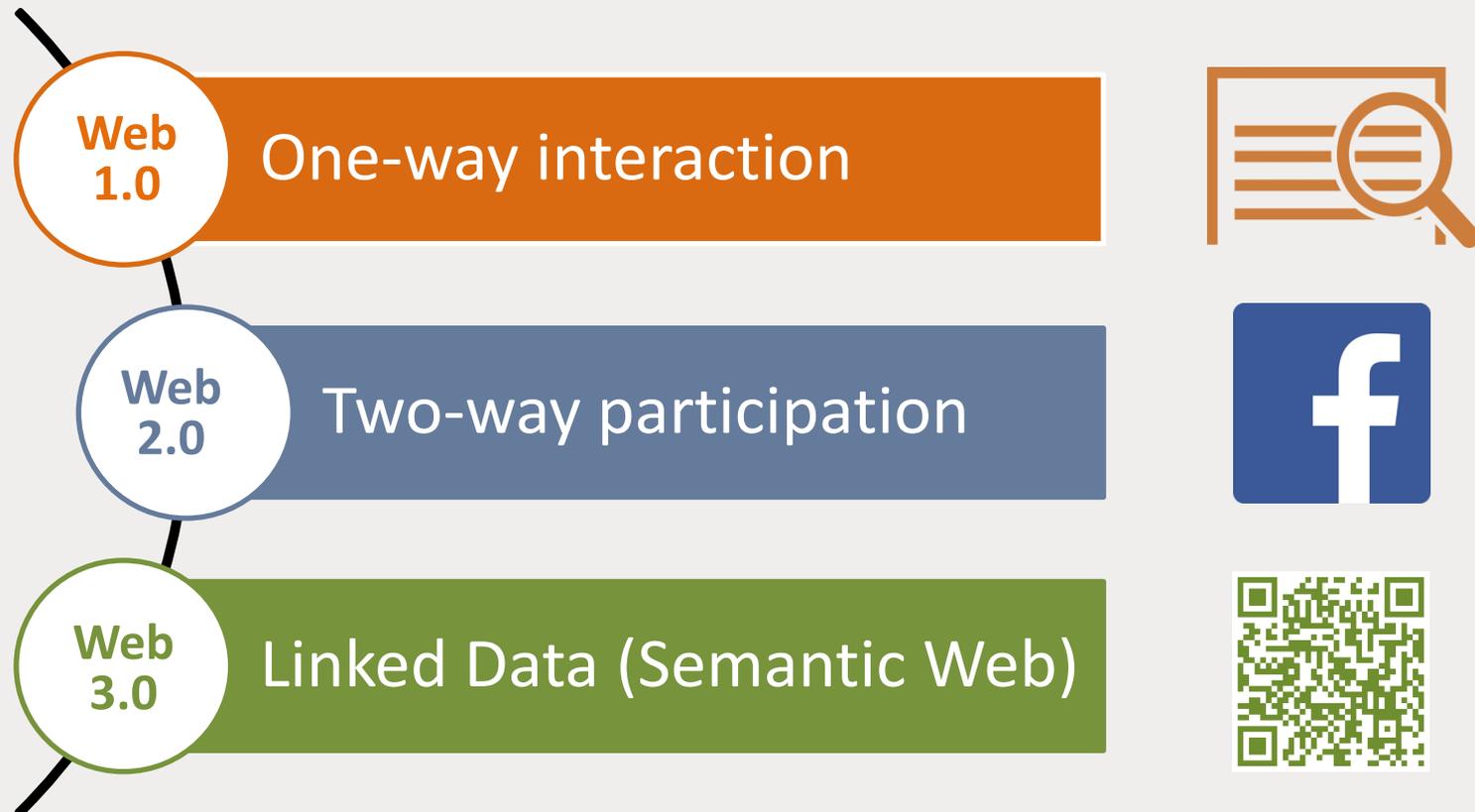
On the Horizon

5 Star Rating
Linked Open Data

Usability
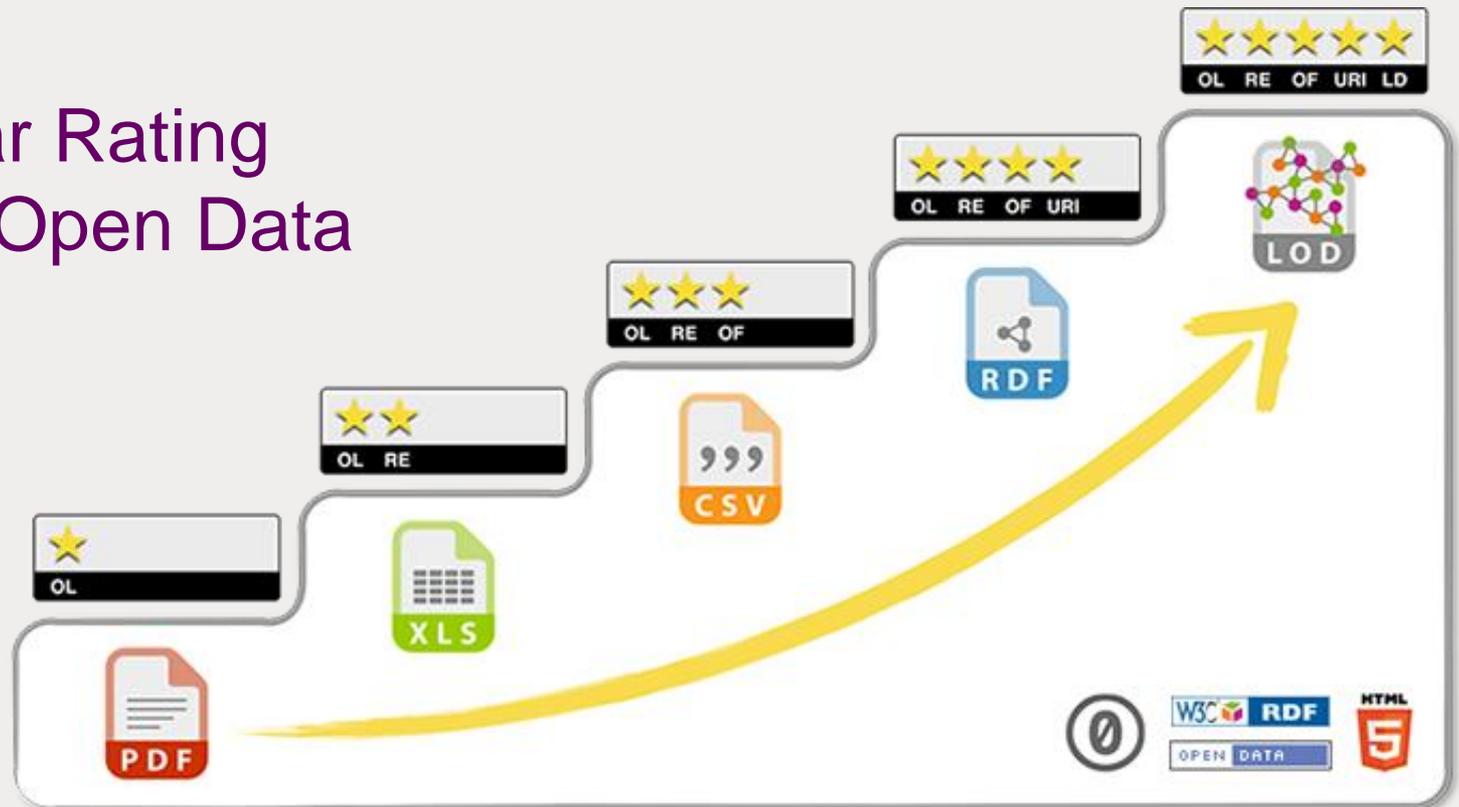
Diagram courtesy: 5 Star Open Data
http://5stardata.info/en/

On the Horizon

Minting Persistent URIs

↓

Real World
Objects
(schools)

Supports Traceability

Schools
http://education.data.gov.uk/id/school

303 See Other

List
http://education.data.gov.uk/doc/school

School
http://education.data.gov.uk/id/school/101140

