

# Agricultural Research Federation (AgReFed)

A federated agricultural research data community

Bruce Simons, Andrew MacLeod  
Federation University Australia



# Project Partners

Partners in the project were the University of Adelaide, the University of Western Australia, the Western Australian Department of Primary Industries and Regional Development (DPIRD), the University of New England, Federation University Australia and the Commonwealth Scientific and Industrial Research Organisation.



## Acknowledgements

This research was supported by the Australian Research Data Commons (ARDC). ARDC is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy Program (NCRIS).



# Stewardship and Governance Framework

- Aims to improve the sharing and reuse of agricultural data.
- Describes a socio-technical system that:
  - Brings independent organisations together
  - Will guide agricultural data providers; and
  - Be implemented by data provider using common technical infrastructure.
- The ‘social architecture’ key concepts:
  - Independent and autonomous **Data Provider Communities**
  - The collective **AgReFed Community** within which they participate;
- Guiding principles
  - FAIR Data
  - Trusted Repository (Core Data Seal)

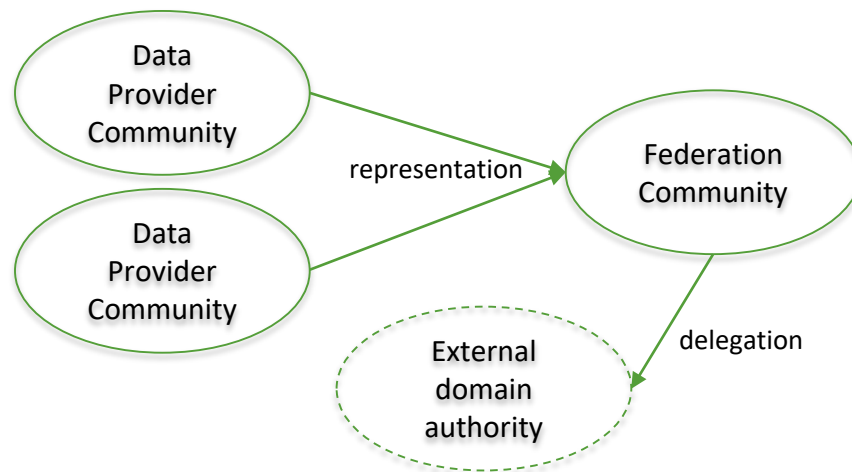
Findable Interoperable  
Accessible Reusable



# AgReFed communities

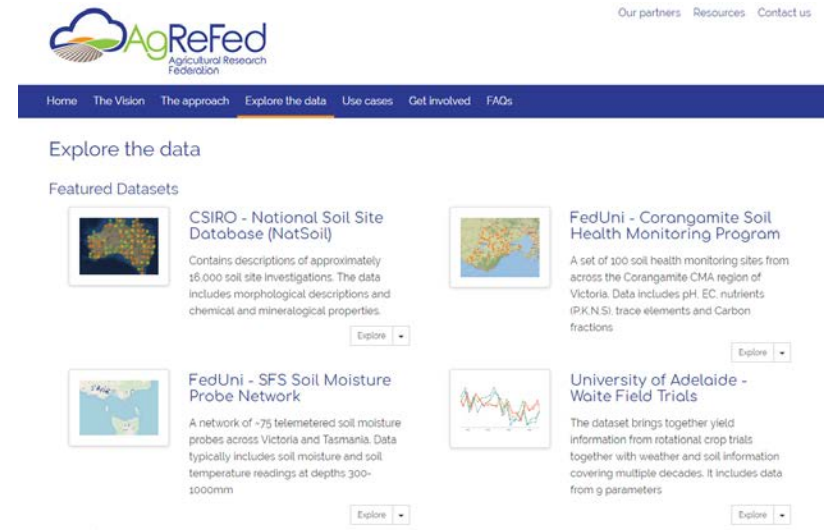
## Communities

- **The Federation Community** – this is the federated community composed of all the various roles that enable the AgReFed to operate;
- **Data Provider Communities** – a community involved in the provision of a data collection or dataset to the AgReFed, represented by multiple roles filled by actors from one or more organisation(s).
- **Domain Authority Communities** – the (virtual) organisations responsible for developing, publishing and governing vocabularies, information models and other kinds of standards. These may be pre-existing external communities that govern standards that are relevant for AgReFed.



# Exemplar datasets

- CSIRO - National Soil Site Database (NatSoil)
- CSIRO – Soil and Landscape Grid of Australia
- FedUni - SFS Soil Moisture Probe Network
- FedUni - Corangamite Soil Health Monitoring Program
- University of Adelaide - Waite Field Trials
- UWA/DPIRD – Frost Nursery Trials
- UNE – SMART Farm SensorNETS

