

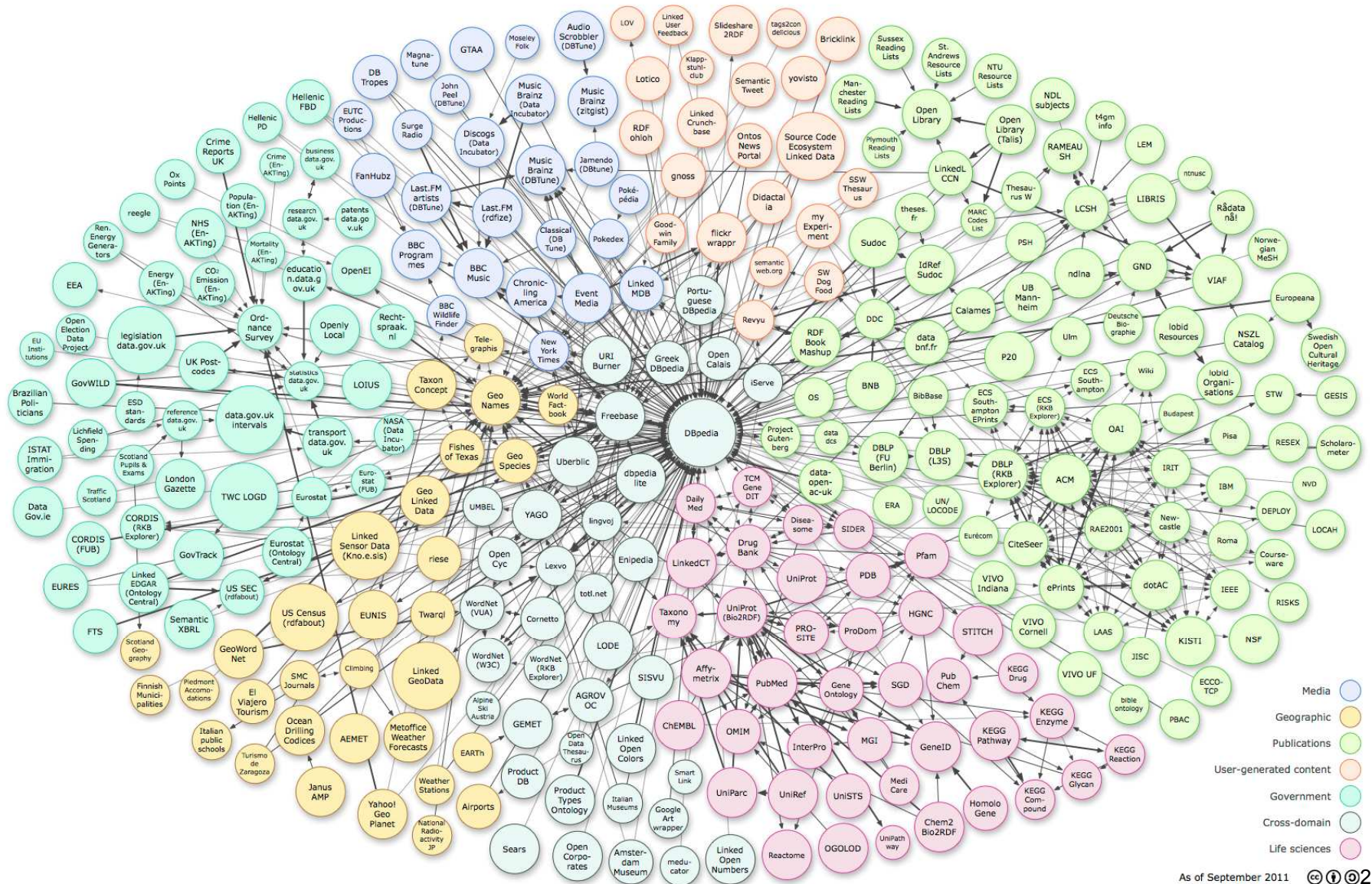
Ontology Driven Data Integration & Big Linked Open Data

By [Kingsley Idehen](#)

Founder & CEO, [OpenLink Software](#)



Linked Open Data (LOD) Cloud





Situation Analysis – Big Data

Data is growing exponentially
along the following dimensions:

- Volume
- Velocity
- Variety

All of this happens while the total
hours in day remains 24 hrs.



