Publishing Linked Data

Jeremy Tandy

November 2014

Why do we care about Linked Data?

2.41 The Commission believes that adopting Open Data principles and removing barriers to re-use can generate economic benefits in the order of £33 billion a year.

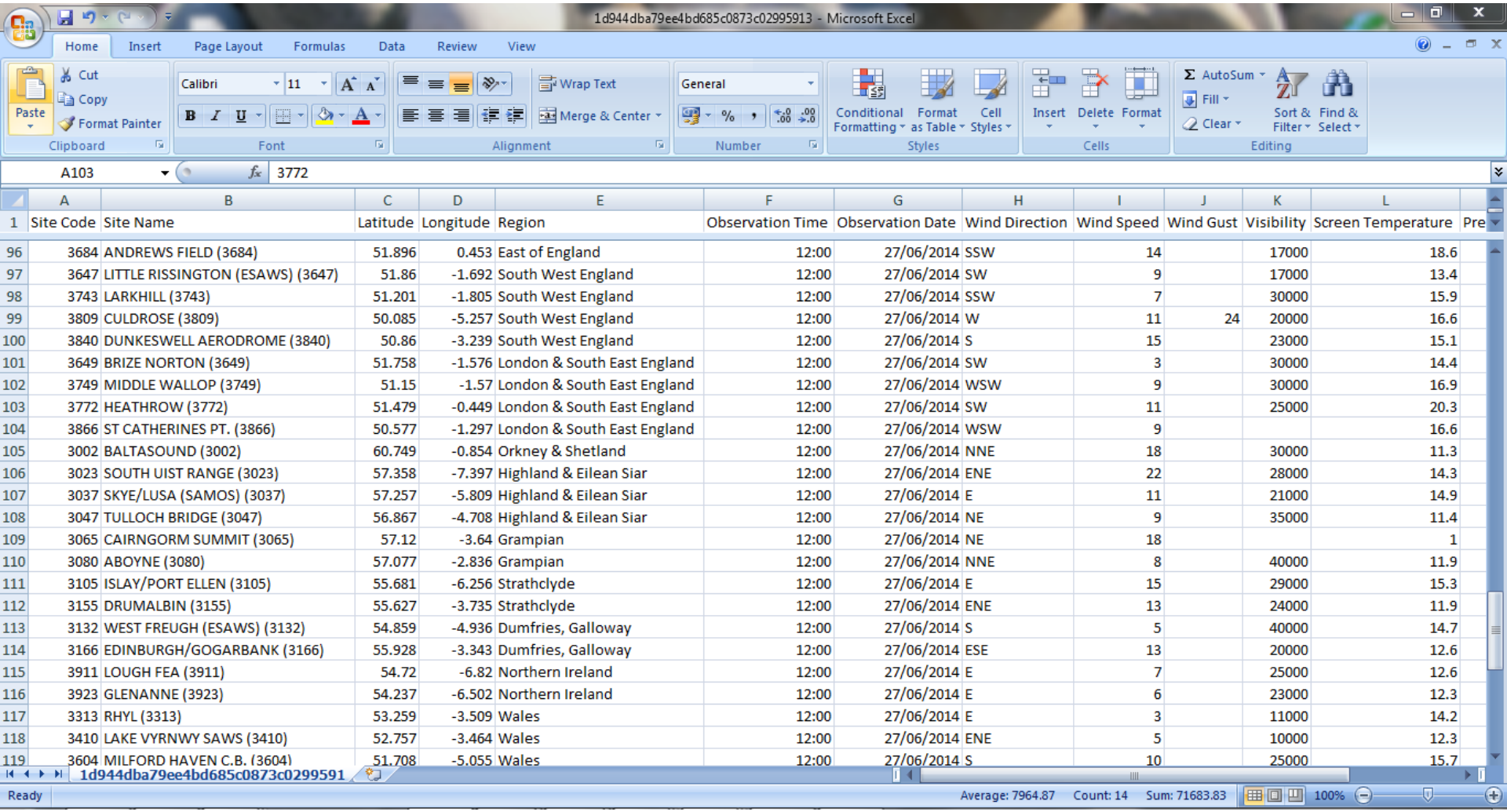
(9) Release data quickly, and then work to make sure that it is available in open standard formats, including **linked data** forms .

2.51 [...] there is a growing realisation of the power of linked data for exposing, sharing and connecting pieces of data and information [...] to realise efficiencies in the public sector.

Disclaimer

Met Office URIs used in this presentation are illustrative and will not resolve.

Does your data look like this?



1d944dba79ee4bd685c0873c02995913 - Microsoft Excel

Home | Insert | Page Layout | Formulas | Data | Review | View

Cut | Copy | Format Painter | Paste | Clipboard

Calibri | 11 | Font

Wrap Text | Merge & Center | Alignment

General | Number

Conditional Formatting | Format as Table | Cell Styles | Styles

Insert | Delete | Format | Cells

AutoSum | Fill | Clear | Sort & Filter | Find & Select | Editing

A103 | 3772

	A	B	C	D	E	F	G	H	I	J	K	L	
1	Site Code	Site Name	Latitude	Longitude	Region	Observation Time	Observation Date	Wind Direction	Wind Speed	Wind Gust	Visibility	Screen Temperature	Pre
96	3684	ANDREWS FIELD (3684)	51.896	0.453	East of England	12:00	27/06/2014	SSW	14		17000	18.6	
97	3647	LITTLE RISSINGTON (ESAWS) (3647)	51.86	-1.692	South West England	12:00	27/06/2014	SW	9		17000	13.4	
98	3743	LARKHILL (3743)	51.201	-1.805	South West England	12:00	27/06/2014	SSW	7		30000	15.9	
99	3809	CULDROSE (3809)	50.085	-5.257	South West England	12:00	27/06/2014	W	11	24	20000	16.6	
100	3840	DUNKESWELL AERODROME (3840)	50.86	-3.239	South West England	12:00	27/06/2014	S	15		23000	15.1	
101	3649	BRIZE NORTON (3649)	51.758	-1.576	London & South East England	12:00	27/06/2014	SW	3		30000	14.4	
102	3749	MIDDLE WALLOP (3749)	51.15	-1.57	London & South East England	12:00	27/06/2014	WSW	9		30000	16.9	
103	3772	HEATHROW (3772)	51.479	-0.449	London & South East England	12:00	27/06/2014	SW	11		25000	20.3	
104	3866	ST CATHERINES PT. (3866)	50.577	-1.297	London & South East England	12:00	27/06/2014	WSW	9			16.6	
105	3002	BALTASOUND (3002)	60.749	-0.854	Orkney & Shetland	12:00	27/06/2014	NNE	18		30000	11.3	
106	3023	SOUTH UIST RANGE (3023)	57.358	-7.397	Highland & Eilean Siar	12:00	27/06/2014	ENE	22		28000	14.3	
107	3037	SKYE/LUSA (SAMOS) (3037)	57.257	-5.809	Highland & Eilean Siar	12:00	27/06/2014	E	11		21000	14.9	
108	3047	TULLOCH BRIDGE (3047)	56.867	-4.708	Highland & Eilean Siar	12:00	27/06/2014	NE	9		35000	11.4	
109	3065	CAIRNGORM SUMMIT (3065)	57.12	-3.64	Grampian	12:00	27/06/2014	NE	18			1	
110	3080	ABOYNE (3080)	57.077	-2.836	Grampian	12:00	27/06/2014	NNE	8		40000	11.9	
111	3105	ISLAY/PORT ELLEN (3105)	55.681	-6.256	Strathclyde	12:00	27/06/2014	E	15		29000	15.3	
112	3155	DRUMALBIN (3155)	55.627	-3.735	Strathclyde	12:00	27/06/2014	ENE	13		24000	11.9	
113	3132	WEST FREUGH (ESAWS) (3132)	54.859	-4.936	Dumfries, Galloway	12:00	27/06/2014	S	5		40000	14.7	
114	3166	EDINBURGH/GOGARBANK (3166)	55.928	-3.343	Dumfries, Galloway	12:00	27/06/2014	ESE	13		20000	12.6	
115	3911	LOUGH FEA (3911)	54.72	-6.82	Northern Ireland	12:00	27/06/2014	E	7		25000	12.6	
116	3923	GLENANNE (3923)	54.237	-6.502	Northern Ireland	12:00	27/06/2014	E	6		23000	12.3	
117	3313	RHYL (3313)	53.259	-3.509	Wales	12:00	27/06/2014	E	3		11000	14.2	
118	3410	LAKE VYRNWY SAWS (3410)	52.757	-3.464	Wales	12:00	27/06/2014	ENE	5		10000	12.3	
119	3604	MILFORD HAVEN C.B. (3604)	51.708	-5.055	Wales	12:00	27/06/2014	S	10		25000	15.7	