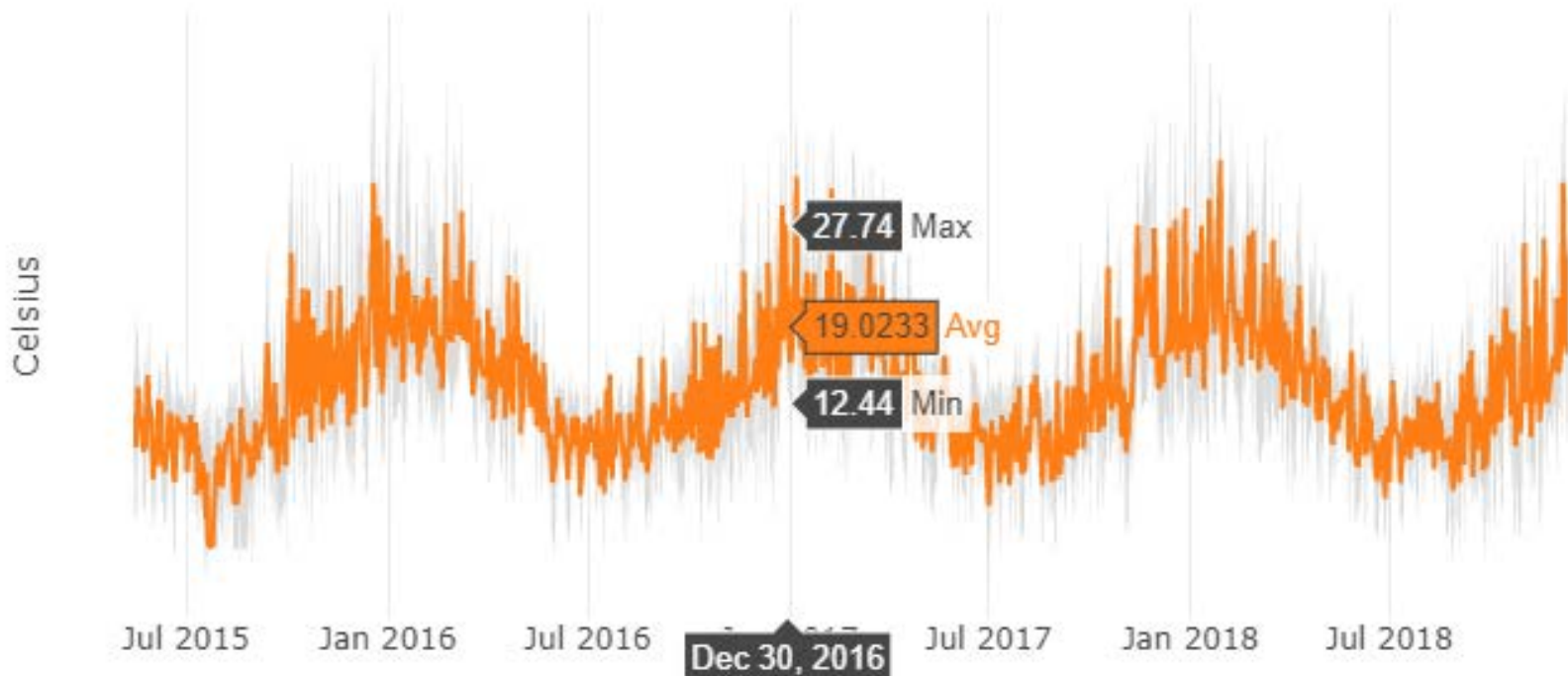


The screenshot shows the AgReFed website interface. At the top right, there are links for "Our partners", "Resources", and "Contact us". Below the logo, a navigation bar contains "Home", "The Vision", "The approach", "Explore the data", "Use cases", "Get involved", and "FAQs". The main content area is titled "Explore the data" and features a "Featured Datasets" section. This section displays four dataset cards, each with a thumbnail image, a title, a brief description, and an "Explore" button. The datasets listed are: CSIRO - National Soil Site Database (NatSoil), FedUni - Corangamite Soil Health Monitoring Program, FedUni - SFS Soil Moisture Probe Network, and University of Adelaide - Waite Field Trials.

# Sensor network - Soil Moisture and Weather



SFS44 - Inverleigh trial site - AirTemperature datastream

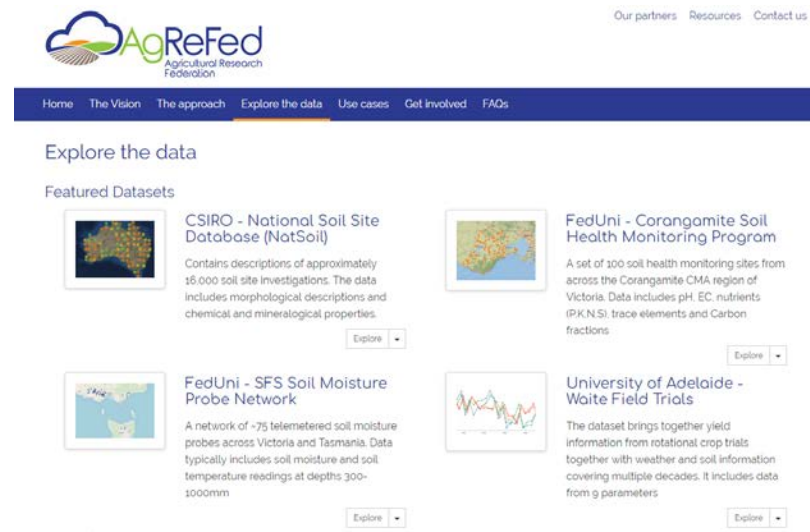


...i?&skip=100",  
...ations(147267368)/Datastream",  
.../Observations(147267368)/FeatureOfInte  
...dand \}\",  
...acedon  
...Rid  
...ations(147267369)/Datastream",  
.../Observations(147267369)/FeatureOfInte  
...)\",  
...Melton  
...ations(147267371)/Datastream",  
.../Observations(147267371)/FeatureOfInte  
...)\",  
...W  
...ations(147267372)/Datastream",  
.../Observations(147267372)/FeatureOfInte  
...)\",  
...Port  
...brysdale  
...Ocean Grove  
...Torquay  
...Anglesea

Circles coloured by current soil moisture  
across depth profile

# Exemplar datasets

- CSIRO - National Soil Site Database (NatSoil)
- CSIRO – Soil and Landscape Grid of Australia
- FedUni - SFS Soil Moisture Probe Network
- FedUni - Corangamite Soil Health Monitoring Program
- University of Adelaide - Waite Field Trials
- UWA/DPIRD – Frost Nursery Trials
- UNE – SMART Farm SensorNETS



The screenshot shows the AgReFed website interface. At the top right, there are links for "Our partners", "Resources", and "Contact us". A dark blue navigation bar contains links for "Home", "The Vision", "The approach", "Explore the data", "Use cases", "Get involved", and "FAQs". Below the navigation bar, the heading "Explore the data" is followed by a section titled "Featured Datasets". This section displays four dataset cards, each with a thumbnail image, a title, a brief description, and an "Explore" button with a dropdown arrow.

