

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: UkGovLd, 19th June 2013, Manchester

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.



The Code of Recommended Practice for Local Authorities on Data Transparency



- 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.

Publishing up to step 3

Star Rating ⁵	In practice	Pros	Cons
	<p>This is typically a static document, perhaps containing data as tables.</p> <p>Often this would be web page as html⁶, a document as pdf⁷, or an image.</p>	<ul style="list-style-type: none"> easy to produce; retains presentation and layout; 	<ul style="list-style-type: none"> cannot further manipulate the data such as sorting, filtering, summing etc; cannot join or compare to other data, or earlier versions;
	<p>This is typically data, such as a spreadsheet, published in the format of the tool that was used to extract it, such as xls⁸.</p>	<ul style="list-style-type: none"> no new tools or skills necessary; data can be downloaded and further processed and analysed. 	<ul style="list-style-type: none"> 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.
	<p>This is typically 2* type data published using an open format, such as csv⁹ or xml¹⁰.</p>	<ul style="list-style-type: none"> there is typically a choice of open source tools available for each open format 	<ul style="list-style-type: none"> 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

	<p>This is explicitly about publishing data to the RDF¹¹ data model, and providing a query service using the SPARQL¹² language.</p>	<ul style="list-style-type: none"> 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. 	<ul style="list-style-type: none"> requires a skill set that most local authorities do not have yet; not suitable to provide directly to residents. Requires some infrastructure
---	---	---	---

	<p>This is Linked Data in the Rdf model, that contains links to external datasets to describe the 'things' that the data refers to.</p>	<ul style="list-style-type: none"> 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. 	<ul style="list-style-type: none"> 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.

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.

- 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?



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



Home Publishing Local 5* Data Events News White Papers Projects Promoted Standards Consultations Governance

Publishing Local 5* Data

This LeGSB guide is currently a draft. Various Linked Data experts have checked it for us and we have been able to improve it as a result. The final version will be available soon when we get endorsements and recommendations from other organisations, but for now, here is the draft.

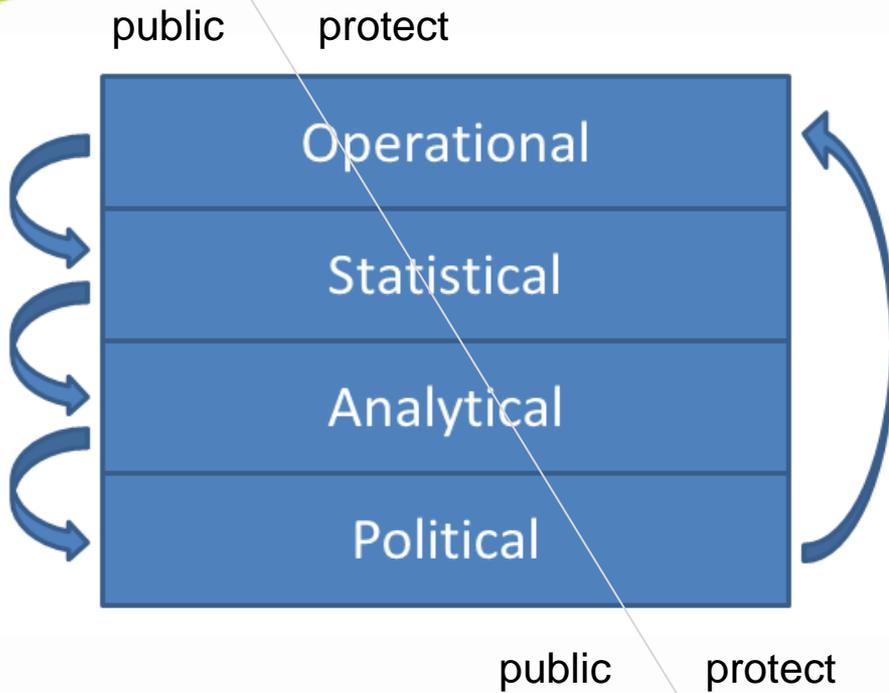
[LeGSB-PublishingLocal5StarData-Rev3](#)

Please give us feedback to this document using the form at the bottom of the page. You will see that we have also provided further pages to discuss some of the issues and 'gaps' that the guide highlights.

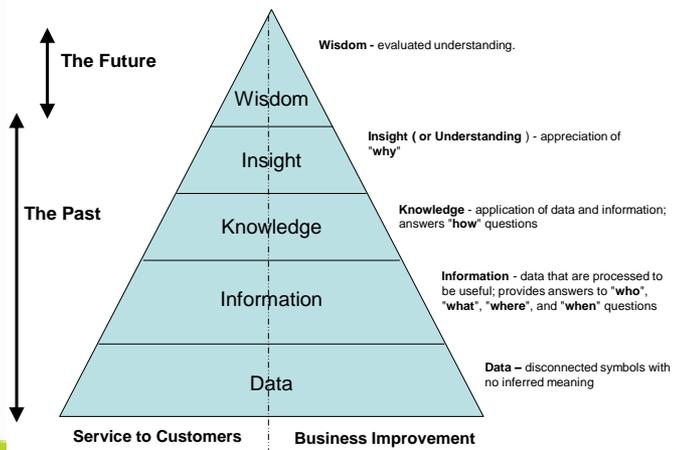
Why do we need this guide?