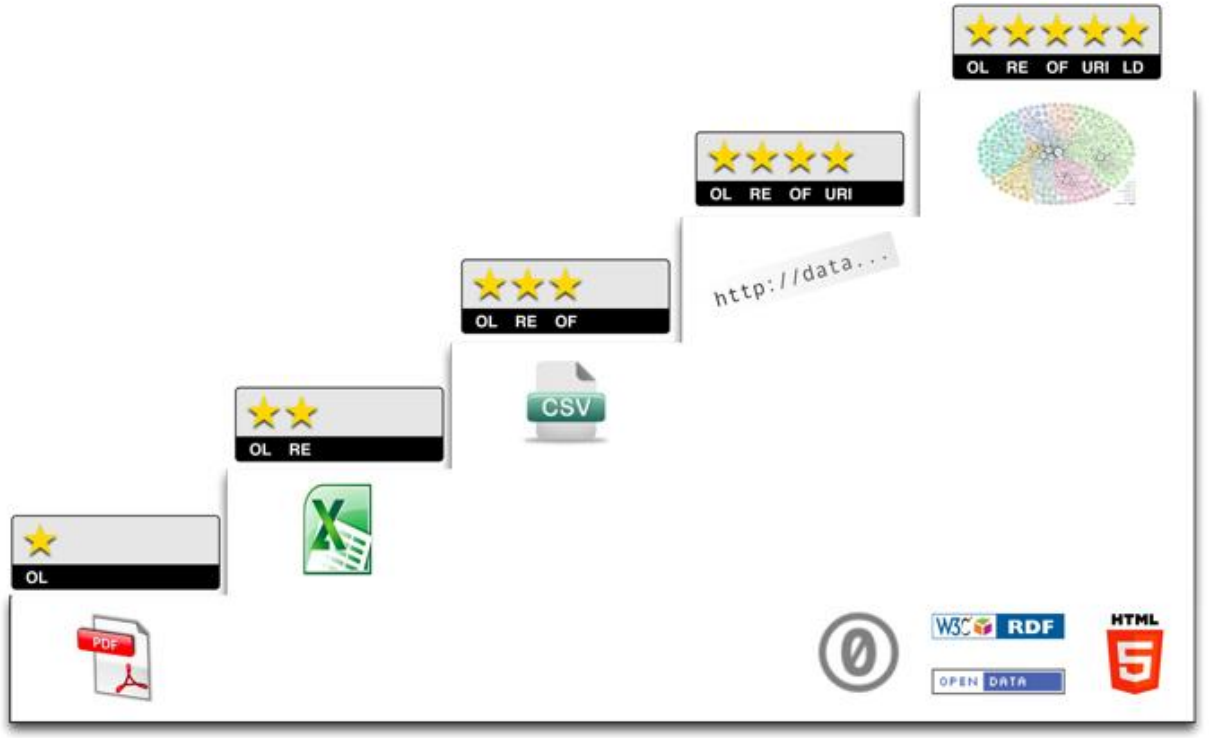
1d944dba79ee4bd685c0873c0299591

Ready | Average: 7964.87 | Count: 14 | Sum: 71683.83 | 100%

Met Office

source: http://data.gov.uk/metoffice-data-archive
(all observing sites, 2014-06-27T12:00)

5★ (Linked) Open Data



on the web published with an open license* ✓

structured data ✓

non-proprietary format ✓

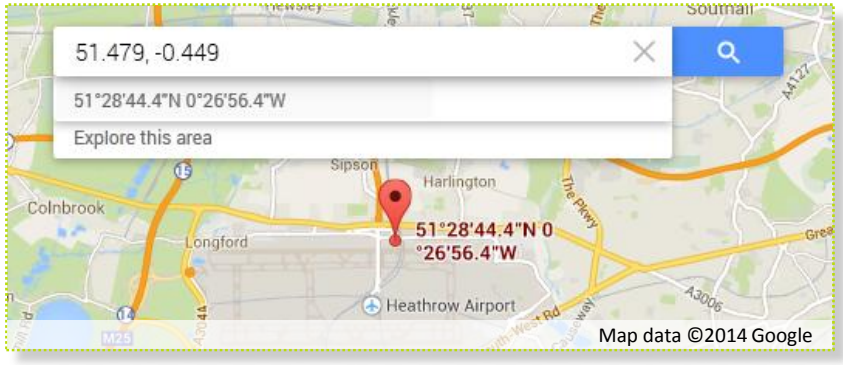
Met Office

source: http://5stardata.info/

4★: use URIs to denote things

	A	B	C	D	E	F	G	H	I	J	K	L	
1	Site Code	Site Name	Latitude	Longitude	Region	Observation Time	Observation Date	Wind Direction	Wind Speed	Wind Gust	Visibility	Screen Temperature	Pre
103	3772	HEATHROW (3772)	51.479	-0.449	London & South East England	12:00	27/06/2014	SW	11		25000	20.3	


```
{
    "Site Code": "3772",
    "Site Name": "HEATHROW (3772)",
    "Latitude": "51.479",
    "Longitude": "-0.449",
    "Region": "London & South East England",
    "Observation Time": "12:00",
    "Observation Date": "27/06/2014",
    "Wind Direction": "SW",
    "Wind Speed": "11",
    "Wind Gust": "",
    "Visibility": "25000",
    "Screen Temperature": "20.3",
    "Pressure": "1012",
    "Significant Weather": "Partly cloudy (day)"
}
```

(16-point compass)
(mph)
(mph)
(m)
(degrees Celsius)
(hPa)
(code)

Referencing code list terms

Canonical labels don’t work well as identifiers as their use is error prone …

```
{
   …
   "Significant Weather": "Partly cloudy (day)”     (code)
}
```

Referencing code list terms

Canonical labels don't work well as identifiers as their use is error prone …

So use an unambiguous global unique identifier to refer to the term … a URI

```
{
    …
    "Significant Weather": {"@id": "http://data.metoffice.com/datapoint/def/weather-type/3"}
}
```

Referencing code list terms

```
{
   "@id": "http://data.metoffice.com/datapoint/def/weather-type/3",
   "notation": "3",
   "label": [
      {"@language": "en", "@value": "Partly cloudy (day)"},
      {"@language": "cy", "@value": "Rhannol gymylog (dydd)”}
   ]
}
```

A URI allows the definition can be reconciled with usage …

A URI doesn’t have to resolve on the web – but it helps; a URI is just as useful as an identifier within a closed system.

```
{
   …
   "Significant Weather": {"@id": "http://data.metoffice.com/datapoint/def/weather-type/3"}
}
```

CHALLENGE: How do I get my URIs to resolve?