Dataset Title	Description
CSIRO - National Soil Site Database (NatSoil)	Contains descriptions of approximately 16,000 soil site investigations. The data includes morphological descriptions and chemical and mineralogical properties.
FedUni - Corangamite Soil Health Monitoring Program	A set of 100 soil health monitoring sites from across the Corangamite CMA region of Victoria. Data includes pH, EC, nutrients (P, K, N, S), trace elements and Carbon fractions.
FedUni - SFS Soil Moisture Probe Network	A network of ~75 telemetered soil moisture probes across Victoria and Tasmania. Data typically includes soil moisture and soil temperature readings at depths 300-1000mm.
University of Adelaide - Waite Field Trials	The dataset brings together yield information from rotational crop trials together with weather and soil information covering multiple decades. It includes data from 9 parameters.

# Soil sampling and observations

CSIRO Linked Data Registry

Browse

About

Advanced ▾



<http://registry.it.csiro.au> / [sandbox](#) / [student](#) / [xavier](#) / [method](#) / [\\_mg-4](#)

## Entry: 15 a1 mG

URI: <http://registry.it.csiro.au/sandbox/student/xavier/method/mg-4>

Exchangeable bases (Ca<sup>2+</sup>,Mg<sup>2+</sup>,Na<sup>+</sup>,K<sup>+</sup>) - 1M ammonium chloride at pH 7.0, no pretreatment for soluble salts

## Definition

alt label	Exch. Mg <sup>++</sup>
definition	Exchangeable bases (Ca <sup>2+</sup> ,Mg <sup>2+</sup> ,Na <sup>+</sup> ,K <sup>+</sup> ) - 1M ammonium chloride at pH 7.0, no pretreatment for soluble salts
description	Exchangeable bases (Ca <sup>2+</sup> ,Mg <sup>2+</sup> ,Na <sup>+</sup> ,K <sup>+</sup> ) - 1M ammonium chloride at pH 7.0, no pretreatment for soluble salts
label	15A1_MG
notation	15A1_MG
pref label	15A1_MG
type	Concept
unit	MilliEquivalent per HectoGram

- Regional soil monitoring program

- National soil database

- Approx 70,000 surveys

- 270,000 samples

- WFS services (complex)

- SF\_Specimen, Soil and

## OM Observation features

## Links

Has unit of measure

- ANZSoilML Spatial information cached into postgres
- Visualised by Geoserver point cluster rendering transformation

```

<observation gml:id="soil_observation_503.SCP.YAN24.1.1.1.15A1_MG.1">
  <identifier codeSpace="http://www.clw.csiro.au/aclep/archive">
    http://www.anzsoil.org/data/csiro-
    soil/anzsoilm1201/observation/soil_observation_503.SCP.YAN24.1.1.1.1.1.
    15A1_MG.1
  </identifier>
  <phenomenonTime>
    <TimeInstant gml:timePosition="2010-01-19T00:00:00Z" gml:timePosition">
    </TimeInstant>
  </phenomenonTime>
  <resultTime>
    <TimeInstant gml:timePosition="2010-01-19T00:00:00Z" gml:timePosition">
    </TimeInstant>
  </resultTime>
  <procedure>
    <name gml:method="http://registry.it.csiro.au/sandbox/student/
    x/15A1_MG.1" gml:title="Exch. Mg++"/>
    </name>
  </procedure>
  <parameter>
    <name gml:method="http://data.anzsoil.org/def/property/x/depth"/>
    </name>
  </parameter>
  <value>
    <anzsml:DepthQuantityRange>
      <anzsml:upperBoundary>
        <swe:Quantity>

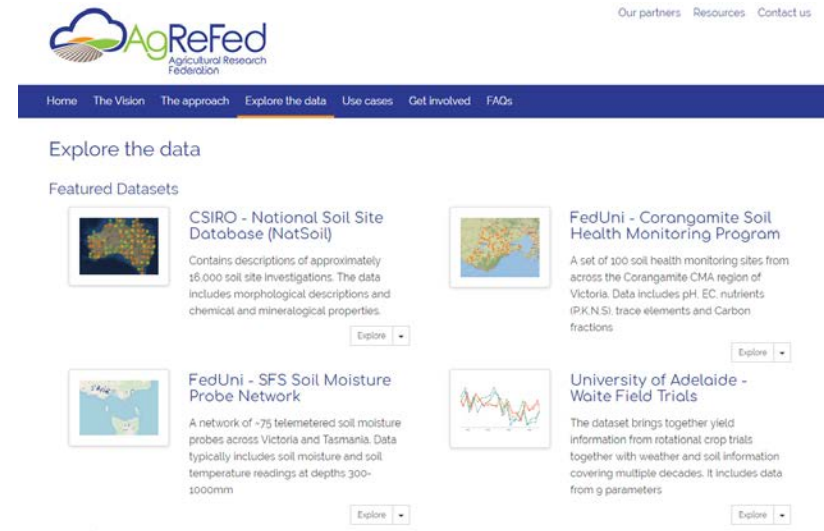
```





# Exemplar datasets

- CSIRO - National Soil Site Database (NatSoil)
- CSIRO – Soil and Landscape Grid of Australia
- FedUni - SFS Soil Moisture Probe Network
- FedUni - Corangamite Soil Health Monitoring Program
- University of Adelaide - Waite Field Trials
- **UWA/DPIRD – Frost Nursery Trials**
- UNE – SMART Farm SensorNETS



The screenshot shows the AgReFed website interface. At the top right, there are links for "Our partners", "Resources", and "Contact us". Below the logo, a navigation bar contains "Home", "The Vision", "The approach", "Explore the data", "Use cases", "Get involved", and "FAQs". The main content area is titled "Explore the data" and lists "Featured Datasets".

