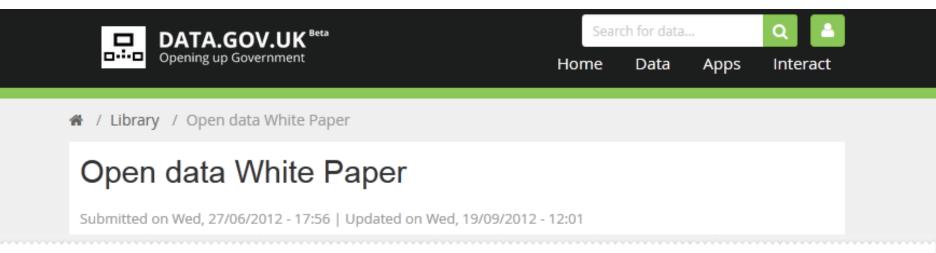


Publishing Linked Data

Jeremy Tandy November 2014



Why do we care about Linked Data?



- 2.41The 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?

					2		e4bd685c0873c02995913	- Microsoft E	xcel					-		x
Pas	Home		Page Layout Formula	A A	Review	View Wrap Text Image: A conter +	General	Condition			nsert Delete Format	Σ AutoSt	ZI Sort 8	Find &	9 -	
-	Clipboar		Font		AI	ignment 😡	Number	-	r as Table Styles		Cells	Clear •	Editing	r Select *		
<u></u>	A103	-	f _x 3772			-		<u>A</u>								×
	А		В	С	D	E	F		G	Н	I	J	К	L		
1 5	Site Code	Site Name	e	Latitude	Longitude F	Region	Observation Tin	e Observa	tion Date	Wind Direct	tion Wind Speed	Wind Gust	Visibility	Screen Ten	nperature	Pre 🔻
96	3684	ANDREWS	S FIELD (3684)	51.896	0.453 E	ast of England	12	00 2	7/06/2014	SSW	14		17000		18.6	
97			SINGTON (ESAWS) (3647)	51.86		South West England	12	00 2	7/06/2014	SW	9		17000		13.4	
98	3743	LARKHILL	(3743)	51.201	-1.805 9	South West England	12	00 2	7/06/2014	SSW	7	,	30000		15.9	
99	3809	CULDROSE	E (3809)	50.085	-5.257 \$	South West England	12	00 2	7/06/2014	W	11	. 24	20000		16.6	;
100	3840	DUNKESW	VELL AERODROME (3840)	50.86	-3.239 9	South West England	12	00 2	7/06/2014	S	15		23000		15.1	
101	3649	BRIZE NO	RTON (3649)	51.758	-1.576 l	ondon & South East Engla	and 12	00 2	7/06/2014	SW	3		30000		14.4	
102	3749	MIDDLE W	/ALLOP (3749)	51.15	-1.57 l	ondon & South East Engla	and 12	00 2	7/06/2014	WSW	9		30000		16.9	
103	3772	HEATHRO	W (3772)	51.479	-0.449 L	ondon & South East Engla	and 12	00 2	7/06/2014	SW	11		25000		20.3	
104	3866	ST CATHE	RINES PT. (3866)	50.577	-1.297 L	ondon & South East Engla	and 12	00 2	7/06/2014	WSW	9				16.6	
105	3002	BALTASOU	UND (3002)	60.749	-0.854 (Orkney & Shetland	12	00 2	7/06/2014	NNE	18		30000		11.3	
106	3023	SOUTH UI	ST RANGE (3023)	57.358	-7.397 H	Highland & Eilean Siar	12	00 2	7/06/2014	ENE	22		28000		14.3	
107	3037	SKYE/LUS	A (SAMOS) (3037)	57.257	-5.809 H	Highland & Eilean Siar	12	00 2	7/06/2014	E	11		21000		14.9	
108	3047	TULLOCH	BRIDGE (3047)	56.867	-4.708 H	Highland & Eilean Siar	12	00 2	7/06/2014	NE	9		35000		11.4	
109	3065	CAIRNGO	RM SUMMIT (3065)	57.12	-3.64 (Grampian	12	00 2	7/06/2014	NE	18				1	
110	3080	ABOYNE (3080)	57.077	-2.836 (Grampian	12	00 2	7/06/2014	NNE	8		40000		11.9	
111	3105	ISLAY/PO	RT ELLEN (3105)	55.681	-6.256 \$	Strathclyde	12		7/06/2014		15		29000		15.3	
112	3155	DRUMALB	BIN (3155)	55.627	-3.735 9	Strathclyde	12		7/06/2014		13		24000		11.9	
113			UGH (ESAWS) (3132)	54.859	-4.936 [Dumfries, Galloway	12		7/06/2014		5		40000		14.7	
114			GH/GOGARBANK (3166)	55.928	-3.343 [Dumfries, Galloway	12		7/06/2014		13		20000		12.6	
115		LOUGH FE		54.72	-6.82	Northern Ireland	12		7/06/2014		7		25000		12.6	
116	3923	GLENANN	IE (3923)	54.237	-6.502	Northern Ireland	12		7/06/2014		6	i	23000		12.3	
117		RHYL (331		53.259	-3.509 \	Vales	12		7/06/2014		3		11000		14.2	
118			NWY SAWS (3410)	52.757	-3.464 \		12		7/06/2014		5		10000		12.3	
119			HAVEN C.B. (3604) ee4bd685c0873c029959	51.708	-5.055 \	Vales	12	00 2	7/06/2014	S	10		25000		15.7	
Read		5-1400a/90	eenuudojiuo/jiu29959							Average: 7964.		m: 71683.83	mam	100% (-)	U	
near	,				-					Andruger 75044	evanari su			10000	V	0