URIs don't have to resolve – but Linked Data is predicated on the idea that they *do* resolve to provide useful info.

The trouble is: URIs are hard.

Publishers need to have control over some suitable web domain and the ability to publish data there. Yet they:

- frequently lack the authority to manage parts of their organisation's namespace, and
- may be required to submit pages for publication via a content management system that neither allows control over the resulting URIs nor supports content negotiation required for data publication.

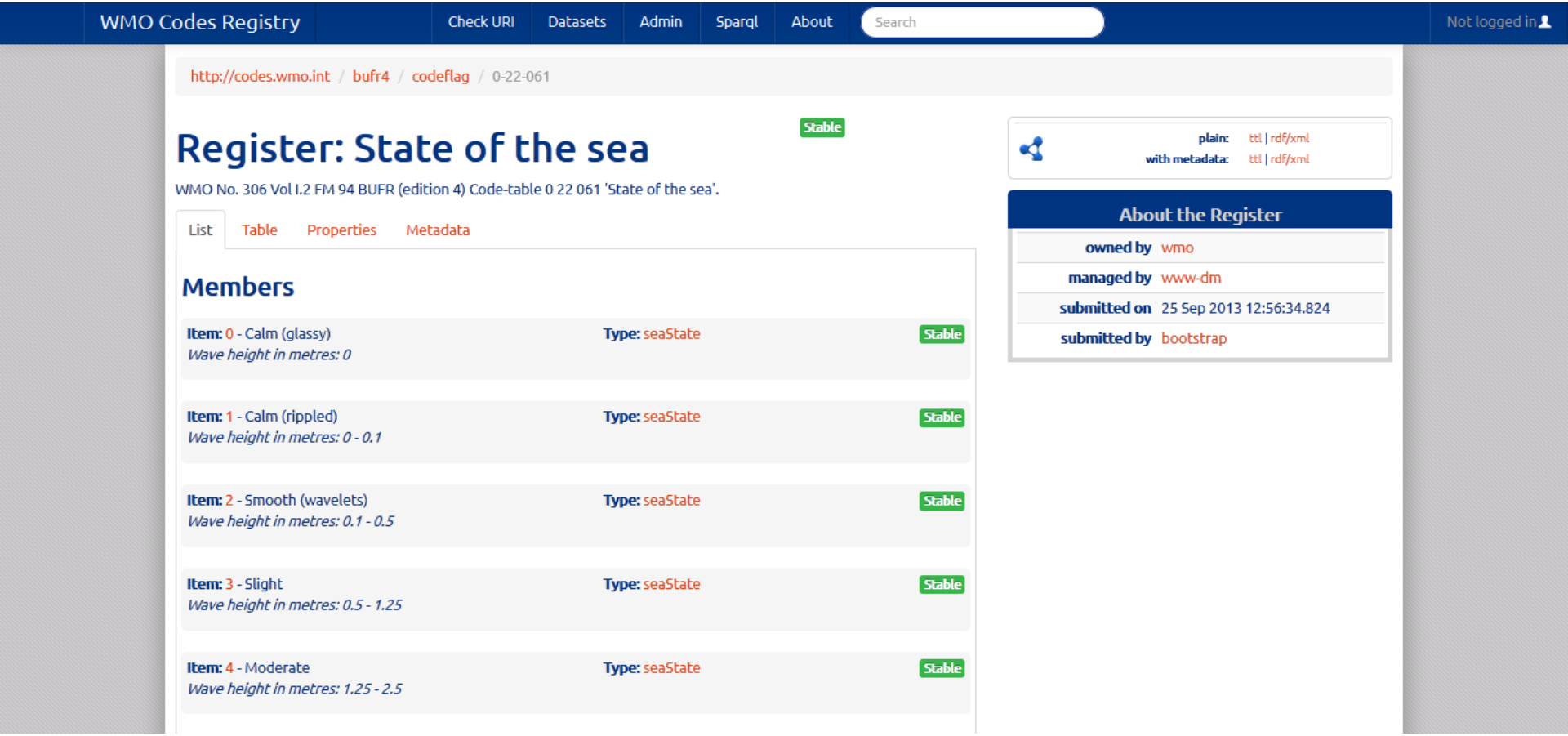
UKGovLD Registry software

Funded by the UK Government, UKGovLD and partners Epimorphics, have developed an open source software solution: the UKGovLD Registry.

- A register is a controlled list; governance is explicit
- Provides authoritative point of reference
- Container pattern; hackable URIs allow enable traversal from member to containing register …

  http://data.metoffice.com/datapoint/def/weather-type~~/3~~
- Validation service: check that a term is valid member
- Simple Linked Data publication platform
- Not limited to SKOS Collections or Concept Schemes