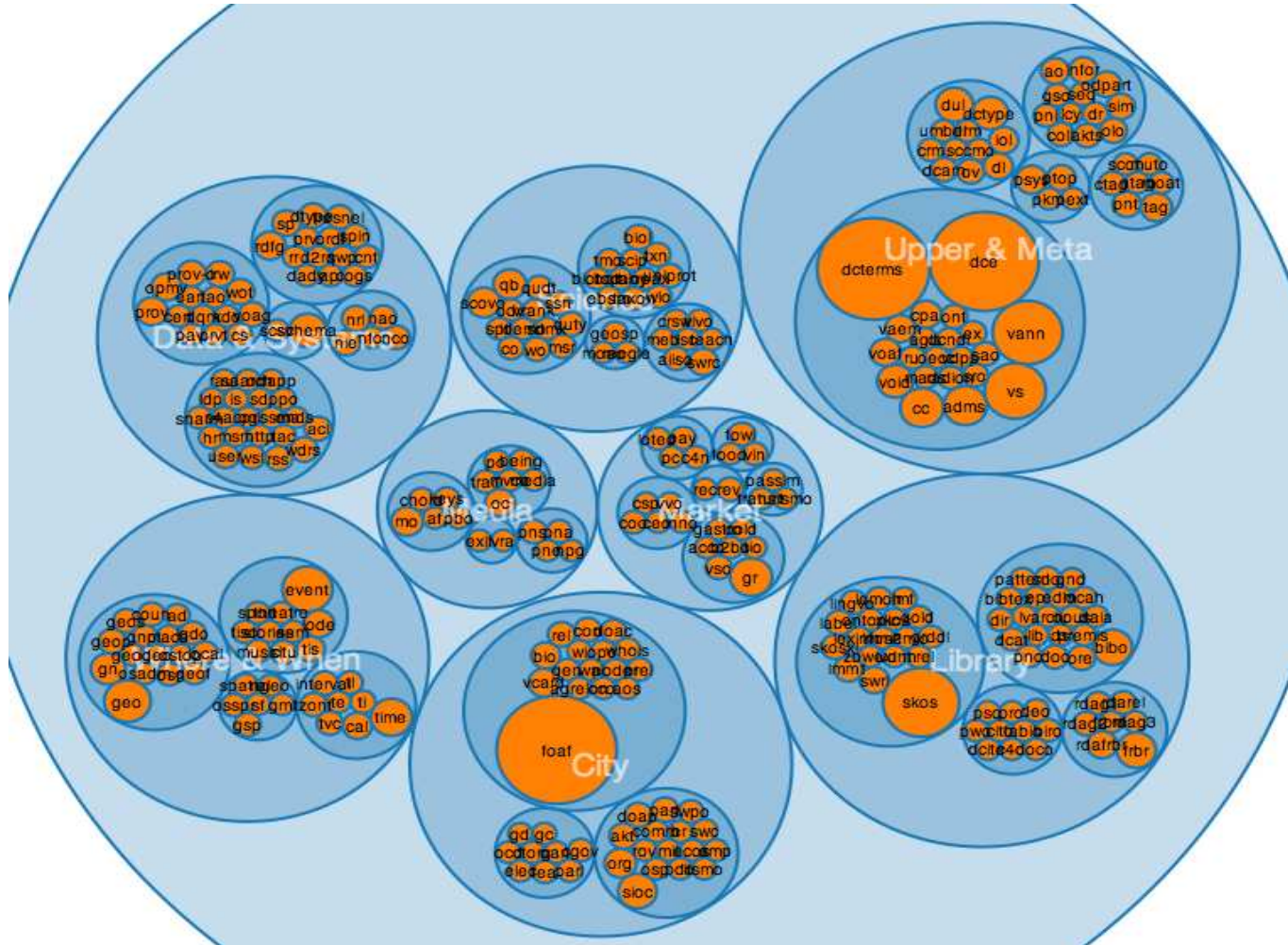
Situation Analysis – Ontologies

- Definition Conflicts
- Creation Conflicts
 - Notations
- Publication Conflicts
 - Formats
- Discovery Conflicts
- Usage Conflicts





Linked Open Vocabularies (LOV) Cloud





Linked Open Data Creation

- Observe
- Describe Observations in a Document
 - Entity Types (TBox)
 - Relation Types (RBox)
 - Entity Relationships (ABox)
- Publish Document



Shared Ontology Creation

- Know a Domain (or World)
- Describe Domain in a Linked Open Data Document
 - Entity Types (TBox)
 - Relation Types (RBox)
- Publish Document to the Web



Order of Operations?

- Build Ontology First?
- Create & Publish Linked Open Data First?
- A bit of both?