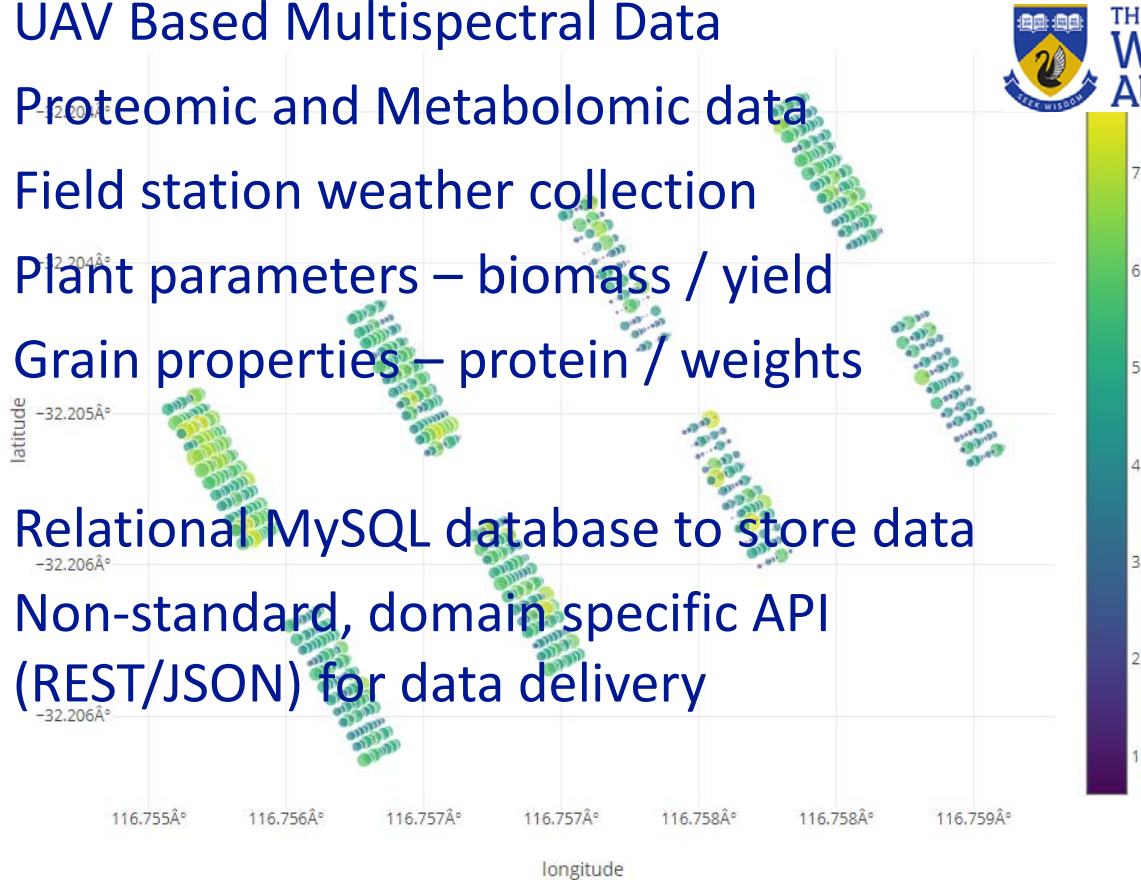
Dataset Name	Description
CSIRO - National Soil Site Database (NatSoil)	Contains descriptions of approximately 16,000 soil site investigations. The data includes morphological descriptions and chemical and mineralogical properties.
FedUni - Corangamite Soil Health Monitoring Program	A set of 100 soil health monitoring sites from across the Corangamite CMA region of Victoria. Data includes pH, EC, nutrients (P, K, N, S), trace elements and Carbon fractions.
FedUni - SFS Soil Moisture Probe Network	A network of ~75 telemetered soil moisture probes across Victoria and Tasmania. Data typically includes soil moisture and soil temperature readings at depths 300-1000mm.
University of Adelaide - Waite Field Trials	The dataset brings together yield information from rotational crop trials together with weather and soil information covering multiple decades. It includes data from 9 parameters.

# Field and greenhouse trial research

University of Western Australia - plot yield vs location data

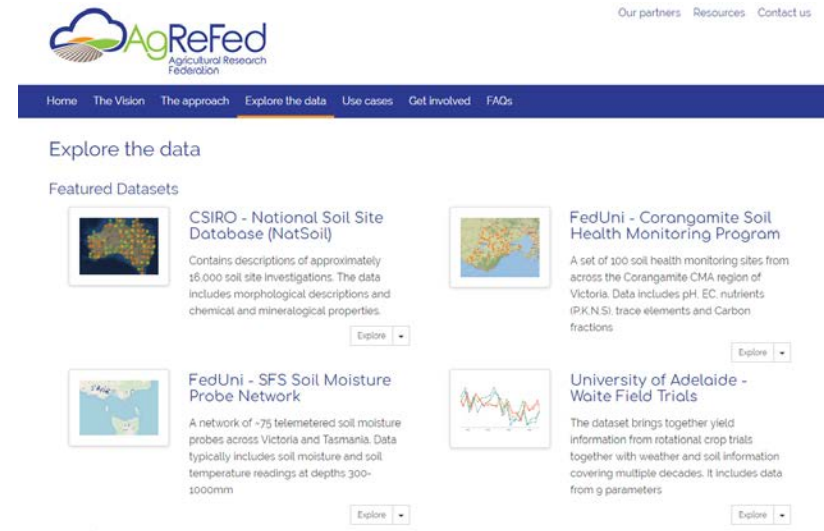
- Ground Based Hyperspectral Data
- UAV Based Multispectral Data
- Proteomic and Metabolomic data
- Field station weather collection
- Plant parameters – biomass / yield
- Grain properties – protein / weights
- Relational MySQL database to store data
- Non-standard, domain specific API (REST/JSON) for data delivery

Plot yield vs location



# Exemplar datasets

- CSIRO - National Soil Site Database (NatSoil)
- CSIRO – Soil and Landscape Grid of Australia
- FedUni - SFS Soil Moisture Probe Network
- FedUni - Corangamite Soil Health Monitoring Program
- **University of Adelaide - Waite Field Trials**
- UWA/DPIRD – Frost Nursery Trials
- UNE – SMART Farm SensorNETS



The screenshot shows the AgReFed website interface. At the top right, there are links for "Our partners", "Resources", and "Contact us". Below the logo is a navigation bar with links for "Home", "The Vision", "The approach", "Explore the data", "Use cases", "Get involved", and "FAQs". The main content area is titled "Explore the data" and features a section for "Featured Datasets".

