

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)

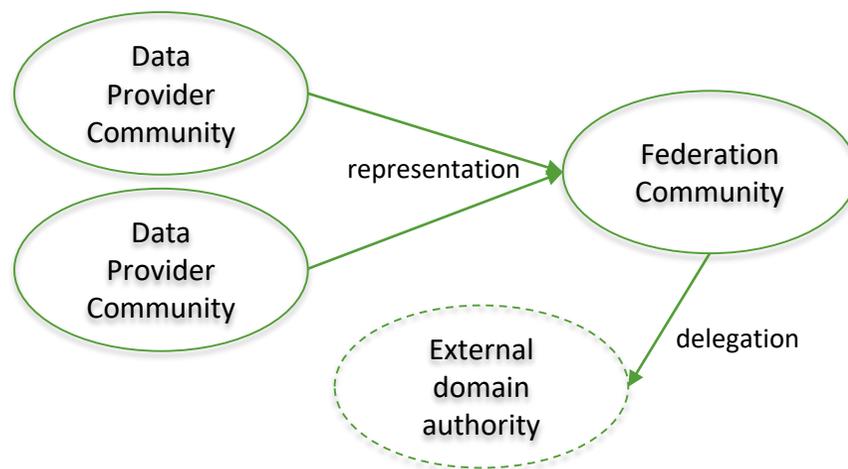
Findable Interoperable
Accessible Reusable



AgReFed communities

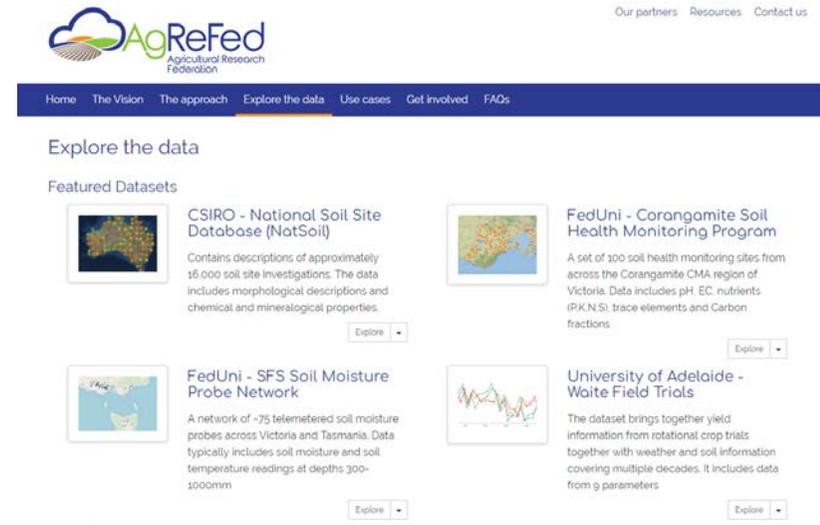
Communities

- **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.

Our partners Resources Contact us

Home The Vision The approach Explore the data Use cases Get involved FAQs

Explore the data

Featured Datasets

 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. [Explore](#)

 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. [Explore](#)

 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. [Explore](#)

 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. [Explore](#)

Sensor network - Soil Moisture and Weather



SFS44 - Inverleigh trial site - AirTemperature datastream



Map showing the location of the Inverleigh trial site (SFS44) in Victoria, Australia. The map includes labels for locations such as Port Fairy, Warrnambool, Koroit, Cobden, Anglesea, Torquay, Ocean Grove, and Brysdale. A data stream is overlaid on the map, showing the location of the sensor network. The data stream is represented by a series of colored circles (red, orange, yellow, green, blue) indicating soil moisture levels across different depths. The data stream is labeled with the following information:

```
...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
```

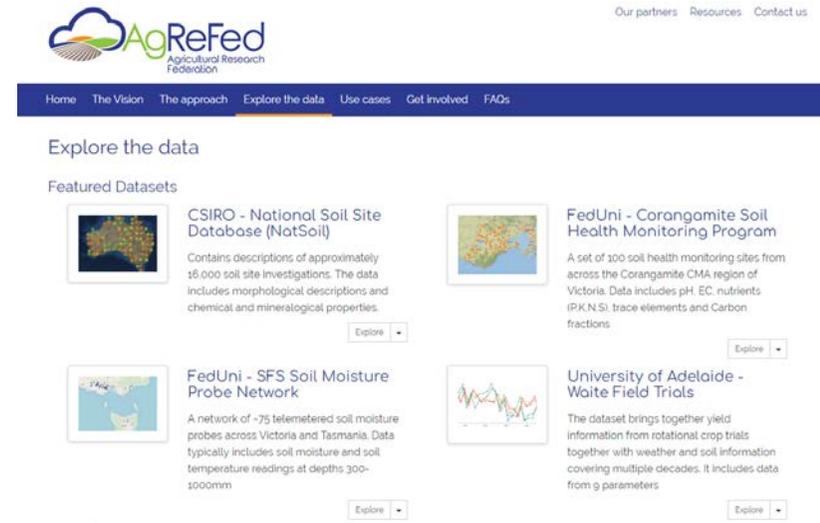
Circles coloured by current soil moisture across depth profile

Southern Farming Systems



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 this is a dark blue 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 "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.

