LeGSB
Local e-Government Standards Body

Publishing local 5 star data.

- A Technical Companion to the DCLG ‘Code of Recommended Practice for Local Authorities on Data Transparency
- A ‘sandbox’ for Open Public Services
- Collaborating via Profiles for Linked Data

Paul Davidson, The Local e-Government Standards Body (LeGSB), and CIO of Sedgemoor District Council

To: Socitm South West, 21st June 2013, Exeter

LeGSB
Local e-Government Standards Body

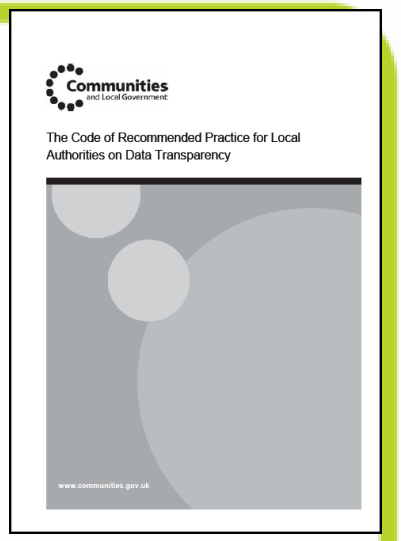
LeGSB's aim

To promote eStandards for Efficiency, Transformation, and Transparency of Local Services

- A Standard that supports all three, is particularly attractive
  - Predominantly about data standards
  - … and in particular, Open Data Standards
- Standards that can be re-used across many …
  - Disciplines / Sectors – e.g. Welfare, Health, Justice
  - Information Uses – e.g. Interoperability, public open data, evidencing policy
- Standards that local public services need …
  - As directed by our sponsors
- Standards that build into an architecture
  - e.g. a data ecosystem

Code of recommended practice …

- expenditure over £500, (including costs, supplier and transaction information)
- senior employee salaries, names, budgets and responsibilities of staff paid over £58,200 - equivalent to the lowest Senior Civil Service pay band
- an organisational chart
- the 'pay multiple' - the ratio between the highest paid salary and the median average salary of the whole of the authority's workforce
- councillor allowances and expenses
- copies of contracts and tenders to businesses and to the voluntary community and social enterprise sector
- grants to the voluntary community and social enterprise sector should be clearly itemised and listed
- policies, performance, external audits and key inspections and key indicators on the authorities' fiscal and financial position
- the location of public land and building assets and key attribute information that is normally recorded on asset registers
- data of democratic running of the local authority including the constitution, election results, committee minutes, decision - making processes and records of decisions.



- Publication should be in open and machine-readable formats. The recommended 5 step journey to a fully open format is:
  - * Available on the web (whatever format) but with an open license
  - ** As for one star plus available as machine-readable structured data (e.g. Excel instead of image scan of a table)
  - *** As for two star plus use a non-proprietary format (e.g. CSV and XML)
  - **** All the above plus use open standards from the World Wide Web Consortium (such as RDF and SPARQL) and
  - ***** All the above plus link your data to other people's data to provide context.

LeGSB
Local e-Government Standards Body

Publishing up to step 3

Star Rating[5]	In practice	Pros	Cons
★ DL	This is typically a static document, perhaps containing data as tables. Often this would be web page as html[6], a document as pdf[7], or an image.	• easy to produce; • retains presentation and layout;	• cannot further manipulate the data such as sorting, filtering, summing etc; • cannot join or compare to other data, or earlier versions;
★★ DL RE	This is typically data, such as a spreadsheet, published in the format of the tool that was used to extract it, such as xls[8].	• no new tools or skills necessary; • data can be downloaded and further processed and analysed.	• assumes that the consumer has the same tool as the producer, or is able to use the format. • very large data sets might not be attractive to download.
★★★ DL RE OF	This is typically 2* type data published using an open format, such as csv[9] or xml[10].	• there is typically a choice of open source tools available for each open format	• Some conversion necessary. • Where data has greater structure than a simple 2-dimensional table, many files may be necessary, which may become disconnected; • The meaning and scope of columns and rows can be hard to express.