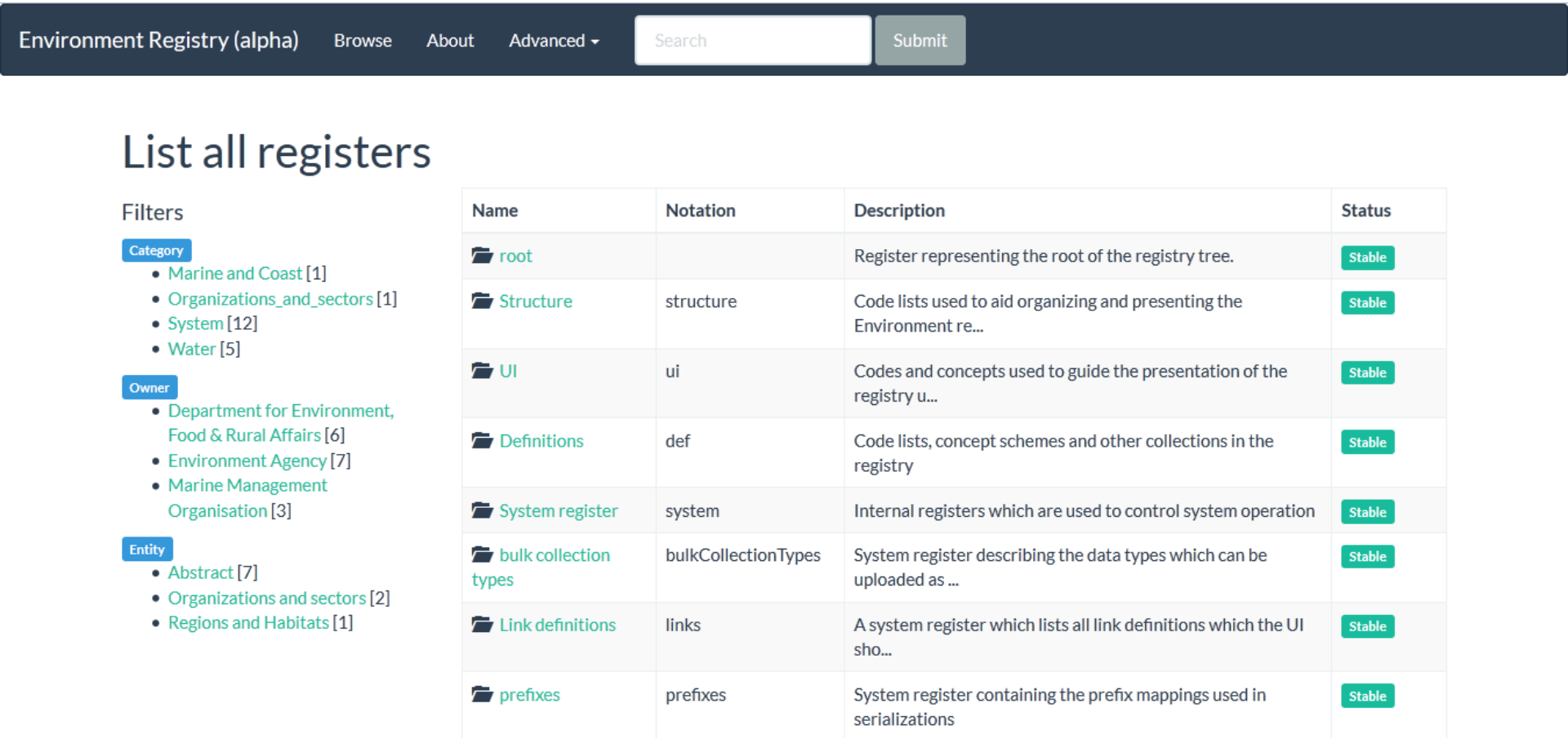
Deployment example: WMO Codes



Met Office

http://codes.wmo.int/

Deployment example: Environment Registry



Met Office

http://environment.data.gov.uk/registry/

4★: use URIs to denote code-list terms

```
{
  "@context": {"skos": "http://www.w3.org/2004/02/skos/core#"},
  "Site Code": "3772",
  "Site Name": "HEATHROW (3772)",
  "Latitude": "51.479",
  "Longitude": "-0.449",
  "Region": "London & South East England",
  "Observation Time": "12:00",
  "Observation Date": "27/06/2014",
  "Wind Direction": {
    "@id": "http://location.data.gov.uk/def/direction/compass-point/SW",
    "skos:notation": "SW“ },
  "Wind Speed": "11",
  "Wind Gust": "",
  "Visibility": "25000",
  "Screen Temperature": "20.3",
  "Pressure": "1012”,
  "Significant Weather": {
    "@id": "http://data.metoffice.com/datapoint/def/weather-type/3",
    "skos:prefLabel": "Partly cloudy (day)“ }
}
```

CHALLENGE: How do determine the entities in my data?


```
{
  "@context": {"skos": "http://www.w3.org/20(
  "Site Code": "3772",
  "Site Name": "HEATHROW (3772)",
  "Latitude": "51.479",
  "Longitude": "-0.449",
  "Region": "London & South East England",
  "Observation Time": "12:00",
  "Observation Date": "27/06/2014",
  "Wind Direction": {
    "@id": "http://location.data.gov.uk/def/direction/compass-point/SW",
    "skos:notation": "SW“ },
  "Wind Speed": "11",
  "Wind Gust": "",
  "Visibility": "25000",
  "Screen Temperature": "20.3",
  "Pressure": "1012”,
  "Significant Weather": {
    "@id": "http://data.metoffice.com/datapoint/def/weather-type/3",
    "skos:prefLabel": "Partly cloudy (day)“ }
}
```


What's the subject? There are two entities described in each row …

- site (location); and
- observation event

4★: use URIs to denote resources

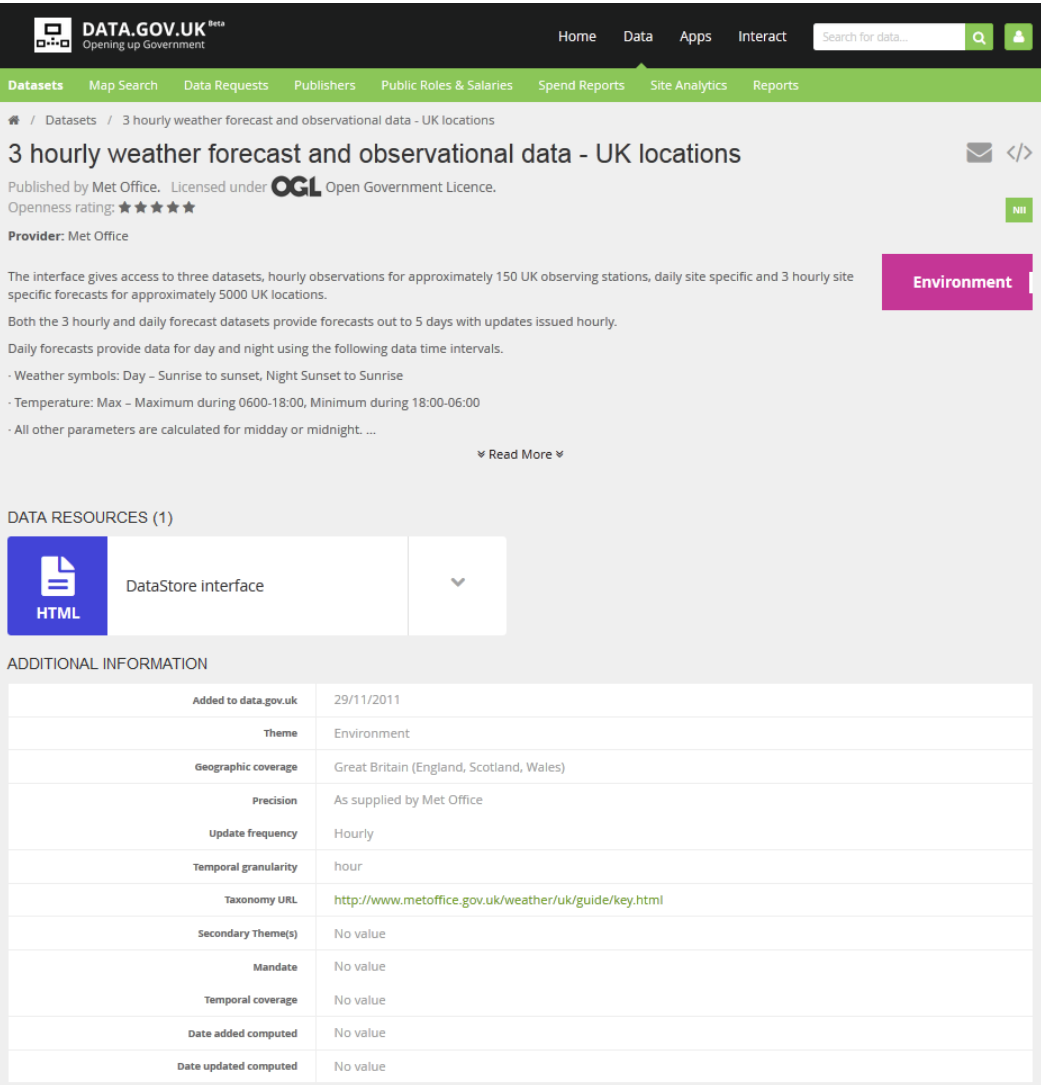

```
{
  "@context": {"skos": "http://www.w3.org/2004/02/skos/core#"},
  "site": {
    "@id": "http://data.metoffice.com/uk/locations/obs/site/3772#id",
    "Site Code": "3772",
    "Site Name": "HEATHROW (3772)",
    "Latitude": "51.479",
    "Longitude": "-0.449",
    "Region": "London & South East England"
  },
  "@id": "http://data.metoffice.com/uk/data/weather-observations/site/3772/date-time/20140627T1200Z",
  "Observation Time": "12:00",
  "Observation Date": "27/06/2014",
  "Wind Direction": {
    "@id": "http://location.data.gov.uk/def/direction/compass-point/SW",
    "skos:notation": "SW"
  },
  "Wind Speed": "11",
  "Wind Gust": "",
  "Visibility": "25000",
...
```

