Missing(?) Building Blocks for a Linked Data Ecosystem

Axel Polleres
Siemens AG Österreich
State of LOD affairs:

• Is Linked Data the “Silver Bullet” for data integration?
  – We’ve come up with a nice interoperable data format (RDF) & a bunch of publishing principles
  – Enables lightweight data integration in a bottom-up, schema-less fashion

• LOD Cloud and table of datasets 339 datasets

<table>
<thead>
<tr>
<th></th>
<th># of LOD Datasets on CKAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jun-13</td>
<td>339 (+44)</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>Sep-11</td>
<td>295 (+92)</td>
</tr>
<tr>
<td>Sep-10</td>
<td>203 (+108)</td>
</tr>
<tr>
<td>Jul-09</td>
<td>95</td>
</tr>
</tbody>
</table>

• … slower growth rate or people too busy to let us know?
• More worrying or even harmful: Stale datasets…

→ “2010-06-14”
• Can we collect and use relevant open Data for a concrete domain (in our case: indicators that determine sustainability aspects of cities) in an automated fashion?
Suspicion/Experience/Lessons learnt

• Openness may raise eyebrows... API keys, different license models, etc. Different implications on reuse, e.g.
  – republication
  – aggregation
  ➔ How to implement pricing models? access control mechanisms? provenance tracking?

• Dynamicity vs. stale Datasets makes LOD idea hard to sell in an industry context:
  – e.g. NYT dataset (not even speaking of stream-data and the like here)

• OWL+RDFS are not enough for integration
  – Numerical data within implicit information
Missing(?) Building Blocks for a Linked Data Ecosystem:

• Not all Linked Data is open, needs mechanisms to declare & compose licenses.

• Linked Data needs standard mechanisms to deal with Dynamicity & Evolution.

• Linked Data Quality needs standard mechanisms to deal with Provenance (& Trust)

• Linked Data needs more (and less) than OWL

• and BTW: not all Linked data is RDF (JSON, CSV)...
Goal:

My expectation on the seminar is a road-map outlining

• 1. how these building blocks can implemented in terms of industry-strength standards & best practices that can interplay and scale on the Web
• 2. where further fundamental research is necessary