303 See Other

Document
http://education.data.gov.uk/doc/school/101140

Diagram courtesy data.gov.uk
https://data.gov.uk/resources/uris
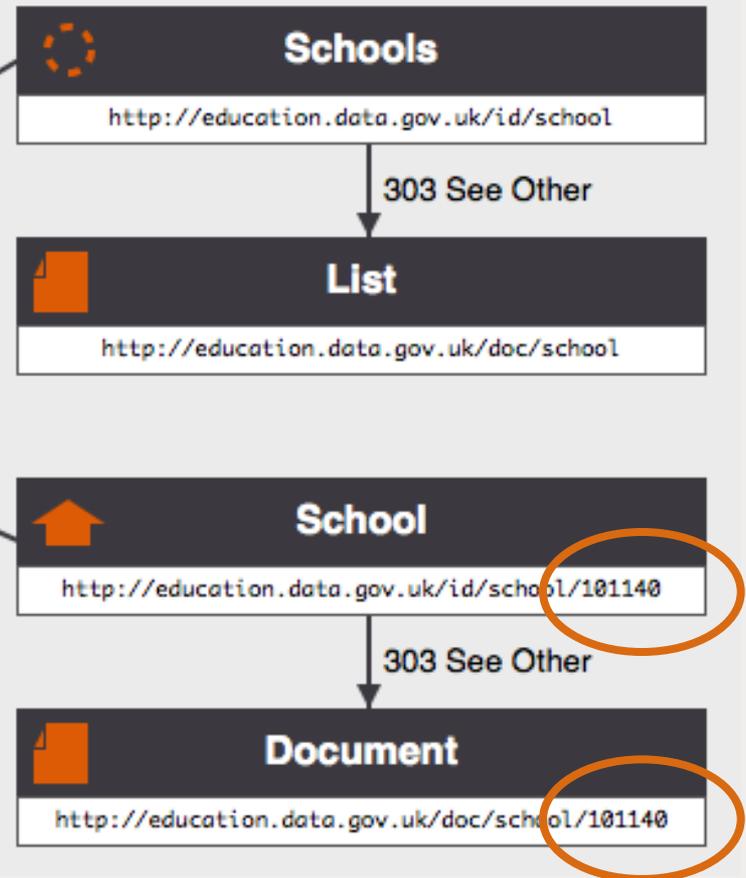
Summary

- Confidence to Open up Access to Data
  - Data sensitivity
  - Individual privacy
  - Copyright
- Simplicity
  - Data custodians set parameters in metadata
- Accountability
- Global standard possible
  - rules created once and reused
  - Unique requirements possible

Thank you