CHALLENGE: How do I get my URIs to resolve?


```
{
  "@context": {"skos": "http://www.w3.org/2004/02/skos/core#"},
  "site": {
    "@id": "http://data.metoffice.com/uk/locations/obs/site/3772#id",
    "Site Code": "3772",
    "Site Name": "HEATHROW (3772)",
    "Latitude": "51.479",
    "Longitude": "-0.449",
    "Region": "London & South East England"
  },
  "@id": "http://data.metoffice.com/uk/data/weather-observations/site/3772/date-time/20140627T1200Z",
  "Observation Time": "12:00",
  "Observation Date": "27/06/2014",
  "Wind Direction": {
    "@id": "http://location.data.gov.uk/def/direction/compass-point/SW",
    "skos:notation": "SW"
  },
  "Wind Speed": "11",
  "Wind Gust": "",
  "Visibility": "25000",
...
```


¡Click!

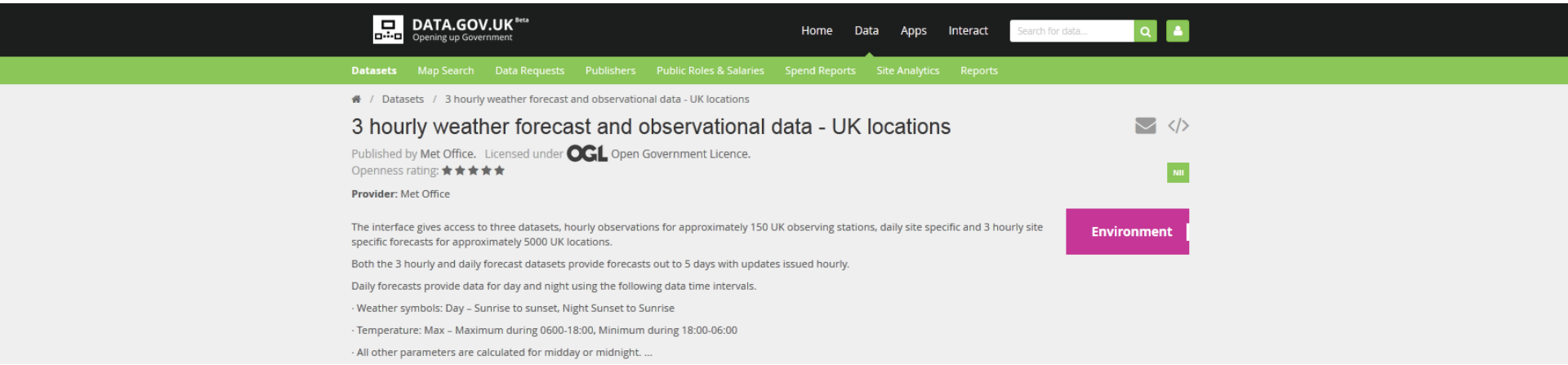
Resolve to a dataset description page?



Met Office

http://data.gov.uk/dataset/metoffice_uklocs3hr_fc

Resolve to a dataset description page?

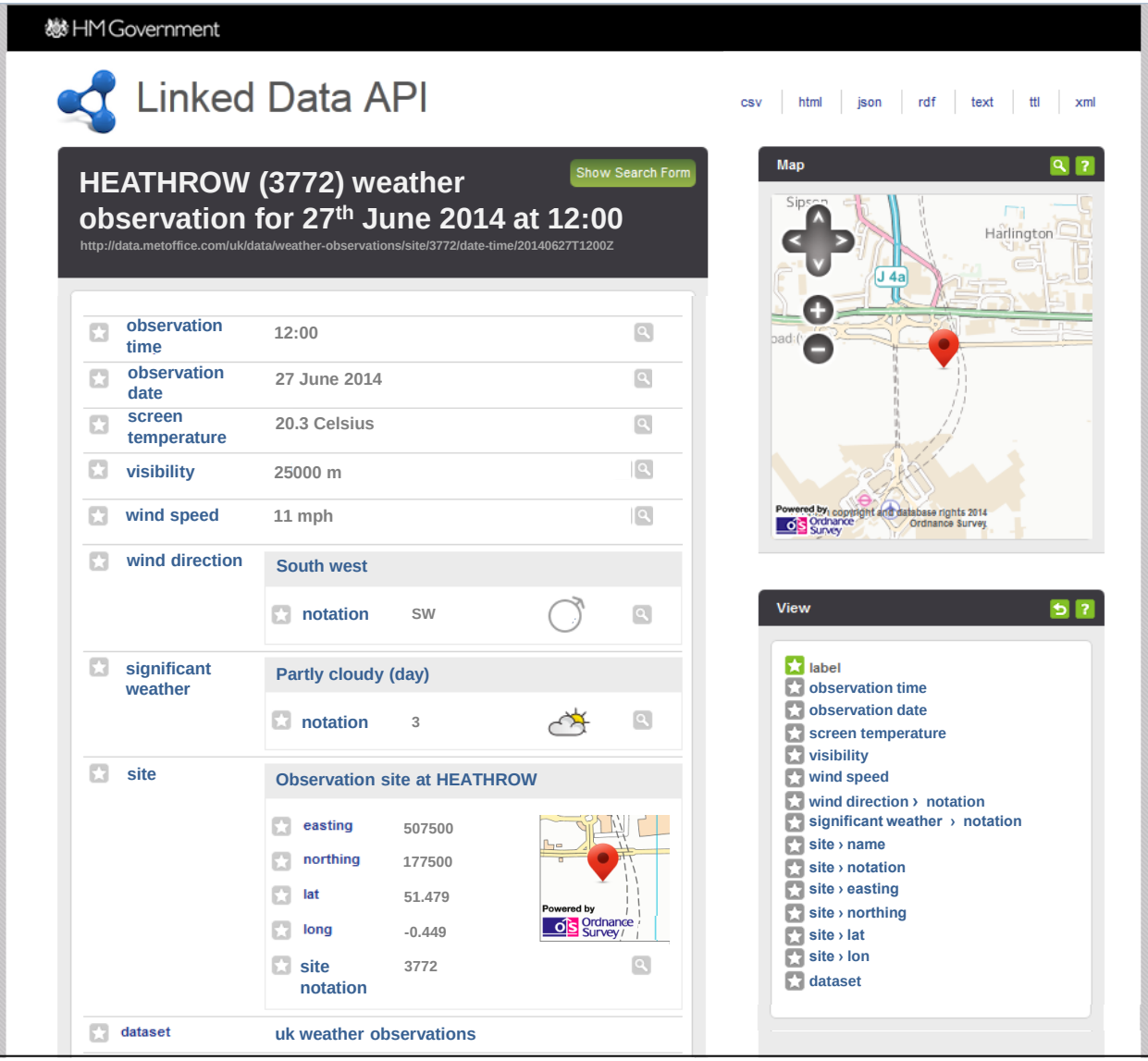


In fact, you should *always* publish a data description page to allow folks to discover and evaluate your data.