Dataset Name	Description
CSIRO - National Soil Site Database (NatSoil)	Contains descriptions of approximately 16,000 soil site investigations. The data includes morphological descriptions and chemical and mineralogical properties.
FedUni - Corangamite Soil Health Monitoring Program	A set of 100 soil health monitoring sites from across the Corangamite CMA region of Victoria. Data includes pH, EC, nutrients (P, K, N, S), trace elements and Carbon fractions.
FedUni - SFS Soil Moisture Probe Network	A network of ~75 telemetered soil moisture probes across Victoria and Tasmania. Data typically includes soil moisture and soil temperature readings at depths 300-1000mm.
University of Adelaide - Waite Field Trials	The dataset brings together yield information from rotational crop trials together with weather and soil information covering multiple decades. It includes data from 9 parameters.

# Historical trial research

## Waite Permanent Rotation Trial

Plot **Permanent rotation trial since 1920**

Crop/Pasture **Wheat**

Rotation type **wheat followed by fallows (W\_F)**

Observed property **Grain Yield**

Start Year **1925**

End Year **1931**

Chart type **Plot types by rotation type**

- Data collected by CSIRO into bespoke spreadsheet
- O&M-relational DB developed – FedUni and UofA

```
id: 9999,
year: "1925",
plot number:
```



THE UNIVERSITY  
of ADELAIDE

ag-semantic.org/gi

"Grain Yield",  
<http://id.agrsemantic.org/gacs/C2241>,

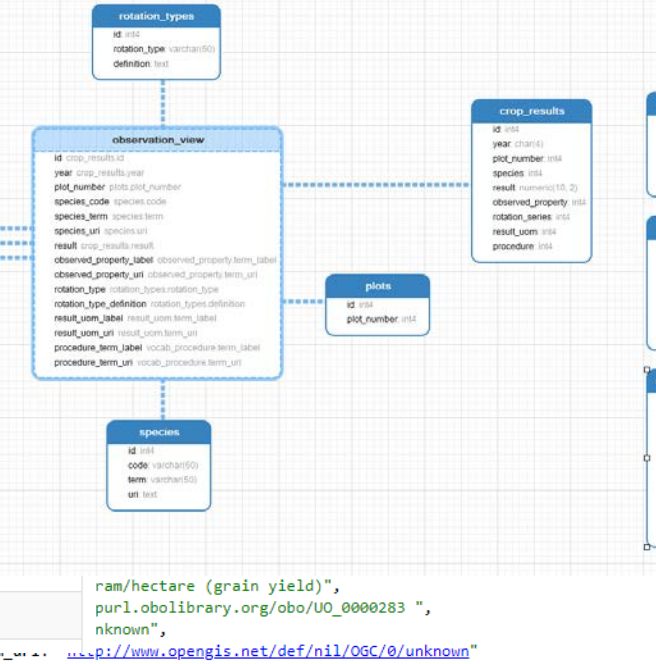
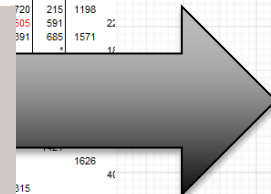
: "wheat followed by fallows",  
ram/hectare (grain yield)",  
purl.obolibrary.org/obo/UO\_0000283",  
nknown",  
p://www.opengis.net/def/nil/OGC/0/unknown"

✕ Clear

Download CSV

Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33		
1925	2484		2654	746	2681		2468	5383		2472	716		2252	5621		2504	2688	5982	800	2501	6045	800	3002	5717			549	5673		2888					
1926		2965	2937	1102		2664		4523	2688	3568	2675	1136	2580	4623	2641	1520	4553	1122	3266	4321	1257	3024	4439		3152	2876			3649	5673	3011				
1927	2381		1794	1450	1891	2664		1913	4822	3568	2681	276	2580	4623	2641	1520	4553	1122	3266	4321	1257	3024	4439		3152	2876			3649	5673	3011				
1928		2044																																	

Output is JSON using an O&M design pattern.  
Use of standardised vocabularies



```
ram/hectare (grain yield)",
purl.obolibrary.org/obo/UO_0000283",
nknown",
p://www.opengis.net/def/nil/OGC/0/unknown"
```



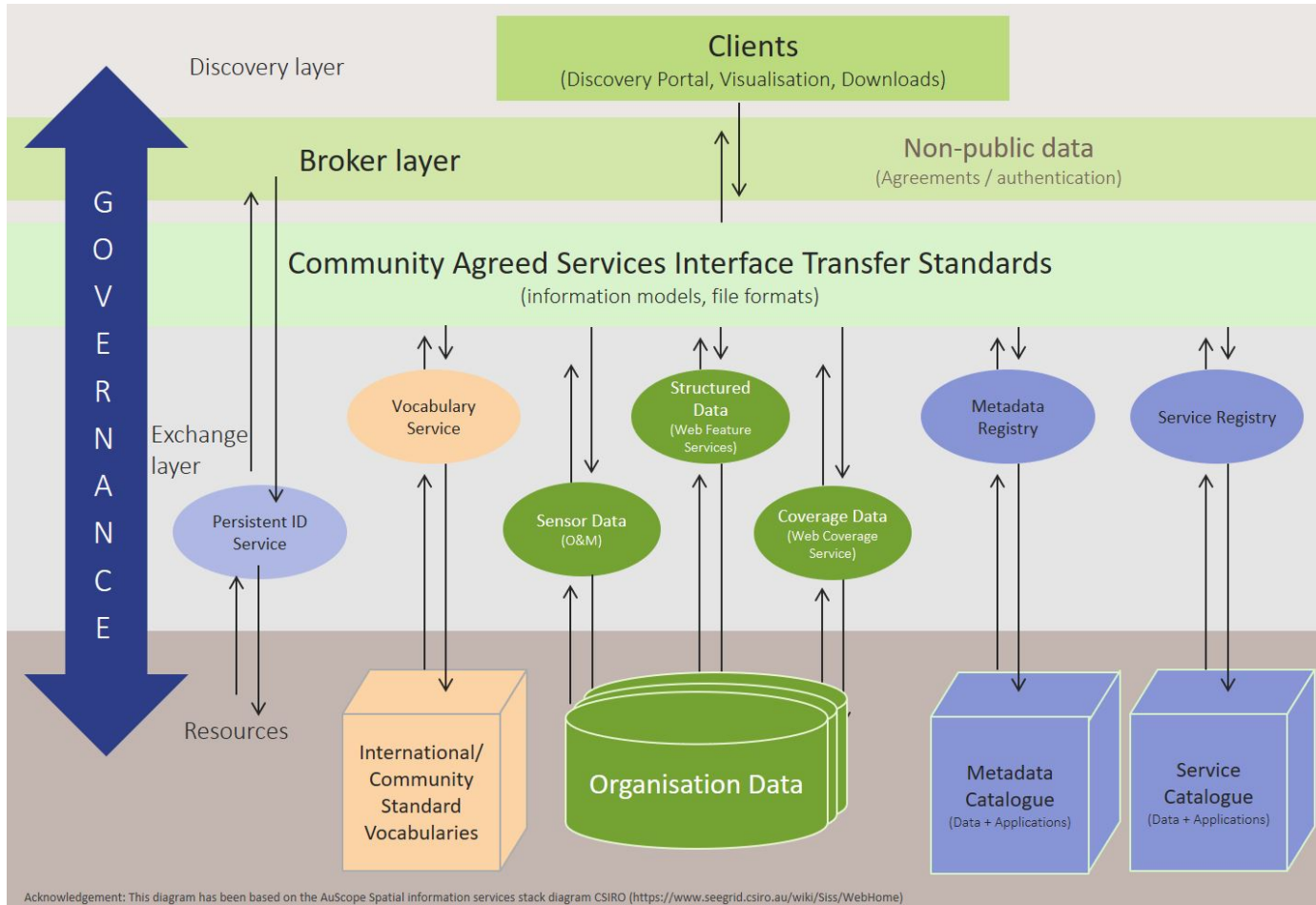
# From 'my data' to 'our FAIR data'