Our partners Resources Contact us

Home The Vision The approach Explore the data Use cases Get involved FAQs

Explore the data

Featured Datasets

 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. [Explore](#)

 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. [Explore](#)

 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. [Explore](#)

 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. [Explore](#)

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 (Ca²⁺,Mg²⁺,Na⁺,K⁺) - 1M ammonium chloride at pH 7.0, no pretreatment for soluble salts

Definition

alt label	Exch. Mg ⁺⁺
definition	Exchangeable bases (Ca ²⁺ ,Mg ²⁺ ,Na ⁺ ,K ⁺) - 1M ammonium chloride at pH 7.0, no pretreatment for soluble salts
description	Exchangeable bases (Ca ²⁺ ,Mg ²⁺ ,Na ⁺ ,K ⁺) - 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 postgis
- Visualised by Geoserver point cluster rendering transformation

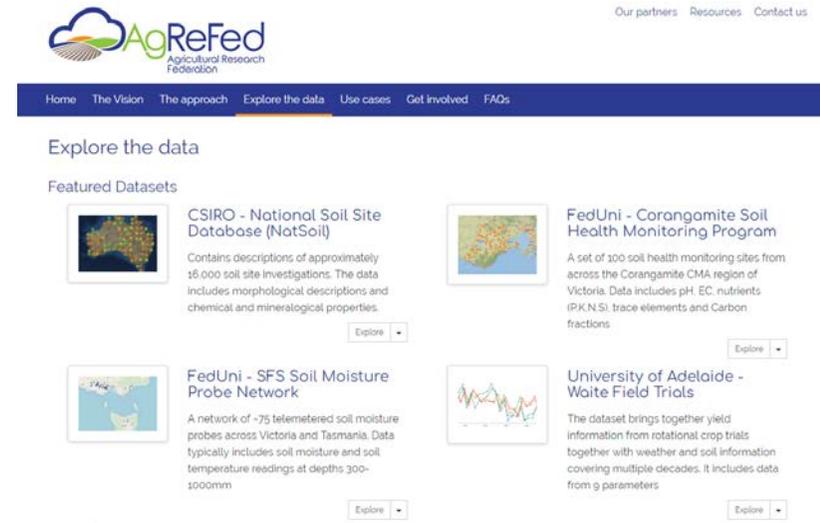
```

<gml:observation gml:id="soil_observation_503.SCP.YAN24.1.1.1.15A1_MG.1">
  <gml:identifier codeSpace="http://www.ciw.csiro.au/aclep/archive">
    http://www.anzsoil.org/data/csiro-
  </gml:identifier>
  <gml:identifier codeSpace="http://www.anzsoil.org/data/csiro-
    soil/anzsoilm1201/observation/soil_observation_503.SCP.YAN24.1.1.1.1.1.1.15A1_MG.1">
    http://www.anzsoil.org/data/csiro-
  </gml:identifier>
  <gml:phenomenonTime>
    <gml:TimeInstant gml:id="phenomenonTime_503.SCP.YAN24.1.1.1.1.15A1_MG.1.0000">
      <gml:timePosition>0101-94-19</gml:timePosition>
    </gml:TimeInstant>
  </gml:phenomenonTime>
  <gml:sultTime>
    <gml:TimeInstant gml:id="sultTime_503.SCP.YAN24.1.1.1.15A1_MG.1.0000">
      <gml:timePosition>0000</gml:timePosition>
    </gml:TimeInstant>
  </gml:sultTime>
  <gml:procedure>
    <gml:Text gml:id="procedure_503.SCP.YAN24.1.1.1.15A1_MG.1.0000">
      Exchangeable bases (Ca2+, Mg2+, Na+, K+) - 1M ammonium chloride at pH 7.0, no pretreatment for soluble salts
    </gml:Text>
  </gml:procedure>
  <gml:unitOfMeasure>
    <gml:Text gml:id="unitOfMeasure_503.SCP.YAN24.1.1.1.15A1_MG.1.0000">
      MilliEquivalent per HectoGram
    </gml:Text>
  </gml:unitOfMeasure>
  <gml:depth>
    <gml:Text gml:id="depth_503.SCP.YAN24.1.1.1.15A1_MG.1.0000">
      DepthQuantityRange
    </gml:Text>
  </gml:depth>
  <gml:upperBoundary>
    <gml:Text gml:id="upperBoundary_503.SCP.YAN24.1.1.1.15A1_MG.1.0000">
      Quantity
    </gml:Text>
  </gml:upperBoundary>
</gml:observation>
  
```



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 "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.