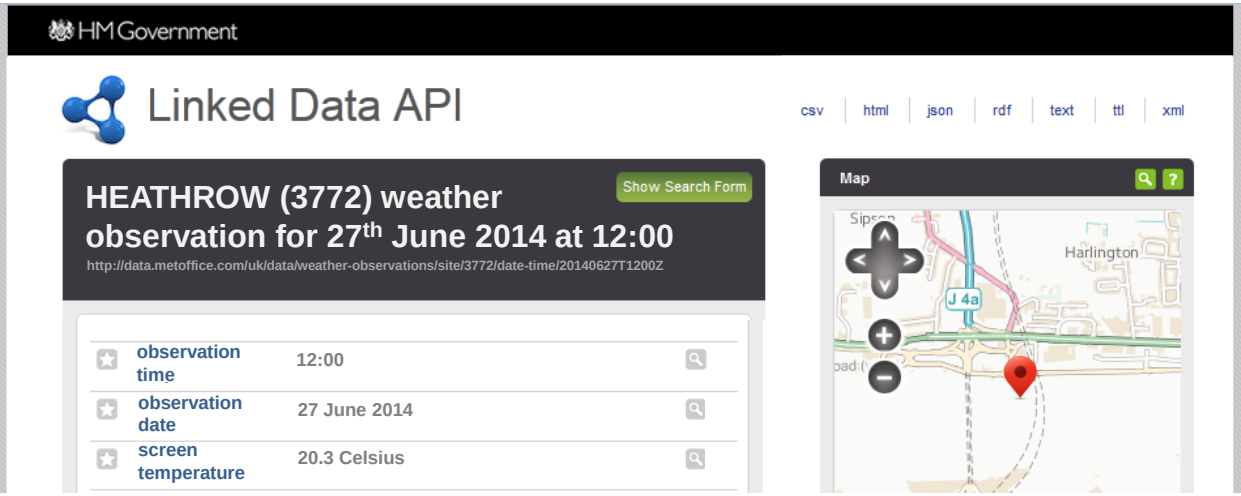
Publishing metadata descriptions for datasets should be familiar to those of you already publishing INSPIRE datasets.

http://data.gov.uk/dataset/metoffice_uklocs3hr_fc

Resolve each data item …



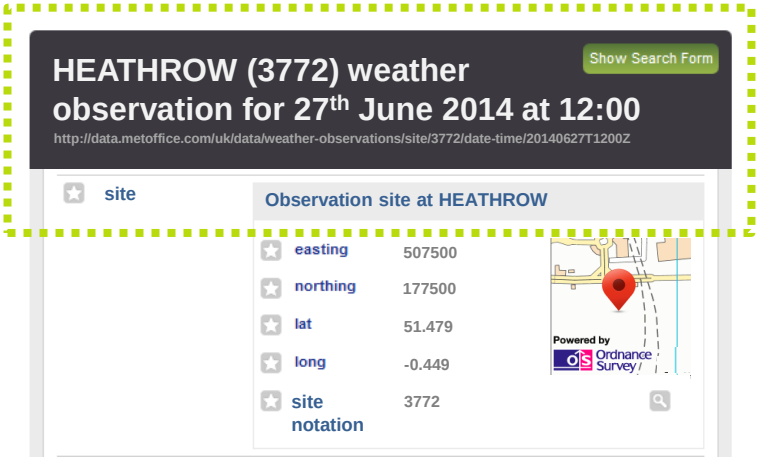
CHALLENGE: How do I deploy and configure item-level publication tools?



In addition to providing a dataset description page, one should publish information about each entity within the dataset.

Whilst the UKGovLD Registry provides a *simple* Linked Data publication platform, tools like ELDA are far more sophisticated. Yet with that sophistication comes the burden of complexity.

RDF triples



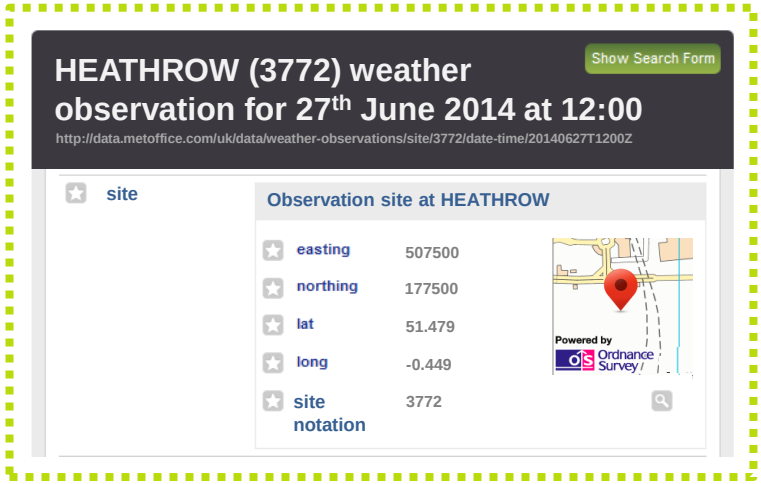
Linked Data is almost always published in RDF (or something that can be *transformed* to RDF).

Each RDF statement is a triple:

subject-predicate-object.

```
 {
S  "@id": "http://data.metoffice.com/uk/data/weather-observations/site/3772/date-time/20140627T1200Z",
P  "site": {
O    "@id": "http://data.metoffice.com/uk/locations/obs/site/3772#id"
     }
 }
```

RDF triples and graphs



Triples can be grouped together into graphs to convey complex information.

It is this ability that provides the *essence* of Linked Data.

```
{
  "@id": "http://data.metoffice.com/uk/data/weather-observations/site/3772/date-time/20140627T1200Z",
  "site": {
    "@id": "http://data.metoffice.com/uk/locations/obs/site/3772#id",
    "notation": "3772",
    "label": "HEATHROW (3772)",
    "lat": "51.479",
    "lon": "-0.449"
    }
}
```

… even linking data across organisation boundaries in 5★ data.

Predicates: they are URIs too

Predicates provide the semantics for the links between *subject* and *object*. To ensure that those semantics are unambiguous, each predicate is identified with a URI.

```
{
  "@context": {
  "ex" : "http://data.metoffice.com/datapoint/def/observation-sites#",
  "rdfs" : "http://www.w3.org/2000/01/rdf-schema#",
  "skos": "http://www.w3.org/2004/02/skos/core#",
  "geo": "http://www.w3.org/2003/01/geo/wgs_84#"
  },
  "@id": "http://data.metoffice.com/uk/data/weather-observations/site/3772/date-time/20140627T1200Z",
  "ex:site": {
    "@id": "http://data.metoffice.com/uk/locations/obs/site/3772#id",
    "skos:notation": "3772",
    "rdfs:label": "HEATHROW (3772)",
    "geo:lat": "51.479",
    "geo:lon": "-0.449"
    }
}
```

… often, we can reuse properties defined elsewhere – but you can't avoid creating your own; ex:site.

RDF Vocabularies: data models by another name

A vocabulary (or ontology) is the collection of classes and properties needed to describe your data. To create a vocabulary, you need to do a bit of data modelling.