Publishing up to step 5

	This is explicitly about publishing data to the RDF[11] data model, and providing a query service using the SPARQL[12] language.	• enables others to make statements over the web about individual lines of data; • can be queried over a data service so that a complete data set does not have to be downloaded; • gives precise definitions to the meaning of the data.	• requires a skill set that most local authorities do not have yet; • not suitable to provide directly to residents. • Requires some infrastructure
	This is Linked Data in the RDf model, that contains links to external datasets to describe the 'things' that the data refers to.	• adds context; • can build into a 'data ecosystem'. • can enable 3rd parties to join data from many sources together to provide new targeted information services and insight.	• requires a 'spine' of core reference data to make links to.

Why do we need this guide?

- Feedback from local authorities tells us that, while many wish to publish data in a smarter, joined up way, that will benefit their residents, they have not been able to find relevant material that explains the practicalities of 5* data publishing.

- At LGA events across England in March 2013, LeGSB presented the material in this guide to roomfuls of local authority practitioners, who then fed back that they would value it being turned into a proper guide, and were likely to take part in supporting web collaborations.

"Understanding this document will allow your organisation to be at the forefront of the open linked data agenda melding it to be of use to your citizens and your organisation."

Roger Hampson, Chief Executive LB Redbridge

a really good piece of work - simple to read, comprehensive and helpful.

Jos Creese, CIO Hampshire C.C.

LeGSB
Local e-Government Standards Body

Get the guide

- http://legsb.i-network.org.uk/resources/publishinglocal5stardata/
- Explanations of Linked Data concepts and techniques

If the example illustrates a fundamental point about linked data, and how it applies to local public data, we will highlight the point with a box like this.

- Examples of local authorities publishing 5* data
  - Bristol City Council – Air Quality Data
  - Hampshire County Council – Land Use Projections
  - Devon County Council – Community Neighbourhood Budgeting
- Step by step walk through of
  - Modelling
  - Transformation into RDF
  - Querying
  - Building applications

If a point needs further explanation than we have space for here, we will provide more at the companion website, and highlight it with a box like this.

If the example draws out a design consideration for local public data, we will highlight that with a box like this.

- Draws out ‘gaps’
  - which ‘concepts’ commonly recur in local public service data?
  - which properties/URI Sets/controlled vocabularies should we consistently use?
  - what existing linked data is already available that could be linked to?
  - what ‘core reference data’ is missing, which we will need to be able to make links in linked data?