- Individual providers' heterogeneous data
- Data providers assessing their data FAIR-ness and Trusted-ness
- Maturity of provider starting point determined level of standardisation reached
  - Structure - full OGC schema vs OGC pattern vs local structure
  - Content – external vocabulary services vs local terminology
  - Technology – OGC compliant stack vs custom API's



# Extra Slides

# AgReFed Computational Viewpoint Stack



This diagram provides a high level view of the interactions between provider and users via provider and common computational components.

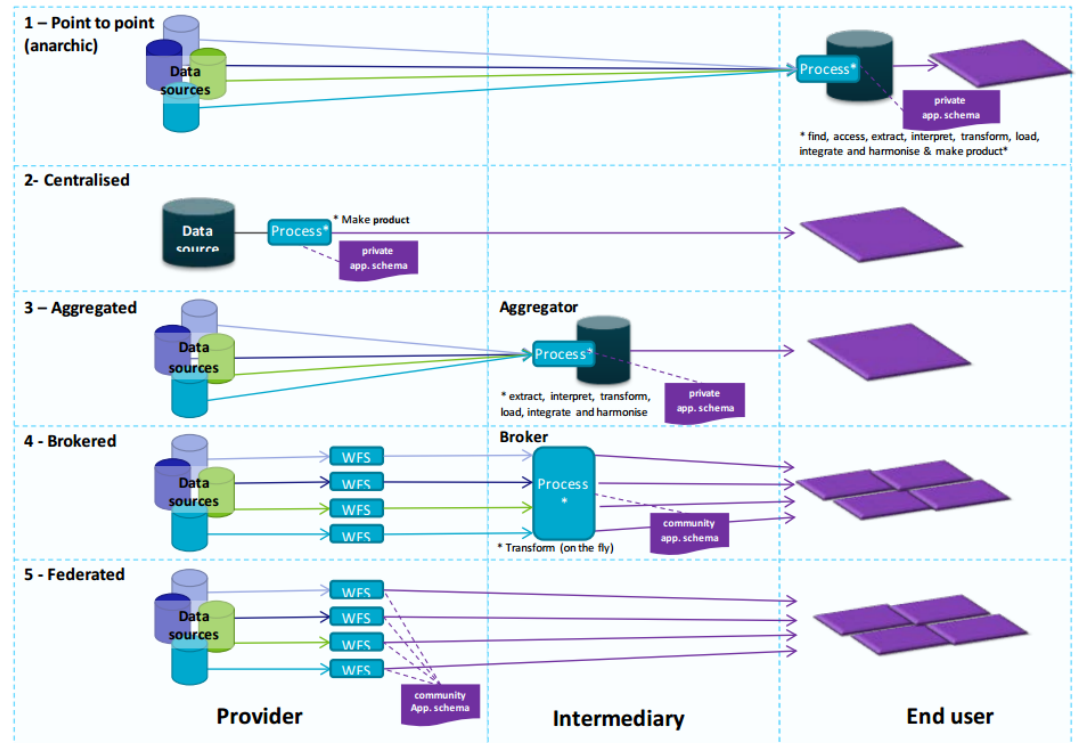


# Distributed data supply chain patterns

The preference for the AgReFed is to use a **federated** approach to data supply.

However, where it makes integration easier to achieve, elements of the brokering and aggregation patterns will also be used.