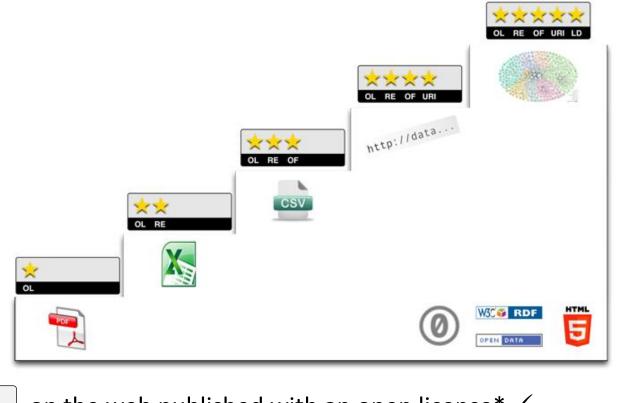


source: http://data.gov.uk/metoffice-data-archive



(all observing sites, 2014-06-27T12:00)

5 **★** (Linked) Open Data





on the web published with an open license* \checkmark



structured data \checkmark



non-proprietary format \checkmark





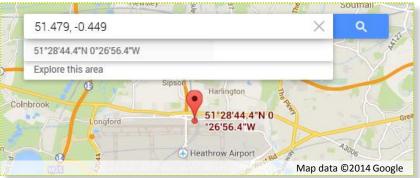
source: http://5stardata.info/



4★: use URIs to denote things 🗵

	Α	В	С	D	E	F	G	Н	1	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	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)"







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"}



#9

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: <u>URIs are hard</u>.

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, <u>UKGovLD</u> and partners <u>Epimorphics</u>, have developed an open source software solution: the <u>UKGovLD Registry</u>.