Data modelling is the same process undertaken to create Application Schema – as defined in the ISO 19100-suite of Geographic Information standards.

ISO/DIS 19150-2 Geographic information – Ontology – Part 2: Rules for developing ontologies in the Web Ontology Language (OWL) provides a mechanism to automatically convert your Application Schema to an RDF vocabulary.

5★: link to other people's data

By linking to other people's data, you increase the size of the *global graph*; it allows data consumers to easily reconcile your data with that from other publishers.

- By referencing (authoritative) definitions of common terminology, consumers comparing your data with that from other publishers can be sure that they are comparing like with like.
- By referencing definitions of places & regions, consumers can easily find information from multiple publishers relating to their place of interest.

5★: link to other people's data


```
{  "@id": "http://data.metoffice.com/uk/data/weather-observations/site/3772/date-time/20140627T1200Z",
   "ssn:observationSamplingTime": {"time:inXSDDateTime": "2014-06-27T12:00:00Z"},
   "ssn:featureOfInterest": {
      "@id": "http://data.metoffice.com/uk/locations/obs/site/3772#id",
      "skos:notation": "3772",
      "rdfs:label": "HEATHROW (3772)",
      "sam:shape": {"geo:asWKT": "Point(-0.449 51.479)"},
      "sam:sampledFeature": {
         "@id": "http://data.ordnancesurvey.co.uk/id/50kGazetteer/112605",
         "rdfs:label": "Heathrow London Airport"}},
   "ssn:observationResult": {
      "ssn:hasOutput": {
         "qudt:numericValue": "20.3",
         "qudt:unit": {
            "@id": "http://qudt.org/vocab/unit#DegreeCelsius",
            "rdfs:label": "Celsius"},
         "qudt:valueQuantity": {
            "qudt:quantityKind": {
               "@id": "http://codes.wmo.int/common/c-15/me/airTemperature",
               "rdfs:label": "Air temperature"}}}
      ...
}}
```


¡Click!

Met Office

Example amended to use Semantic Sensor Network (SSN) ontology and other well known vocabs

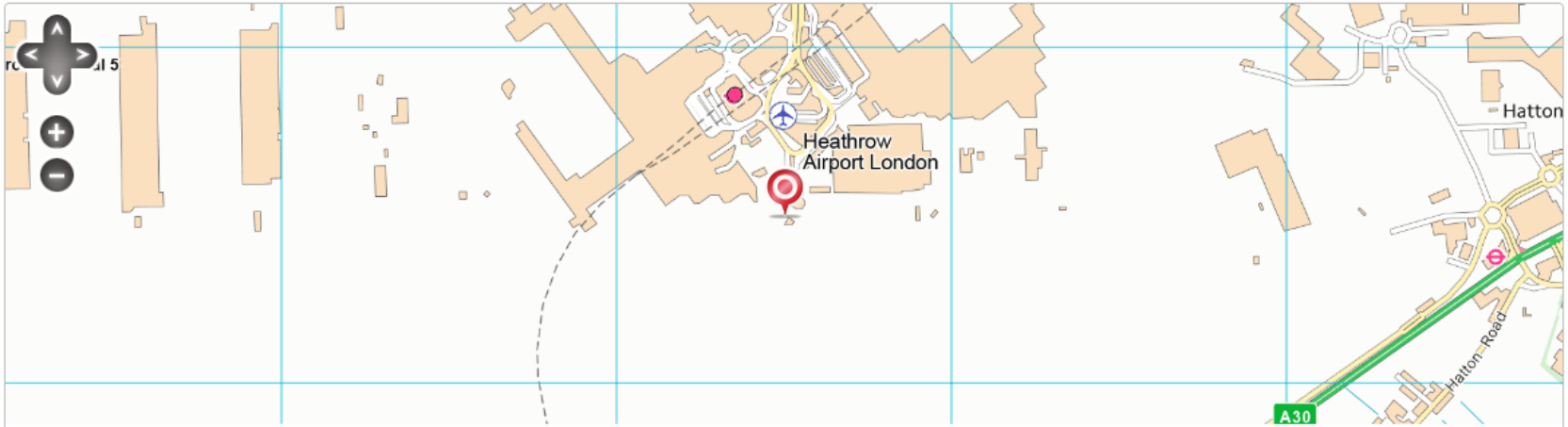
Heathrow London Airport (Named Place)

Ordnance Survey

You are here: linked-data » ordnance-survey-linked-data » heathrow-airport-london

Heathrow Airport London

Map powered by OS OpenSpace



Zoom to: Country County District City Area **City** Street

Heathrow Airport London is a Named Place.

Objects related to "Heathrow Airport London"

TwentyKMGrid Reference	TQ06
Feature Type	Other
OneKMGrid Reference	TQ0775

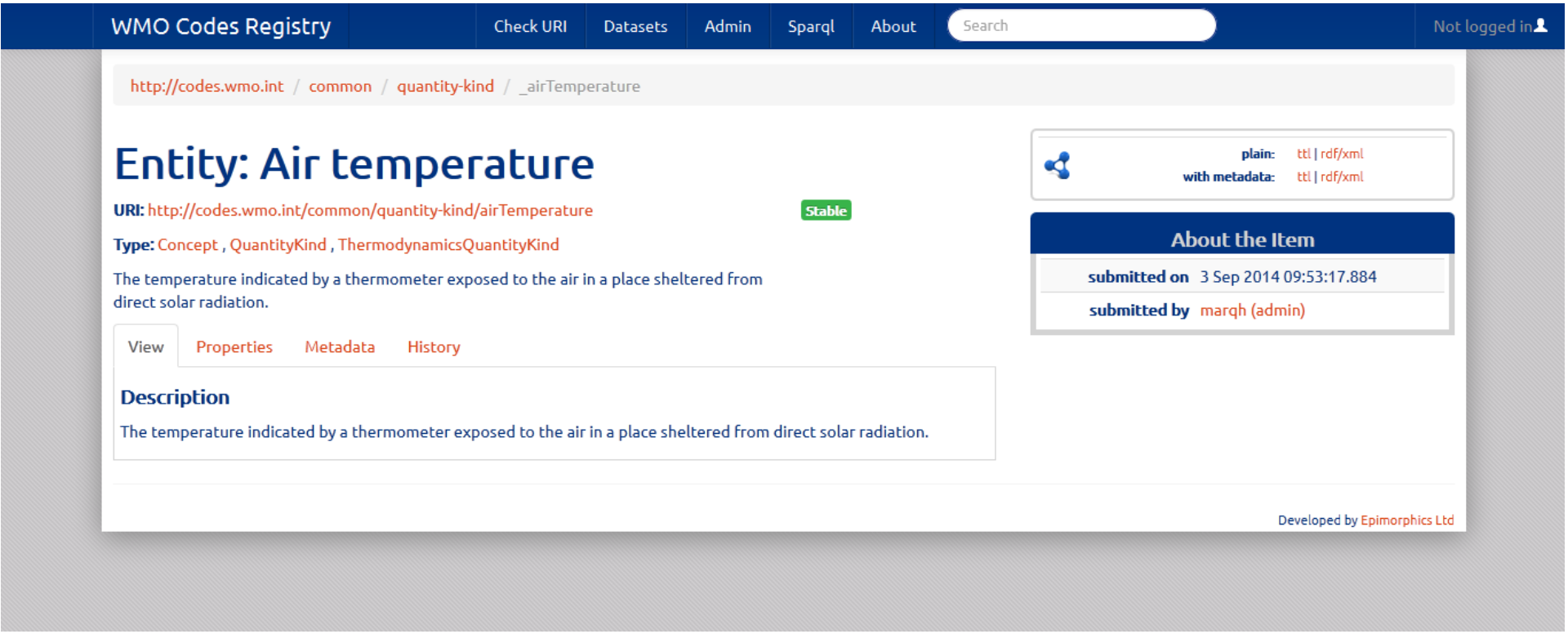
Core facts about "Heathrow Airport London"