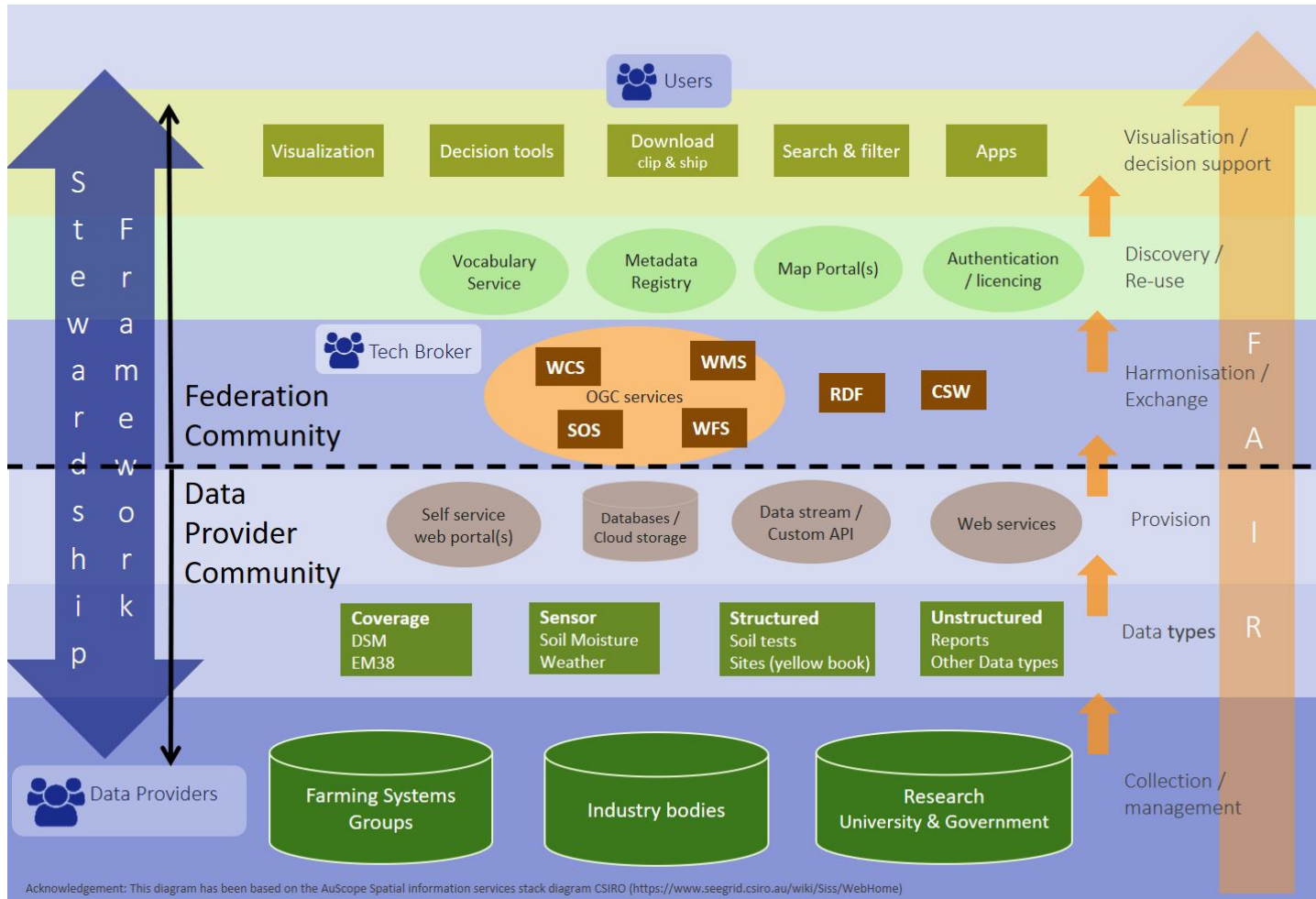
- For federated and brokered data, source data resides with the data provider system. Both solutions encourage currency and validity of data.
- A standards based Service Oriented Architecture ([https://en.wikipedia.org/wiki/Service-oriented\\_architecture](https://en.wikipedia.org/wiki/Service-oriented_architecture)) is utilised, including metadata cataloguing and vocabulary linking. These will provide information about the data using standardised terms.
- Data is transformed from services developed using a community application schema, or in the case of aggregation, using a respected standards based aggregation platform which has broad appeal.



Distributed data supply chain patterns (Box et al., 2015)