Our partners Resources Contact us

Home The Vision The approach Explore the data Use cases Get involved FAQs

Explore the data

Featured Datasets

 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. [Explore](#)

 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. [Explore](#)

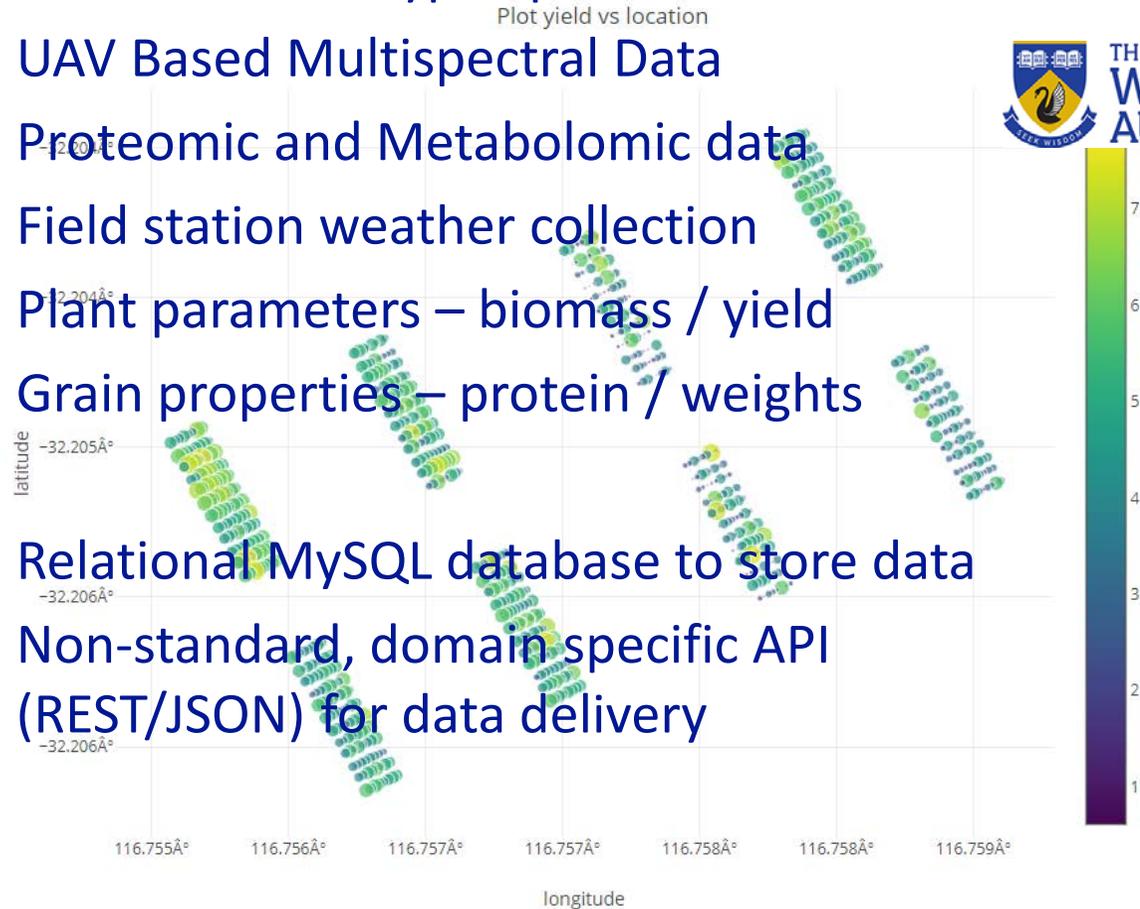
 FedUni - SFS Soil Moisture Probe Network
A network of ~75 telemetered soil moisture probes across Victoria and Tasmania. Data typically includes soil moisture and soil temperature readings at depths 300-1000mm. [Explore](#)

 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. [Explore](#)

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



OpenAPI Search YAML Notebook swagger: ui: 3.22.3 [Gettable](#) [Discover PEB](#)

ale

fact_data/swagger.json

`interval` or `$number interval` allow range queries. e.g. `9`, translates to `attr >= 9`. `9,12` translates to `attr >= 9`. More than 1 comma translates to an `in` query. No comma is an equality query. is explained [here](#).

pret any string with a `%` character as an SQL "like" query.

ma separated list of column names prefixed or postfixed by `>` or `<` so that `col>` or `col<` is ascending and `col<` or `col>` is descending.

`!all data`.

`mask` allows column filtering and its format is explained [here](#).

n UTC+08:00.

[national License](#)

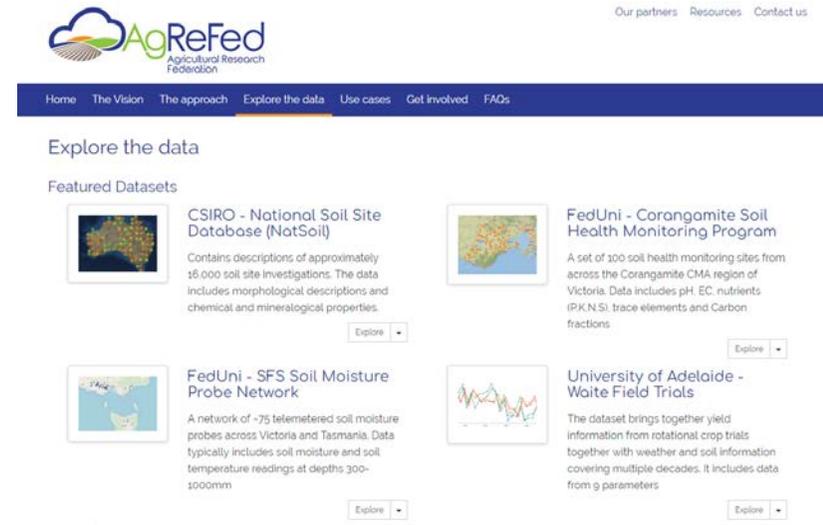
tables

table
Metabolite
DI



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 "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.

