

Publishing local 5 star data.

A Technical Companion to the DCLG 'Code of Recommended Practice for Local Authorities on Data Transparency'.

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

To: The Local Public Data Panel, 5<sup>th</sup> June 2013, London.

- **Mission**

- *To promote eStandards that support Efficiency, Transformation, and Transparency of Local Services*
- *Syntax, Semantics, Quality, Rights, Authentication, Transport, Governance*

- **Funded by sponsorship from**

- **Cabinet-Office**
  - Driving the work of the Information Domain
  - Information Standards to support the Infrastructure initiatives from the government ICT Strategy
- **DWP**
  - Universal Credits, ATLAS message formats
- **Communities and Local Government**
  - Transparency, Comparable Financial Data

- **Governance**

- Originally an ODPM National Project
- A part of the iNetwork
- Accountable body is Tameside Council

- **Executive Steering Board**

- LGA, SOCITM, DWP, CLG, C.O., Health, Education, Intellect, Information Commissioner's Office

- **Representing to**

- CTO Delivery Group
- UK Location Council
- Knowledge Council
- EC, ISA European Interoperability Architecture Working Groups

# 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 <sup>5</sup>	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 html<sup>6</sup>, a document as pdf<sup>7</sup>, or an image.</p>	<ul style="list-style-type: none"> <li>easy to produce;</li> <li>retains presentation and layout;</li> </ul>	<ul style="list-style-type: none"> <li>cannot further manipulate the data such as sorting, filtering, summing etc;</li> <li>cannot join or compare to other data, or earlier versions;</li> </ul>
	<p>This is typically data, such as a spreadsheet, published in the format of the tool that was used to extract it, such as xls<sup>8</sup>.</p>	<ul style="list-style-type: none"> <li>no new tools or skills necessary;</li> <li>data can be downloaded and further processed and analysed.</li> </ul>	<ul style="list-style-type: none"> <li>assumes that the consumer has the same tool as the producer, or is able to use the format.</li> <li>very large data sets might not be attractive to download.</li> </ul>
	<p>This is typically 2* type data published using an open format, such as csv<sup>9</sup> or xml<sup>10</sup>.</p>	<ul style="list-style-type: none"> <li>there is typically a choice of open source tools available for each open format</li> </ul>	<ul style="list-style-type: none"> <li>Some conversion necessary.</li> <li>Where data has greater structure than a simple 2-dimensional table, many files may be necessary, which may become disconnected;</li> <li>The meaning and scope of columns and rows can be hard to express.</li> </ul>

# Publishing up to step 5

	<p>This is explicitly about publishing data to the RDF<sup>11</sup> data model, and providing a query service using the SPARQL<sup>12</sup> language.</p>	<ul style="list-style-type: none"> <li>enables others to make statements over the web about individual lines of data;</li> <li>can be queried over a data service so that a complete data set does not have to be downloaded;</li> <li>gives precise definitions to the meaning of the data.</li> </ul>	<ul style="list-style-type: none"> <li>requires a skill set that most local authorities do not have yet;</li> <li>not suitable to provide directly to residents.</li> <li>Requires some infrastructure</li> </ul>
---	---	---	---

	<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"> <li>adds context;</li> <li>can build into a 'data ecosystem'.</li> <li>can enable 3<sup>rd</sup> parties to join data from many sources together to provide new targeted information services and insight.</li> </ul>	<ul style="list-style-type: none"> <li>requires a 'spine' of core reference data to make links to.</li> </ul>
---	---	--	---

## 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.

## Your thoughts?

- Does the panel support the ‘technical companion’
  - as recommended reading for Local Authorities?
- What type of local data is valuable when provide as 5\*
  - and what is the vision for a ‘local data ecosystem’?
- The companion lists various gaps that need to be address before Local Authorities could make a success of 5\* publishing
  - how do we address those?

# Addressing Gaps

- Content
  - 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?
- Methods / Tools / Architecture / Standards
  - modeling
  - describing quality and provenance;
  - patterns for common scenarios such as statistics;
  - expressing the ‘profile’ of a dataset;
  - data formats to respond with when dereferencing URIs;
  - options for triplestores, APIs, data services;



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.