1 <https://publications.csiro.au/rpr/pub?pid=csiro:EP155525>

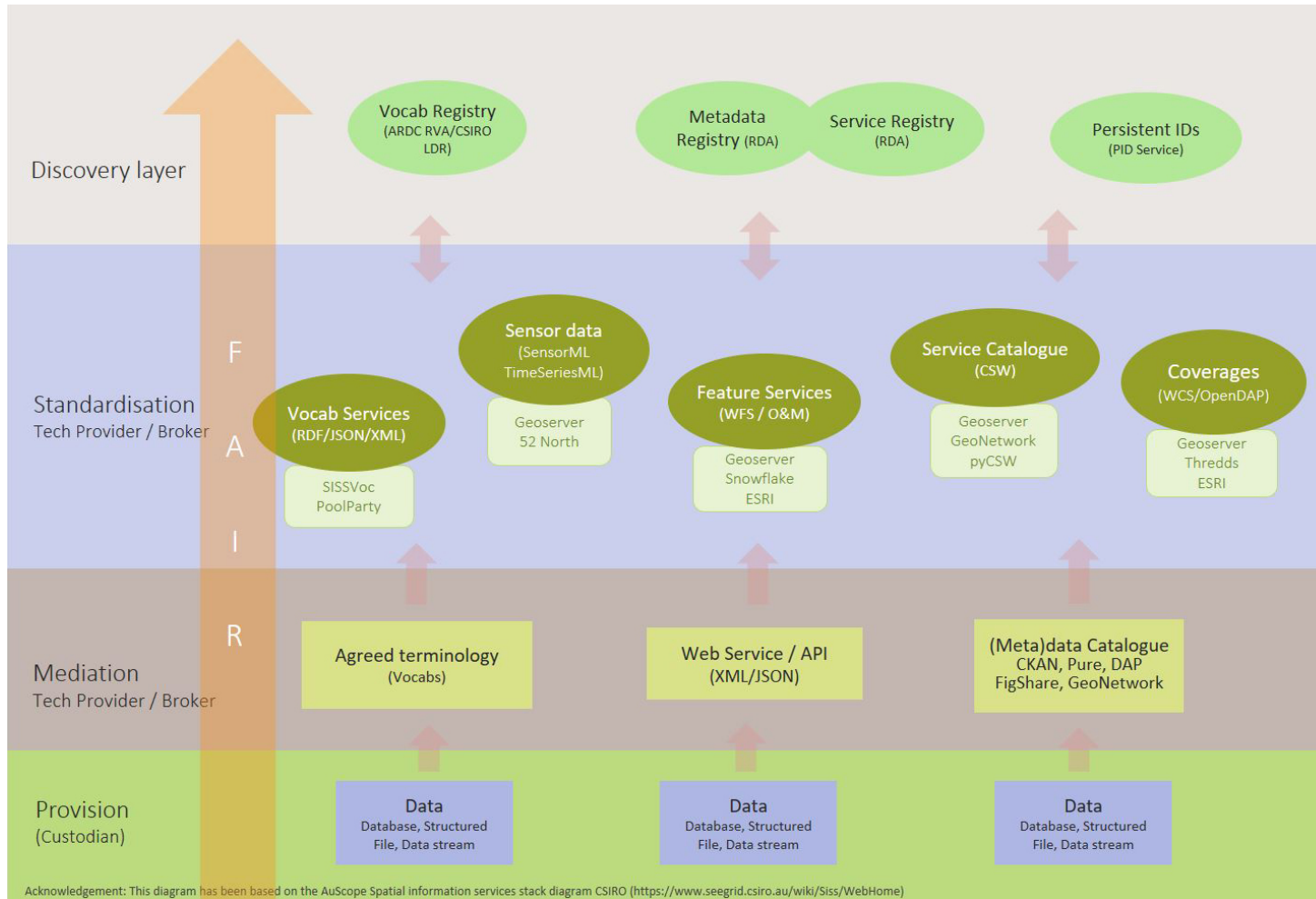
# AgReFed Engineering viewpoint



This diagram provides a view of the components, mechanisms and functions required to support distributed interactions among objects in the system, including whose responsibility they are. When this diagram is fleshed out in more detail it will show where components are deployed.

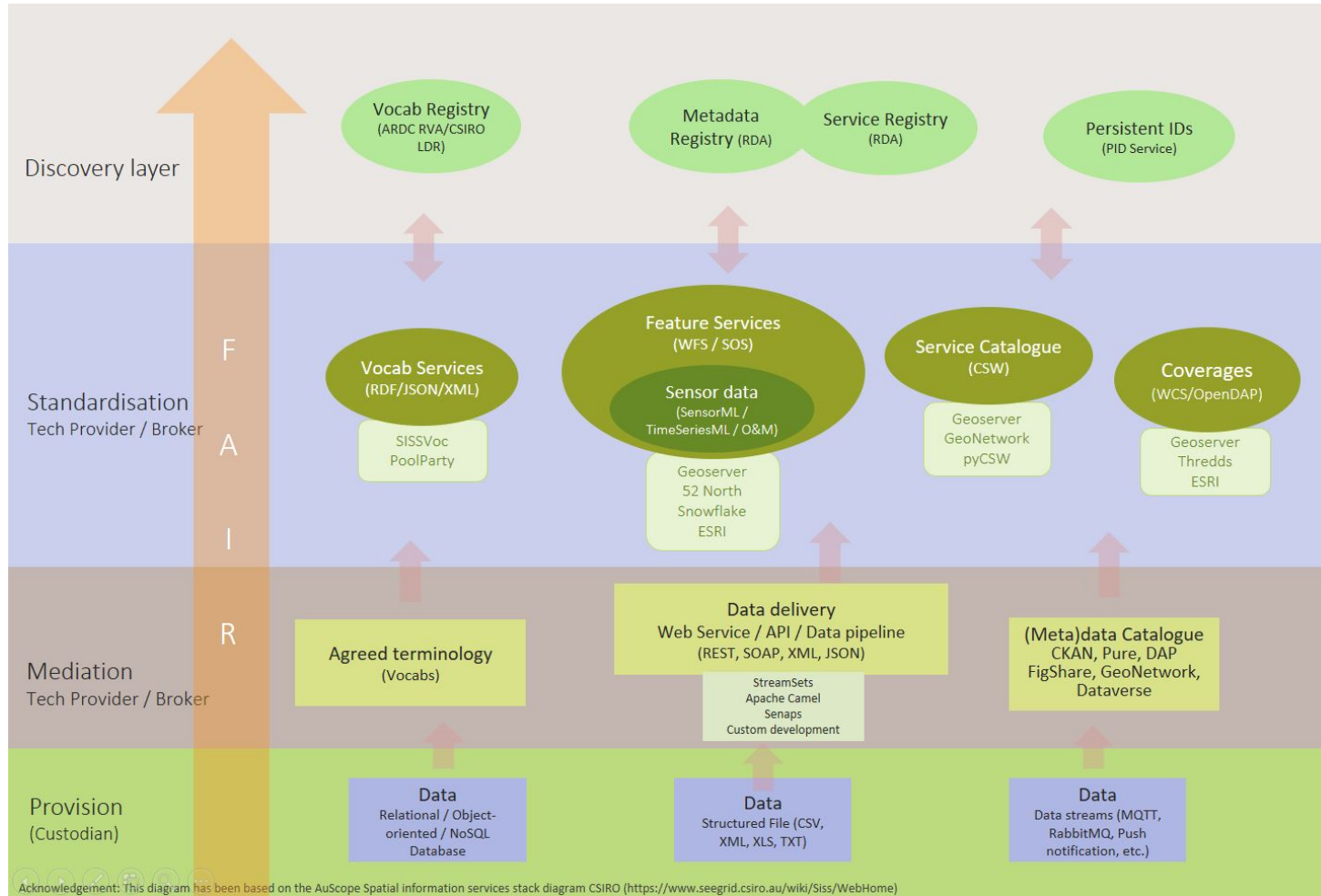
Acknowledgement: This diagram has been based on the AuScope Spatial information services stack diagram CSIRO (<https://www.seegrid.csiro.au/wiki/Siss/WebHome>)

# AgReFed Technology viewpoint – Feature instance



This diagram provides a **(spatial) 'feature oriented' view** of the technology stack which emphasises the delivery of geospatial feature data

# AgReFed Technology viewpoint – Sensors instance



This diagram provides a sensor oriented view of the technology stack which emphasises the delivery of sensor data.