Our partners Resources Contact us

Home The Vision The approach Explore the data Use cases Get involved FAQs

Explore the data

Featured Datasets

 **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. [Explore](#)

 **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. [Explore](#)

 **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. [Explore](#)

 **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. [Explore](#)

Historical trial research

Waite Permanent Rotation Trial

Plot: Crop/Pasture: Rotation type: Observed property:

Start Year: End Year: Chart type:

- Permanent rotation trial since 1920
- Data collected by CSIRO into bespoke spreadsheet
- O&M-relational DB developed – FedUni and UofA

```
{
  id: 9999,
  year: "1925",
  plot_number: 1
}
```



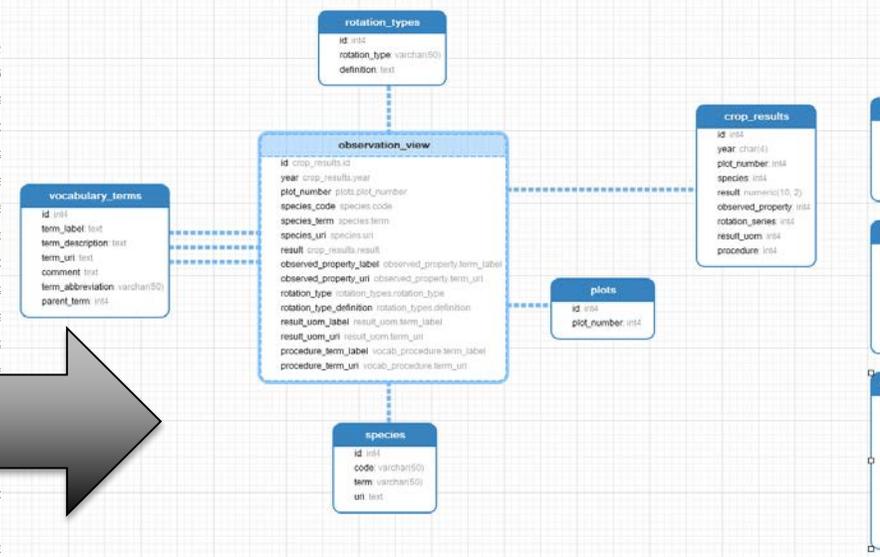
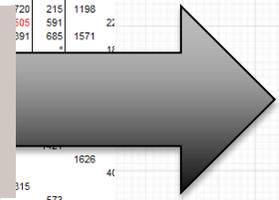
THE UNIVERSITY
of ADELAIDE

<http://id.agrisemantics.org/gi>
 "Grain Yield",
<http://id.agrisemantics.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	3	
1925	2464	2654	746	2681				2468	5383		2472	716		2252	5621		2504	2688	5982	800	2501	6045	800	3002	5717			549	5673		2888			
1926	2937	2965	1102	2664				4523	2688	3568	2675	1136	2580	4623	2641	2641	1520	4503	1122	3266	4321	1257	3024	4439	3152	2876			3649	5673	3011	3080	2i	
1927	2381	2044	1794	1450	1891			1913	4822	3568	2681	276	2755	3716	2018	276	3004	3642	276	2977	4146													
1928	2619	1399	2619	1399	3422			5063	3348	3500	3998	4666	3622	1824	3854	1740	2787	4111	1839	2987	3818	5268												
1929	3287	2848	2854	1266	3099			2948	4598	2673	2195	3500	2556	4240	1228	1505	2964	4351	1689	2959	4210	5421												
1930	1439	1667	1262	988	1547			1650	3111	1679	1799	1754	3434	872	1409	2097	1779	1486	2740	1991														
1931	2675	3076	1801	2285	2992			4075	1866	3426	2238	3902	2238	2238	856	2369	3112	1993	2592	2850	2366													
1932	2675	3076	1801	2285	2992			4075	1866	3426	2238	3902	2238	2238	856	2369	3112	1993	2592	2850	2366													
1933	2675	3076	1801	2285	2992			4075	1866	3426	2238	3902	2238	2238	856	2369	3112	1993	2592	2850	2366													
1934	2675	3076	1801	2285	2992			4075	1866	3426	2238	3902	2238	2238	856	2369	3112	1993	2592	2850	2366													
1935	1404	2523	1144	1152	3143			91	2736	1699	2775	3214	1693	2244	1950	97	1815	1764	681	1373	1674	931												
1936	1936	1898	1792	653	2510			2795	2611	2116	1944	1693	2244	1950	97	1815	1764	681	1373	1674	931													
1937	2122	2236	2046	1764	2510			2795	2611	2116	1944	1693	2244	1950	97	1815	1764	681	1373	1674	931													
1938	1898	2236	2046	1764	2510			2795	2611	2116	1944	1693	2244	1950	97	1815	1764	681	1373	1674	931													
1939	2216	2655	1460	2206	2655			2535	2033	2325	2730	2466	762	481	1493	2167	2562	1724	2230	3134														
1940	2366	2539	1326	1097	2539			2462	2317	2534	1280	2668	1708	675	2025	1696	1395	2091	1564	1325														
1941	2532	2813	2376	2611	2813			4074	2718	2852	2020	3924	3009	1428	3955	3115	2909	4052	3277	2977	2738	5273												
1942	2033	2332	2009	1594	2332			2521	2272	2206	1689	1976	1124	260	2861	1609	18	2833	2214	1247	224													
1943	3222	4123	2183	1856	4123			2521	2272	2206	1689	1976	1124	260	2861	1609	18	2833	2214	1247	224													
1944	1884	2561	1054	1855	2561			1785	2541	1438	2468	1679	2228	725	1114	1229	1978	1788	1476	2085														
1945	3024	3956	1636	1348	3956			371	no d																									
1946	2201	2545	978	1196	2545			1676	no d																									
1947	1072	1652	788	1572	1652			2908	24																									
1948	1197	2376	527	1630	2376			2942	22																									
1949	2168	3569	2134	3569	3569			2942	22																									
1950	2005	3355	1282	407	3355			579	no d																									
1951	210	666	210	745	666			181	no d																									
1952	925	1922	538	477	1922			1472	22																									
1953	760	847	786	67	847			181	no d																									
1954	1305	2374	1230	1243	2374			396	16																									
1955	592	968	538	477	968			974	21																									
1956	547	1215	464	852	1215			974	21																									
1957	1561	2656	1277	638	2656			3011	19																									
1958	1976	3616	2103	1841	3616			3011	19																									
1959	253	414	242	34	414			358	464																									
1960	1566	2380	811	1425	2380			1808	3246																									
1961	424	753	985	611	753			900	605	2748																								
1962	1990	3293	1962	1599	3293			3248	766	895	2200																							
1963	425	784	1702	255	784			786	895	2200																								