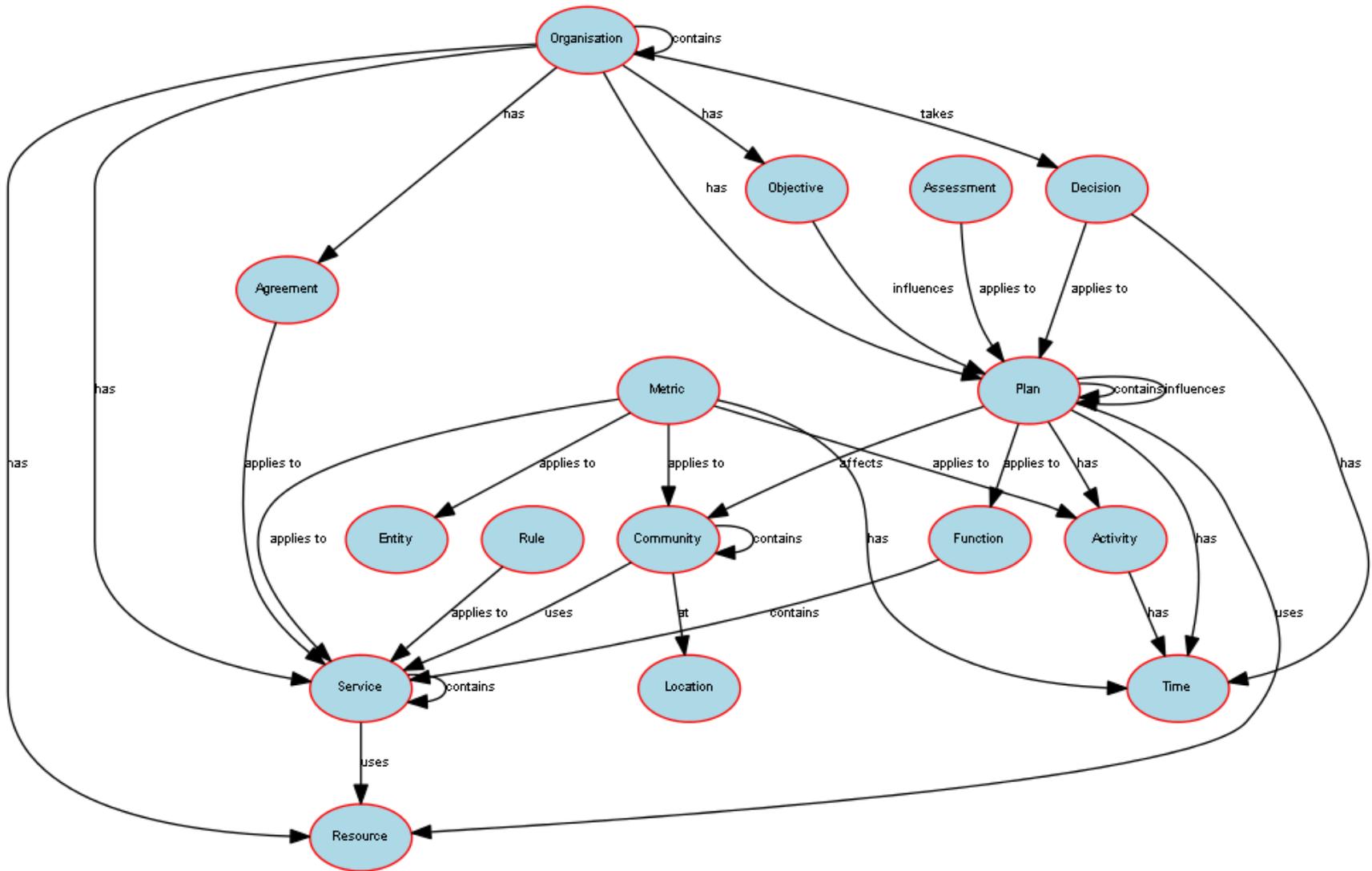
[LeGSB-PublishingLocal5StarData-Rev3](#)

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

# Joining it up?



Legsb  
The Local e-Government Standards Body

[paul.davidson@sedgemoor.gov.uk](mailto:paul.davidson@sedgemoor.gov.uk)

[www.legsb.gov.uk](http://www.legsb.gov.uk)