LeGSB
Local e-Government Standards Body

Bristol City Council – Air Quality Data

- Get 3* Data

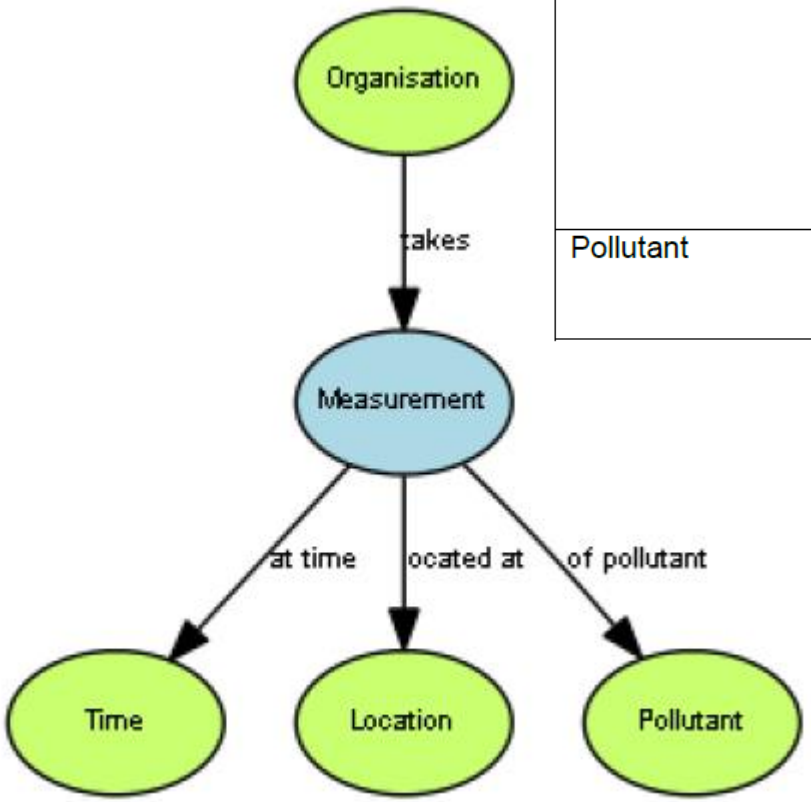
Epimorphics Ltd
Linked data solutions

	A	B	C	D	E	F	G	H	I	J	K
1	SiteID	Date	Time	NO2raw_ppb	NOxraw_ppb	NOraw_ppb	NO2rat_ppb	NOxrat_ppb	NOrat_ppb	CAQI	
2	203	01/01/2001	00:30:00	13.5	65	50.5	4.73	19.44	14.71	2	
3	203	01/01/2001	00:45:00	13	55.5	41.5	4.58	16.57	11.99	2	
4	203	01/01/2001	01:00:00	-10	-10	-10				2	
5	203	01/01/2001	01:15:00	-10	-10	-10				2	
6	203	01/01/2001	01:30:00	12.5	46	32.5	4.43	13.7	9.27	2	
7	203	01/01/2001	01:45:00	12.5	36.5	23.5	4.28	10.83	6.55	2	
8	203	01/01/2001	02:00:00	11	54.5	43	3.82	16.26	12.44	2	
9	203	01/01/2001	02:15:00	9	44	35.5	2.92	13.09	10.17	2	
10	203	01/01/2001	02:30:00	10.5	35.5	24	3.82	10.52	6.7	2	
11	203	01/01/2001	02:45:00	12	34	22	3.97	10.07	6.09	2	
12	203	01/01/2001	03:00:00	11.5	32	20	3.97	9.47	5.49	2	
13	203	01/01/2001	03:15:00	11.5	40.5	28	4.13	12.03	7.91	2	
14	203	01/01/2001	03:30:00	13	32	18	4.58	9.47	4.89	2	
15	203	01/01/2001	03:45:00	11	42.5	31.5	3.67	12.64	8.97	2	
16	203	01/01/2001	04:00:00	12.5	31.5	18	4.43	9.31	4.89	2	
17	203	01/01/2001	04:15:00	11	23	11.5	3.82	6.75	2.92	2	
18	203	01/01/2001	04:30:00	11	17.5	6	3.82	5.09	1.26	2	
19	203	01/01/2001	04:45:00	11	19	7	3.97	5.54	1.56	2	
20	203	01/01/2001	05:00:00	10	12.5	2.5	3.37	3.58	0.2	1	
21	203	01/01/2001	05:15:00	8.5	10	1.5	2.92	3.07	0.15	1	
22	203	01/01/2001	05:30:00	10	12.5	2.5	3.37	3.82	0.45	1	
23	203	01/01/2001	05:45:00	8.5	11	2.5	2.92	3.37	0.45	1	

Bristol City Council – Air Quality Data

- Model it

Organisation	Bristol City Council. • What type of organisation is that? • What geographic area are they responsible for? • What is their air quality strategy? • How much do they spend on monitoring air quality?
Time and Location	What other events have also occurred around that time and location? • traffic counts • health incidents? Who lives at that location? • population • demographics • deprivation
Pollutant	What do we know about the pollutant? • Safe levels? • Health implications



Bristol City Council – Air Quality Data

- Transform it

```
<http://bristol-data-epimorphics.dyndns.org/data/environment/air-quality/observation/conc/location/00203/date/1999-01-01/time/00:15:00/duration/PT15M/substance/NO>
        <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/month> "01" ;

        <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/interval> <http://reference.data.gov.uk/id/gregorian-interval/1999-01-01T00:15:00/PT15M>;

        <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/AirQualityMeasurement>;

        <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://purl.org/linked-data/cube#Observation>;

        <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/monitoringPoint> <http://bristol-loc-epimorphics.dyndns.org/so/ef/MonitoringPoint/aqmp.bcc/00203>;

        <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/year> "1999";

        <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/ratifiedReading> "10";

        <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/pollutant> <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/air-pollutant/NO>;

        <http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/dateTime> "1999-01-01T00:15:00";

        .
```