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"
```



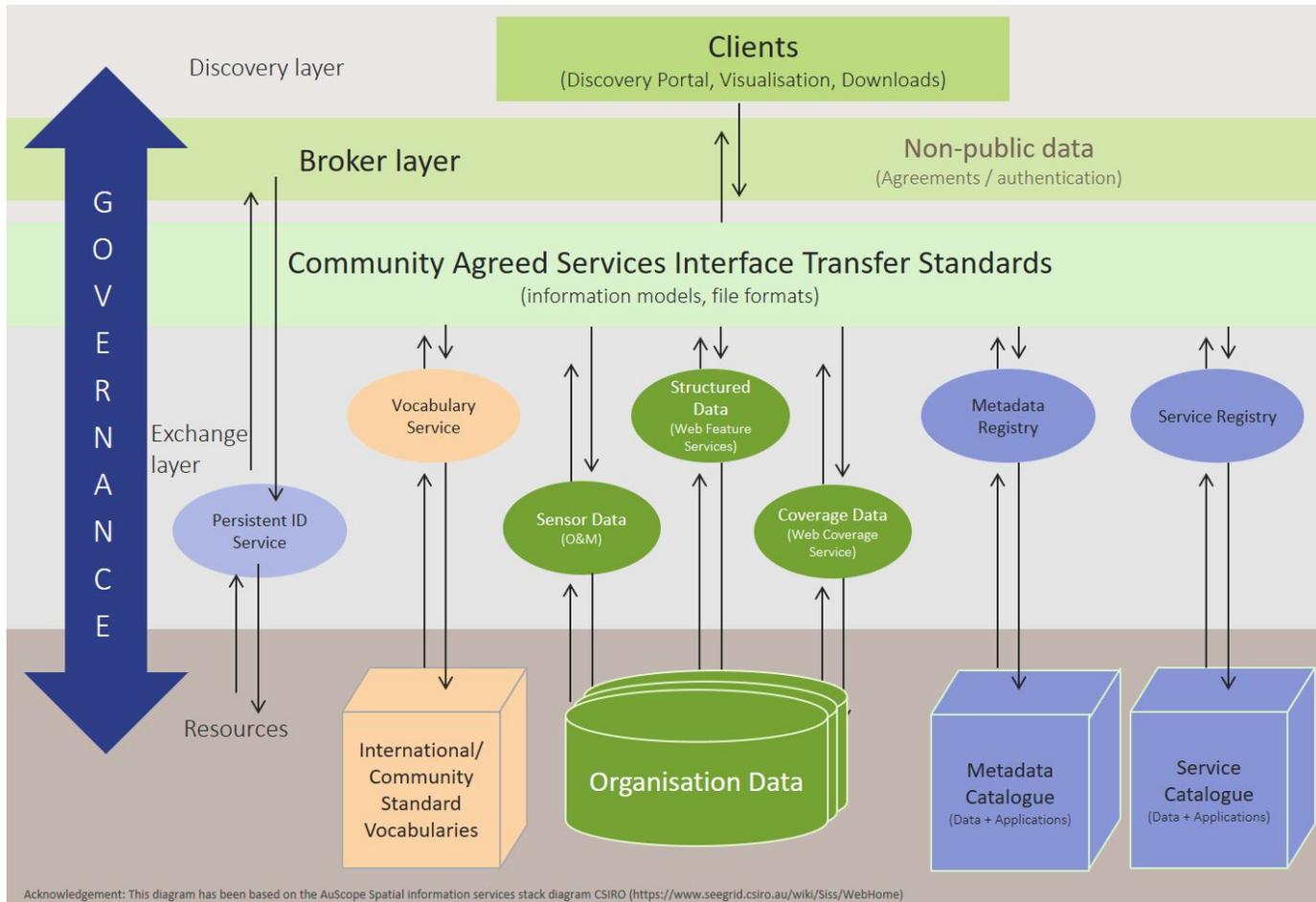
From 'my data' to 'our FAIR data'

- 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



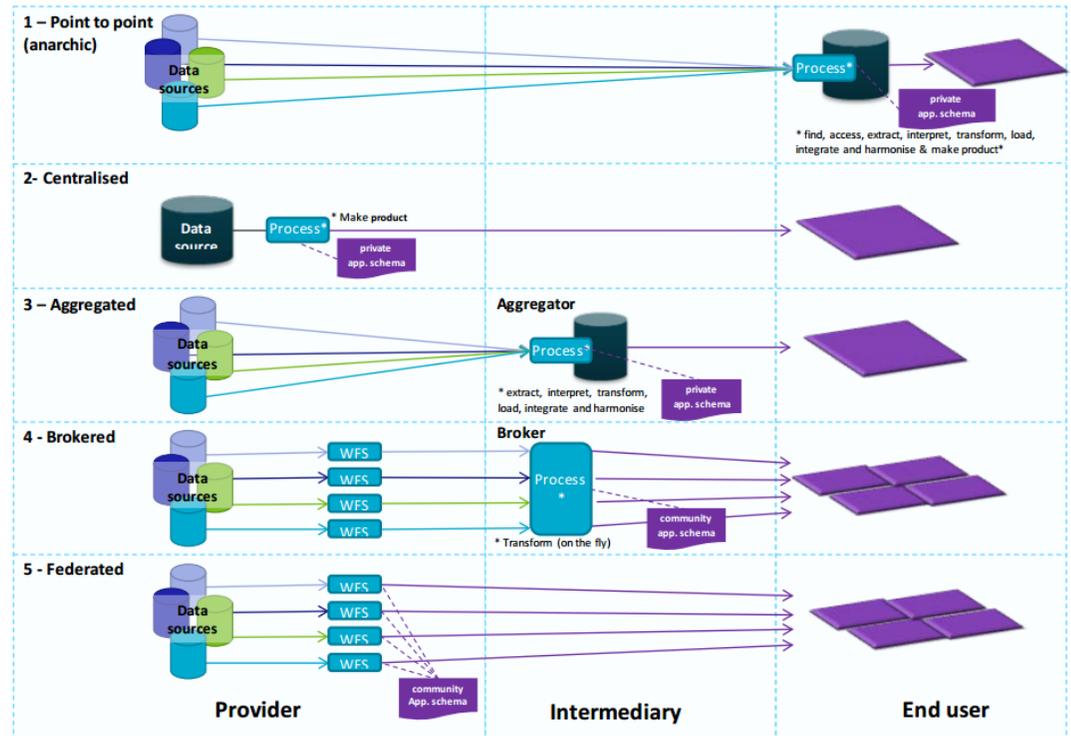
This diagram provides a high level view of the interactions between provider and users via provider and common computational components.