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

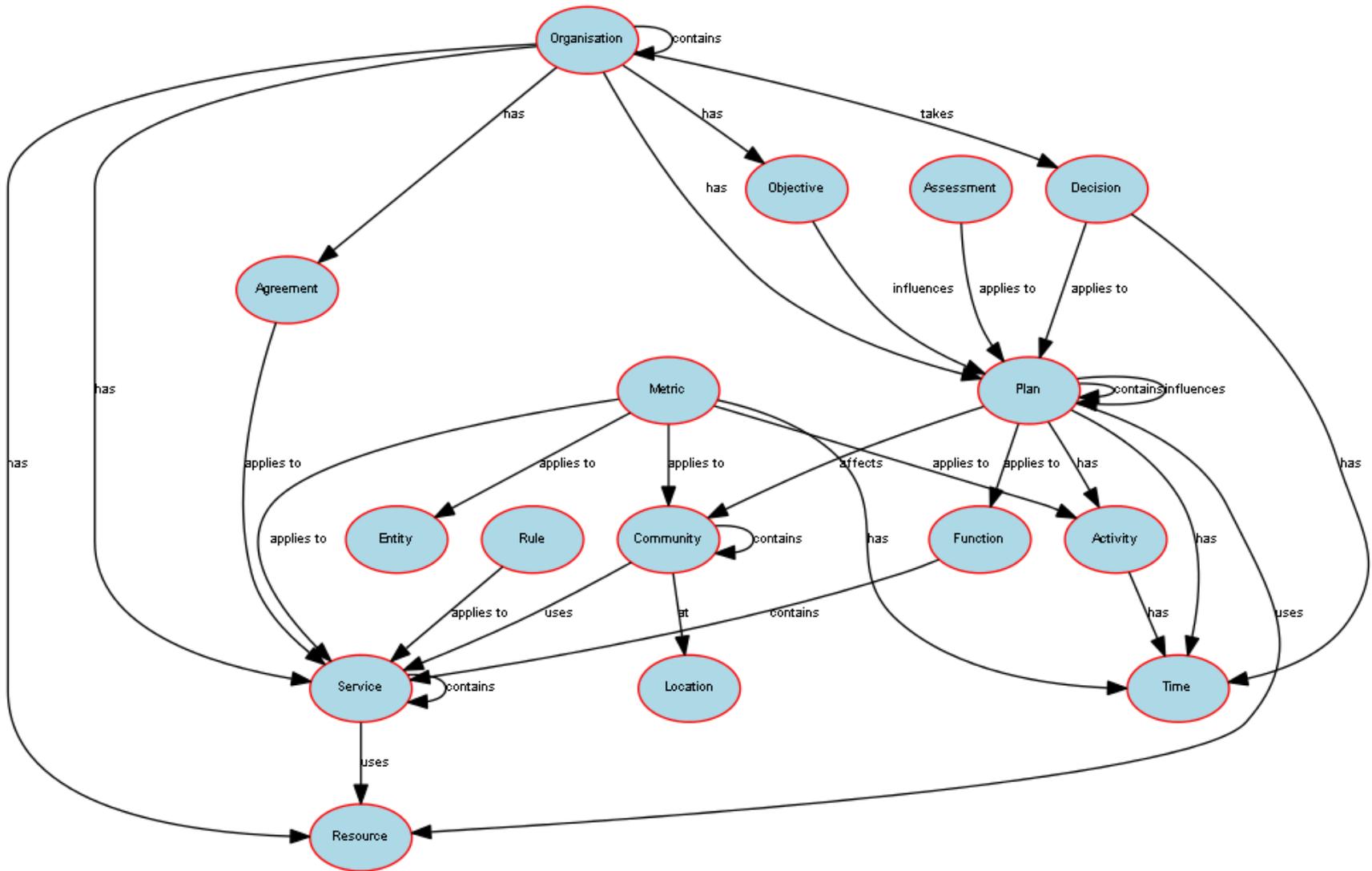
Making local links



- **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.



Joining it up?



- <http://www.openpublicdata.com>



PaulD

Open Public Data- view Dataset

Name [About the Council](#)

View Organisation

Collection council

Published Data

- [Browse](#)
- [Query](#)

[edit](#)

My Data

- [Edit](#)
- [Import](#)
- [Plug Board](#)

Items

org:FormalOrganization

id	source
1	Sedgemoor District Council

1 Sedgemoor District Council

[add](#)

[select uri](#)

foaf:Person

id	source

[add](#)

[select uri](#)

```
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix org: <http://www.w3.org/ns/org#> .
@prefix time: <http://www.w3.org/2006/time#> .
@prefix vcard: <http://www.w3.org/2006/vcard/ns#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix wgs84: <http://www.w3.org/2003/01/geo/wgs84#> .

<http://data.sedgemoor.gov.uk/id/council/organisation/1>
  a org:FormalOrganization ;
  skos:prefLabel ""Sedgemoor District Council"" ;
  org:classification <http://opendatacommunities.org/def/local-government/DistrictCouncil> ;
  dcterms:description ""Sedgemoor District Council is a District Council in the County of Somerset, UK"" ;
  owl:sameAs <http://opendatacommunities.org/doc/district-council/sedgemoor> ;
  .

<http://data.sedgemoor.gov.uk/id/council/organisation/1> org:hasUnit <http://data.sedgemoor.gov.uk/id/council/unit/1> .

<http://data.sedgemoor.gov.uk/id/council/unit/1>
  a org:OrganizationalUnit ;
  skos:prefLabel ""The Executive"" ;
  dcterms:description ""The Executive is the part of the Council which is responsible for most day-to-day decisions. The Executive is responsible for the day-to-day running of the Council. When major decisions are to be discussed or made, these are published in the Forward Plan in so far as they can be anticipated. If these decisions are not published in the Forward Plan, the Executive has to make decisions which are in line with the Council's overall policies and budget. If it wishes to make a decision which is not in line with the Council's overall policies and budget, it must refer the matter to the Council for a decision."";
  .
```

Profiles?

- 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
The Local e-Government Standards Body

paul.davidson@sedgemoor.gov.uk

www.legsb.gov.uk