Bristol City Council – Air Quality Data

- Triples

http://bristol-data-epimorphics.dyndns.org/data/environment/air-quality/observation/conc/location/00203/date/1999-01-01/time/00:00:00/duration/PT15M/substance/NOX

http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/ratifiedReading

"6"

Bristol City Council – Air Quality Data

- Triples

http://bristol-data-epimorphics.dyndns.org/data/environment/air-quality/observation/conc/location/00203/date/1999-01-01/time/00:00:00/duration/PT15M/substance/NOX

http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/core/pollutant

http://bristol-data-epimorphics.dyndns.org/def/environment/air-quality/air-pollutant/NOX

LeGSB
Local e-Government Standards Body

Collaborating

http://legsb.i-network.org.uk/resources/publishinglocal5stardata/

Help us to improve the standards and techniques

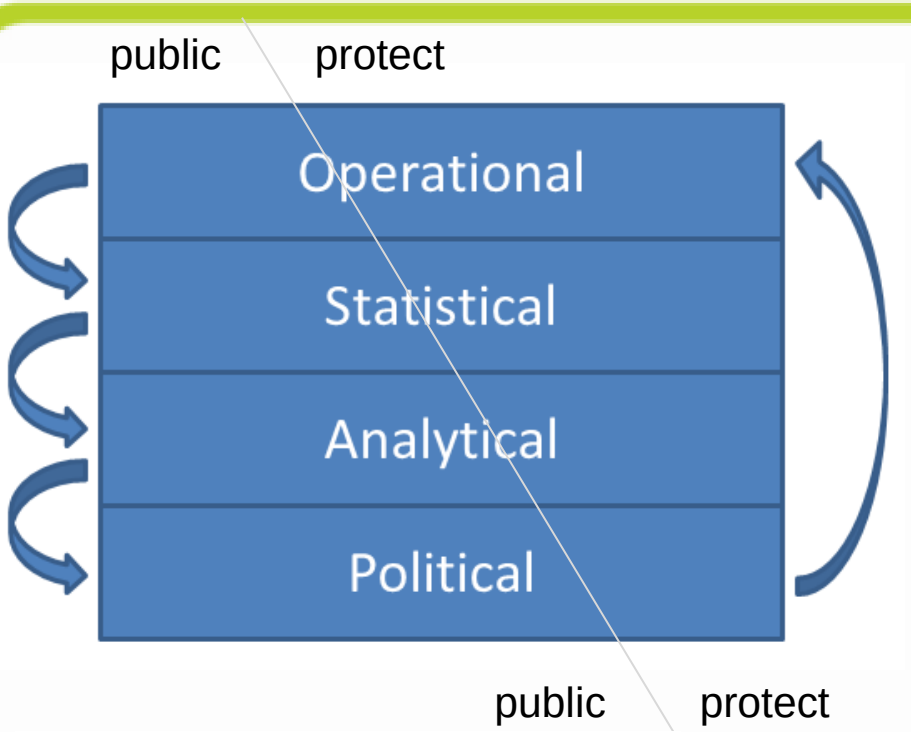
The LeGSB guide raises topics where we provide more information, invite local authorities and linked data expertise to collaborate on