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.

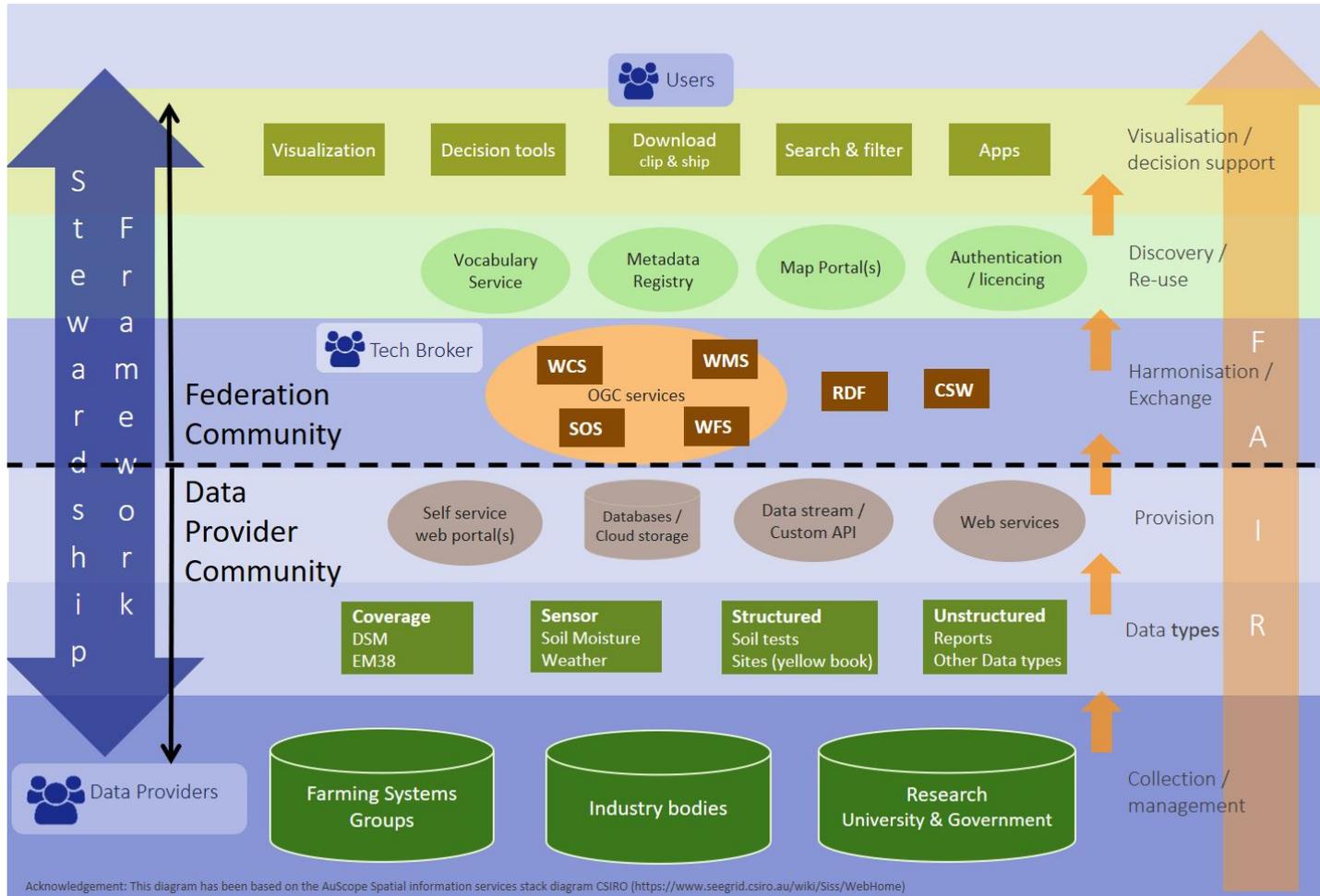
- For federated and brokered data, source data resides with the data provider system. Both solutions encourage currency and validity of data.
- A standards based Service Oriented Architecture (https://en.wikipedia.org/wiki/Service-oriented_architecture) is utilised, including metadata cataloguing and 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)

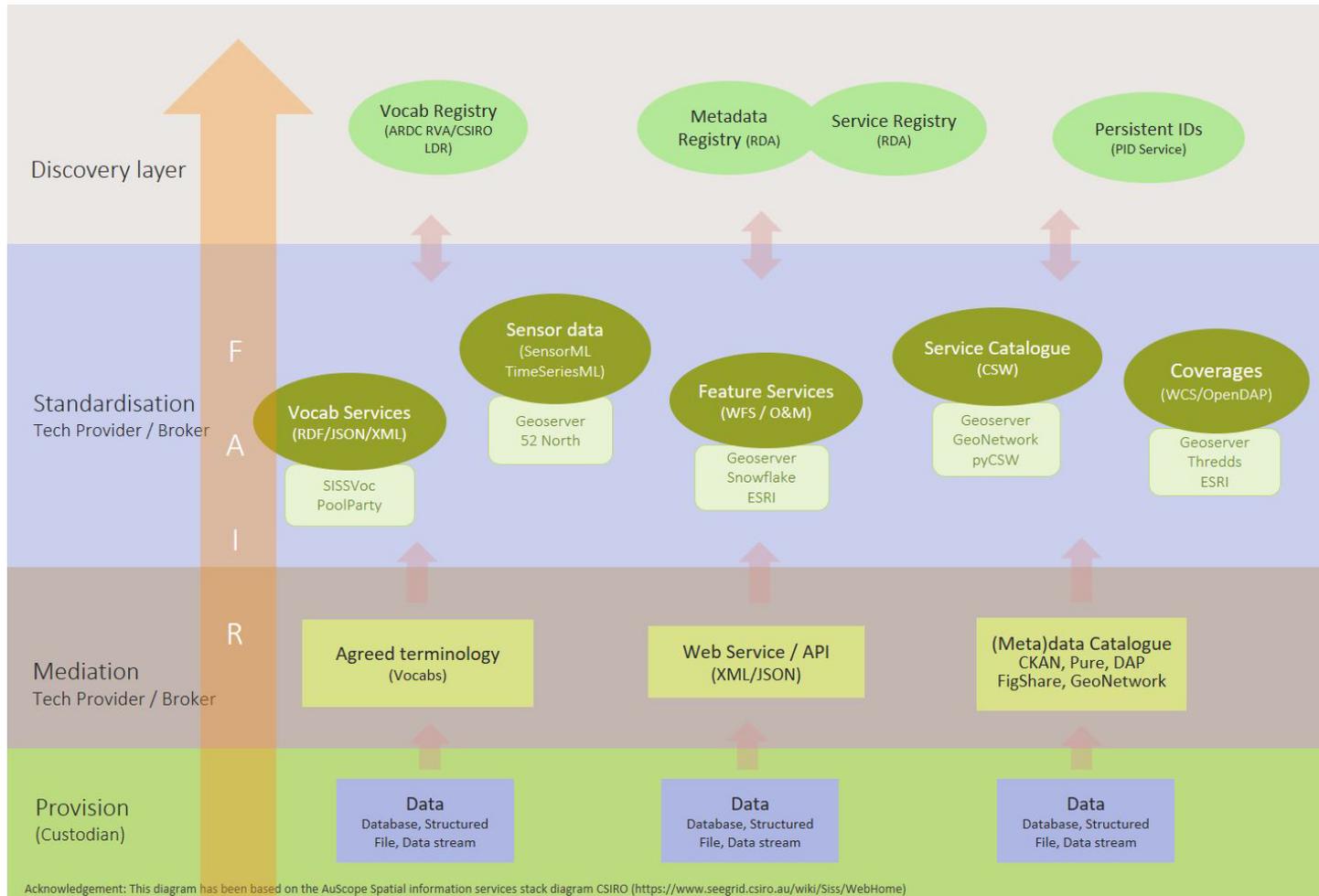
1 <https://publications.csiro.au/rpr/pub?pid=csiro:EP155525>

AgReFed Engineering viewpoint



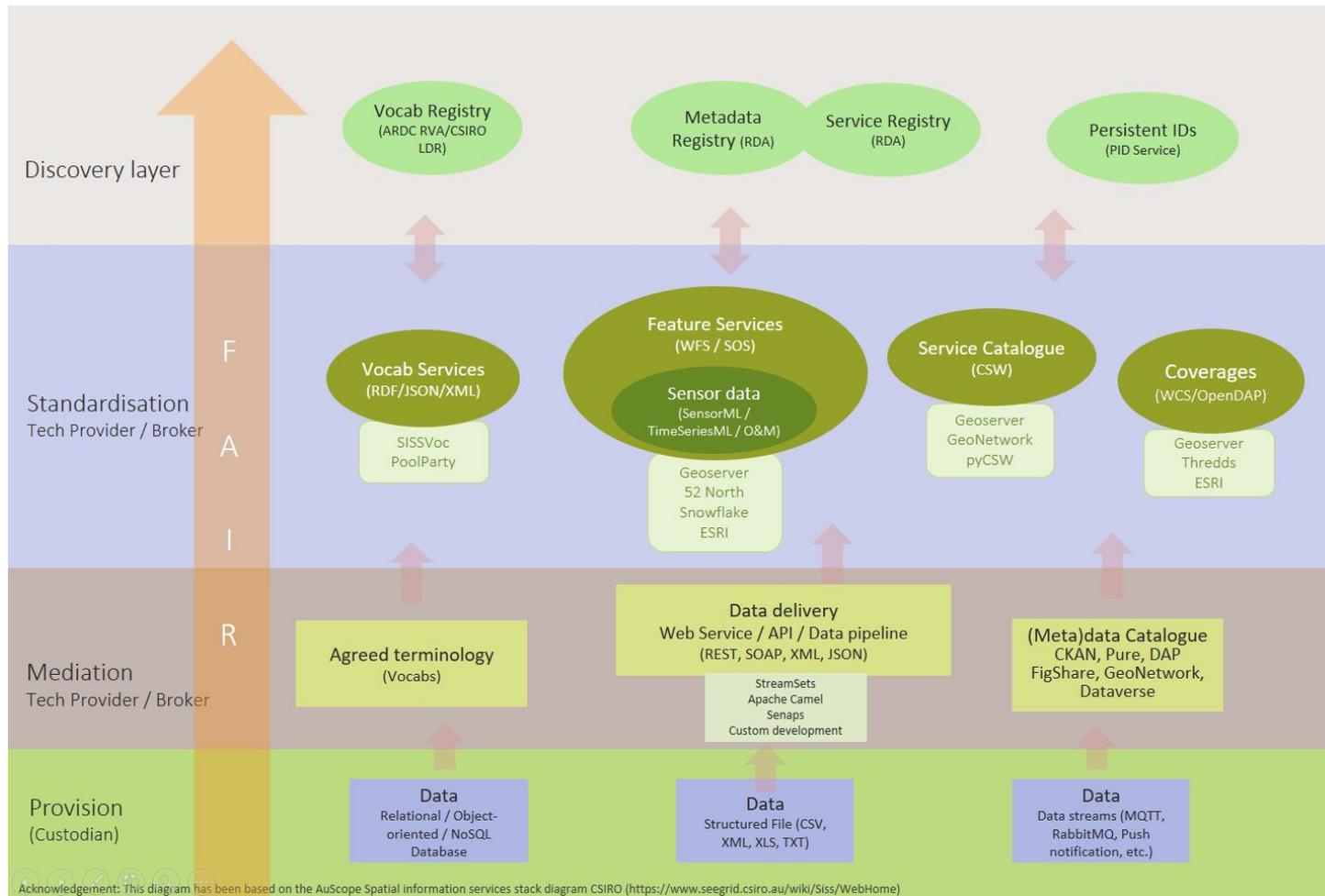
This diagram provides a view of the components, mechanisms and functions required to support distributed interactions among objects in the system, including whose responsibility they are. When this diagram is fleshed out in more detail it will show where components are deployed.

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.