- 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

Met Office

Not limited to SKOS Collections or Concept Schemes





Deployment example: WMO Codes

WMO C	odes Registry	Check URI	Datasets	Admin	Sparql	About	Search					Not logged in 👤
	http://codes.wmo.int / bufr4 / co	odeflag / 0-22-06	51									
	Register: Stat				ea'	Stable		4		-	ttl rdf/xml ttl rdf/xml	
	· · ·	etadata							About th	e Reg	jister	
								_	owned by wmo			
	Members							ma	naged by www	dm		
			_					_	mitted on 25 Se		3 12:56:34.824	
	Item: 0 - Calm (glassy) Wave height in metres: 0		Тур	e: seaState	2		Stable	sub	mitted by boots	trap		
	Item: 1 - Calm (rippled) <i>Wave height in metres: 0 - 0.1</i>		Тур	e: seaState	2		Stable					
	Item: 2 - Smooth (wavelets) <i>Wave height in metres: 0.1 - 0.5</i>		Тур	e: seaState	•		Stable					
	Item: 3 - Slight Wave height in metres: 0.5 - 1.25		Тур	e: seaState	2		Stable					
	Item: 4 - Moderate Wave height in metres: 1.25 - 2.5		Тур	e: seaState	2		Stable					



http://codes.wmo.int/



Deployment example: Environment Registry

Environment Registry (alpha)

About Advanced -

List all registers

Browse

Filters

- Category
 - Marine and Coast [1]
 - Organizations_and_sectors [1]
 - System [12]
 - Water [5]

Owner

- Department for Environment, Food & Rural Affairs [6]
- Environment Agency [7]
- Marine Management
 Organisation [3]

Entity

- Abstract [7]
- Organizations and sectors [2]
- Regions and Habitats [1]

Name	Notation	Description	Status
🚔 root		Register representing the root of the registry tree.	Stable
Structure Structure	structure	Code lists used to aid organizing and presenting the Environment re	Stable
🖆 UI	ui	Codes and concepts used to guide the presentation of the registry u	Stable
Definitions	def	Code lists, concept schemes and other collections in the registry	Stable
着 System register	system	Internal registers which are used to control system operation	Stable
bulk collection types	bulkCollectionTypes	System register describing the data types which can be uploaded as	Stable
Link definitions	links	A system register which lists all link definitions which the UI sho	Stable
👕 prefixes	prefixes	System register containing the prefix mappings used in serializations	Stable







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?

```
What's the subject? There are
"@context": {"skos": "http://www.w3.org/200
"Site Code": "3772",
                                          two entities described in each
"Site Name": "HEATHROW (3772)",
"Latitude": "51.479",
                                          row ...
"Longitude": "-0.449",

    site (location); and

"Region": "London & South East England",
"Observation Time": "12:00",

    observation event

"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)" }
```