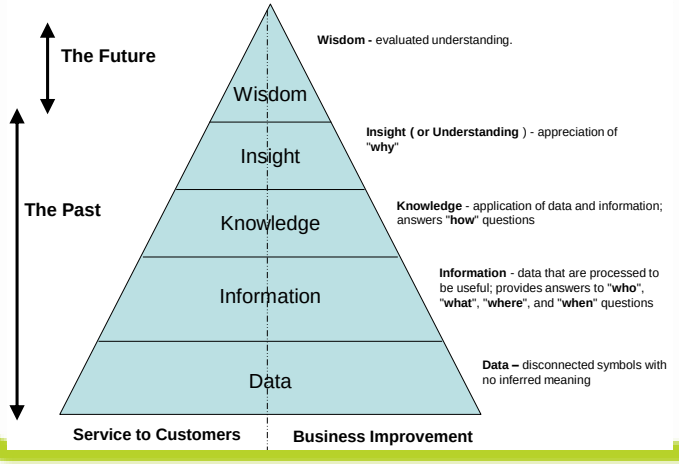
- tools and methods for modelling;
- data formats to respond with when dereferencing URIs;
- options for triplestores;
- describing quality and provenance;
- patterns for common scenarios such as statistics;

The guide demonstrates that 5* data publishing requires external data to link to – made up of definitions of concepts, and identifiers for 'things'. To create a data ecosystem, we need to collaborate to determine

- which 'concepts' commonly recur in local public service data?
- which properties/URI Sets/controlled vocabularies should we consistently use?
- what existing linked data is already available that could be linked to?
- what 'core reference data' is missing, which we will need to be able to make links in linked data?