Type	Named Place
Label	Heathrow Airport London
Name	Heathrow Airport London
Northing	175500
Easting	507500

Met Office

http://data.ordnancesurvey.co.uk/id/50kGazetteer/112605

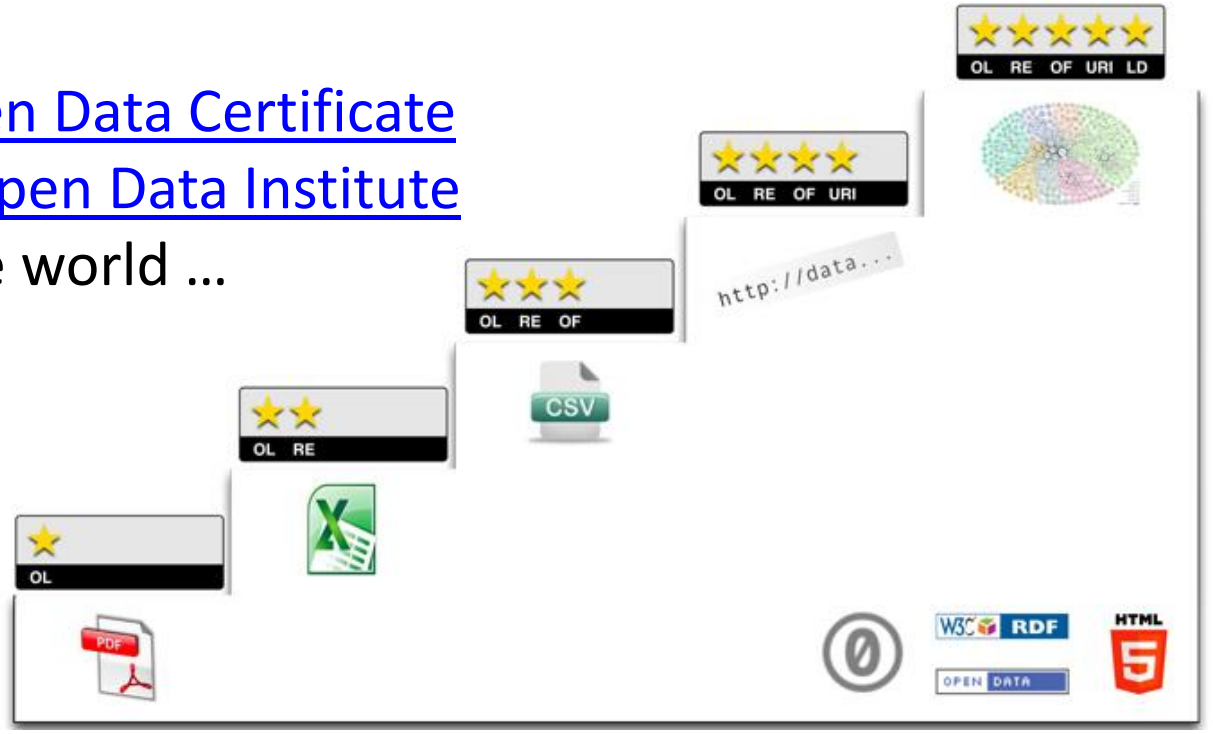
Air temperature (Quantity kind)



http://codes.wmo.int/common/c-15/me/airTemperature

So now you have 5★ Open Data …

Get an Open Data Certificate
from the Open Data Institute
and tell the world …



ODI
OL RE OF URI LD

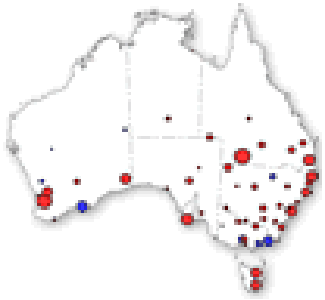
Standard level
self certified

Embed this on your site

A real example: ACORN-SAT

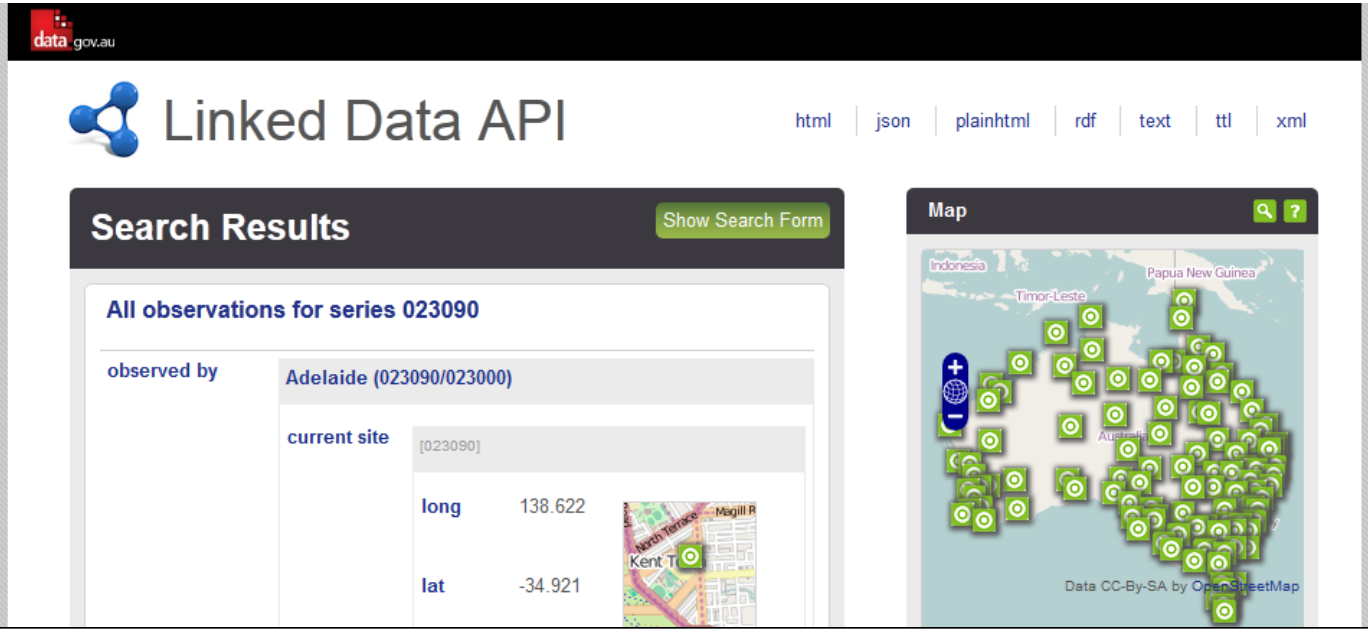
Australian Government
Bureau of Meteorology

[Australian Climate Observations Reference Network – Surface Air Temperature](#) (ACORN-SAT)



[ACORN-SAT: Linked-data innovation area](#)

Experimental Environmental Linked-data published by the Bureau of Meteorology



Met Office

Thank you