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"
                                             iClick!
},
"@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",
```

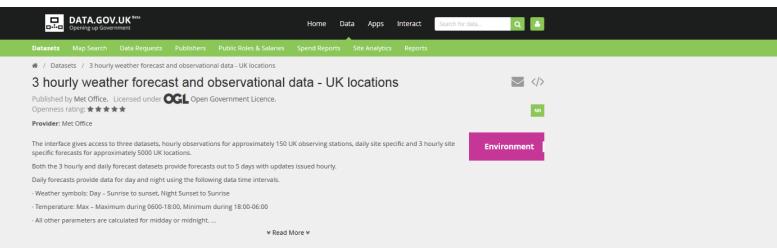


{

...



Resolve to a dataset description page?



DATA RESOURCES (1)





ADDITIONAL INFORMATION

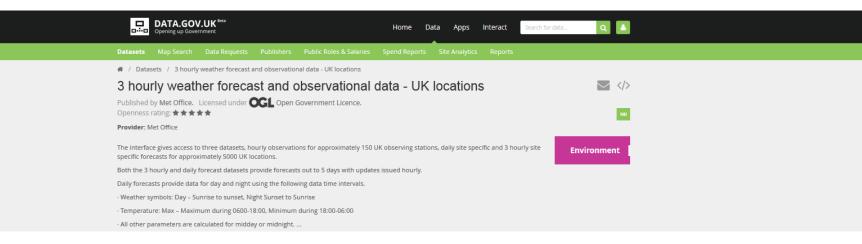
Added to data.gov.uk	29/11/2011
Theme	Environment
Geographic coverage	Great Britain (England, Scotland, Wales)
Precision	As supplied by Met Office
Update frequency	Hourly
Temporal granularity	hour
Taxonomy URL	http://www.metoffice.gov.uk/weather/uk/guide/key.html
Secondary Theme(s)	No value
Mandate	No value
Temporal coverage	No value
Date added computed	No value
Date updated computed	No value



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

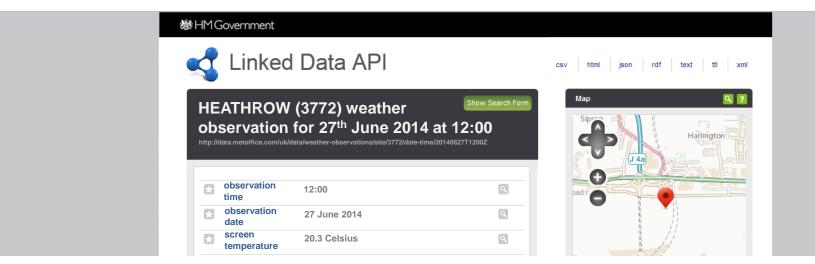
Resolve each data item ...

A HM Government Linked Data API CSV html rdf ttl xml ison text Мар Q ? Show Search Forr HEATHROW (3772) weather observation for 27th June 2014 at 12:00 Harlington observation ٩ 12:00 time observation 27 June 2014 Q, date screen 20.3 Celsius Q, temperature Q, visibility 25000 m Powered by copyright and database rights 2014 Ordnance Survey ٩, wind speed 11 mph wind direction South west 5 ? View ٩ SW notation label 8 significant Partly cloudy (day) observation time weather Construction date ٩ À notation 3 screen temperature visibility site 🚼 wind speed **Observation site at HEATHROW** wind direction > notation significant weather > notation easting 507500 🚼 site > name northing 177500 site > notation site > easting 📄 lat 51.479 site > northing Ordnance long -0.449 site > lat 🚼 site > lon 🚺 site 3772 **dataset** notation dataset uk weather observations





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



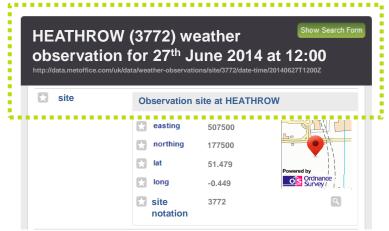
In addition to providing a dataset description page, one should publish information about each entity within the dataset.

Whilst the UKGovLD Registry provides a *simple* Linked Data publication platform, tools like <u>ELDA</u> 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 <u>triple</u>:

<u>s</u>ubject-<u>p</u>redicate-<u>o</u>bject.

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

"@id": "http://data.metoffice.com/uk/locations/obs/site/3772#id"





RDF triples and graphs



Triples can be grouped together into <u>graphs</u> 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

<u>Predicates</u> 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",
                                       ... often, we can reuse properties
  "rdfs:label": "HEATHROW (3772)",
  "geo:lat": "51.479",
                                      defined elsewhere – but you can't
  "geo:lon": "-0.449"
                                        avoid creating your own; ex:site.
```





RDF Vocabularies: data models by another name

A <u>vocabulary</u> (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 <u>data modelling</u>.

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.





\rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow OL RE OF URI LD

}}

Met Office

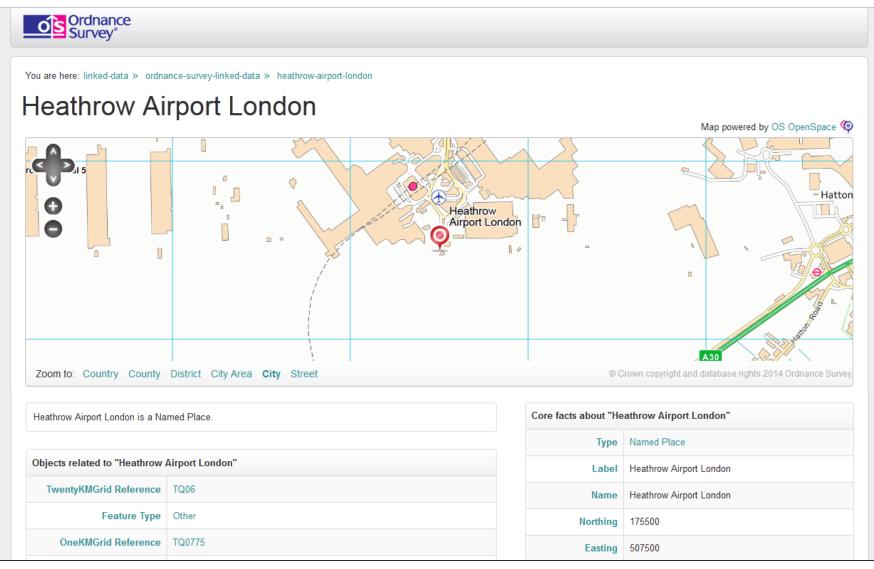
5 **★**: link to other people's data

```
"@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)"},
                                                             iClick!
  "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" }}
```

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



Heathrow London Airport (Named Place)





© Crown copyright #28

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

Air temperature (Quantity kind)

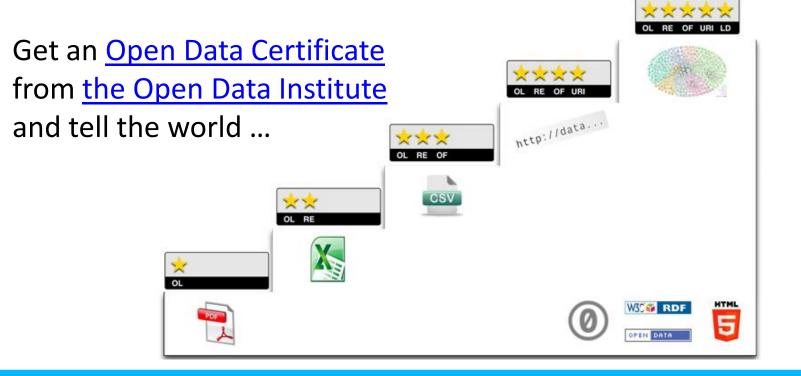
WMO Codes Registry	Check URI	Datasets	Admin	Sparql	About	Search					Not logged in
http://codes.wmo.int / common / quantity	- kind / _airTemp	erature									
Entity: Air tempe	erature	9					4	with	plain: metadata:	ttl rdf/xml ttl rdf/xml	
URI: http://codes.wmo.int/common/quantity-ki Type: Concept , QuantityKind , Thermodynamic	· · · ·	e		Stable				Abo	ut the II	em.	
The temperature indicated by a thermometer e direct solar radiation.		submitted on 3 Sep 2014 09:53:17.884 submitted by margh (admin)									
View Properties Metadata History	/										_
Description The temperature indicated by a thermometer of	exposed to the air	in a place she	ltered from	ı direct solar	radiation.						
									ſ	Developed by Epim	orphics Ltd





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

So now you have 5 🖈 Open Data ...







source: http://5stardata.info/



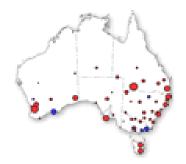
A real example: ACORN-SAT



Australian Government

Bureau of Meteorology

<u>Australian Climate Observations</u> <u>Reference Network – Surface Air</u> <u>Temperature</u> (ACORN-SAT)



ACORN-SAT: Linked-data innovation area

Experimental Environmental Linked-data published by the Bureau of Meteorology

data gov.au								
4	Link	ed Da	ita /	٩PI	htm	l json	plainhtml rdf	text ttl xml
Sea	rch Re	sults			Show Search Form		ap	Q 7
All o	bservatio	ns for series ()23090				Timor-Leste	Papua New Guinea
obser	ved by	Adelaide (023	090/02300)0)			+ <mark>.</mark> 9 6	
		current site	[023090]					
			long	138.622	Kent TO	i		
			lat	-34.921			Data	CC-By-SA by OpenStreetMap







Thank you