Making local links

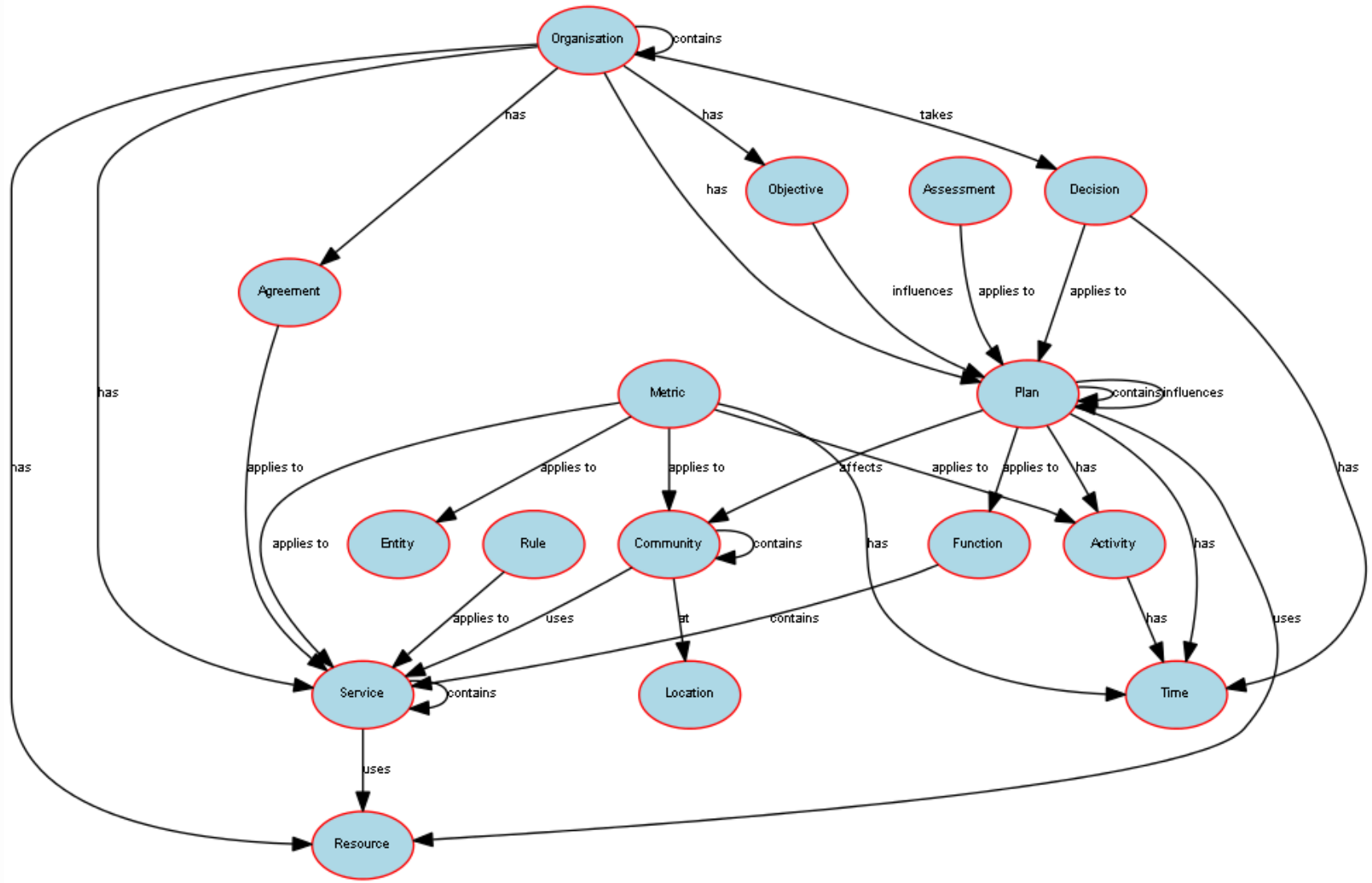




- OPERATIONAL - Data about real people and places, with real needs and circumstances, using real services, i.e. case work
- STATISTICAL - Aggregated operational data – organised using common classifications and segmentations
- ANALYTICAL - The conclusions drawn from an analysis of statistical data
- POLITICAL - The decisions taken to shape services, e.g. budgets, strategies, priorities, targets etc.

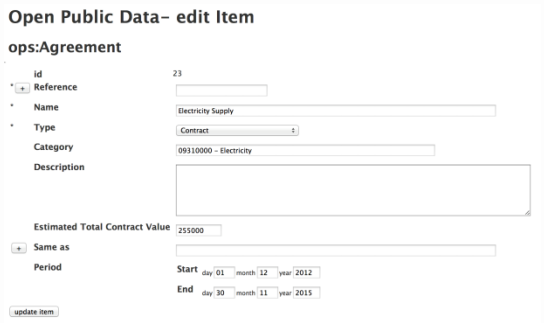
LeGSB
Local e-Government Standards Body

Joining it up?



Sand Box

- http://www.openpublicdata.com




```
<http://data.sedgemoor.gov.uk/id/contracts/agreement/23>
        a ops:Agreement ;
        skos:prefLabel """Electricity Supply""" ;
        ops:classification """09310000 - Electricity""" ;
        dcterms:type """Contract""" ;
        ops:estimatedtotalcontractvalue """255000""" ;
        .

<http://data.sedgemoor.gov.uk/id/contracts/agreement/23> ops:hasPeriod <http://data.sedgemoor.gov.uk/id/contracts/period/23> .

<http://data.sedgemoor.gov.uk/id/contracts/period/23>
        a ops:Period ;
        ops:periodStart """2012-12-01""" ;
        ops:periodEnd """2015-11-30""" ;
        .

<http://data.sedgemoor.gov.uk/id/contracts/agreement/23> ops:awardedBy <http://data.sedgemoor.gov.uk/id/council/organisation/1> .

<http://data.sedgemoor.gov.uk/id/contracts/agreement/23> ops:awardedTo <http://data.sedgemoor.gov.uk/id/contracts/organisation/23> .

<http://data.sedgemoor.gov.uk/id/contracts/organisation/23>
        a org:Organization ;
        skos:prefLabel """Eon UK plc,""" ;
        .
```

LeGSB
Local e-Government Standards Body

Profiles?

- Working with W3C – RDF Validation
  - To define how a dataset has used classes, properties, uri-sets etc, from many ontologies/sources.
  - To invite others with similar data to use the same profile, and therefore find that their data can be queried together.
  - To be able to register a dataset so that it can be discovered by its ‘Profile’.
  - To be able to ‘query’, and present a dataset (or across many datasets) by referring to its ‘profile’
  - To be able to validate that a dataset does conform to a ‘profile’
  - To enable others to create a new profile by adapting an existing one.
  - To be able to create a register of good practice ‘patterns’ for commonly recurring data structures, that can then be re-used.

LeGSB
Local e-Government Standards Body

Legsb
The Local e-Government Standards Body

paul.davidson@sedgemoor.gov.uk

legsb.i-network.org.uk