Problem – LOD Document Creator

- Procrastination due to quest for perfect ontology
- Inability to comprehend perfect ontology
 - Due representation notation
 - Relation semantics



Ontology Creator

- Terminology Disputes
- Relation Semantics Disputes
- Representation Notation Disputes
- Discovery Challenges in regards to related Ontologies.



Ontology Documents Directory

Index of /DAV/home/kidehen/Public/Linked Data Documents/

Name	Date Modified	Size	Content Type	Owner	Group	Permissions
..						
Offers-Licenses-Products	2012-09-12 18:12:47			kidehen	kidehen	rw-r--rm
.DS_Store	2012-09-19 12:19:26	6 K	application/octet-stream	kidehen	kidehen	rw-r--r
._dbpedia_to_sumo_links.nt	2013-03-21 13:30:59	4 K	application/octet-stream	kidehen	kidehen	rw-r--rm
._fao_to_sumo_mappings.txt	2013-03-21 13:30:53	4 K	text/plain	kidehen	kidehen	rw-r--rm
._sumo_to_fao_mappings.txt	2013-03-21 13:30:56	4 K	text/plain	kidehen	kidehen	rw-r--rm
MyProfile.ttl	2013-01-28 17:56:26	3 K	text/turtle	kidehen	kidehen	rw-r--rm
dbpedia_to_fao_mapping.nt	2012-09-06 15:04:17	28 K	text/plain	kidehen	kidehen	rw-r--r
dbpedia_to_fao_mapping.nt,acl	2012-10-30 10:08:21	924 B	text/turtle	kidehen	kidehen	rw-r--r
dbpedia_to_sumo_links.nt	2013-03-21 13:30:59	16 K	application/octet-stream	kidehen	kidehen	rw-r--rm
fao_to_sumo_mappings.txt	2013-03-21 13:30:38	30 K	text/plain	kidehen	kidehen	rw-r--rm
foaf_addressbook.ttl	2013-01-16 17:37:10	445 B	text/turtle	kidehen	kidehen	rw-r--rm
ldp_to_sioc_mapping2.ttl	2013-01-29 10:17:31	2 K	text/turtle	kidehen	kidehen	rw-r--rm
martin-udogle-podcast-with-ifueko-omogui-okauru	2013-02-01 13:55:11	2 K	text/turtle	kidehen	kidehen	rw-r--rm
operating_systems.ttl	2013-01-04 10:27:04	899 K	text/turtle	kidehen	kidehen	rw-r--rm
opsys.rdf	2013-01-04 10:32:38	3 M	application/rdf+xml	kidehen	kidehen	rw-r--rm
opsys.ttl	2013-01-04 10:22:44	899 K	text/turtle	kidehen	kidehen	rw-r--r
personae.ttl	2013-01-03 12:49:15	1711 B	text/turtle	kidehen	kidehen	rw-r--r
profile.ttl	2013-01-08 17:32:24	2021 B	text/turtle	kidehen	kidehen	rw-r--rm
sumo_to_fao_mappings.txt	2013-03-21 13:30:55	30 K	text/plain	kidehen	kidehen	rw-r--rm
test.ttl	2013-01-09 19:22:41	2 K	text/turtle	kidehen	kidehen	rw-r--rm
vcard_addressbook.ttl	2013-01-16 17:50:11	492 B	text/turtle	kidehen	kidehen	rw-r--rm
vcard_foaf_mapping.ttl	2013-01-16 17:42:26	815 B	text/turtle	kidehen	kidehen	rw-r--rm



Class Equivalence Reasoning

foaf:Organization Class Description (LOD Cloud)

OPENLINK
SOFTWARE

[Facets \(new session\)](#)
[Description](#)
[Metadata](#)
[Settings](#)

About: Organization [Sponge](#) [Permalink](#)
 An Entity of Type : [rdfs:Class](#), within Data Space : [lod.openlinksw.com](#) associated with source [dataset\(s\)](#)

Attributes	Values
type	Class
equivalentClass	Organization Organization
is type of	American Council on Renewable Energy European Biomass Association African Wind Energy Association Associação Portuguesa da Indústria Solar Portuguese Association of Independent Renewable Energy Electricity Producers »more»

Page 1 of 35



Faceted Search & Find service v1.13.53
 Alternative Linked Data Views: [iSPARQL](#) | [ODE](#) | Raw Data in: [CXML](#) | [CSV](#) | [RDF](#) ([N-Triples](#) [N3/Turtle](#) [JSON XML](#)) | [OData](#) ([Atom](#) [JSON](#)) | [Microdata](#) ([JSON HTML](#)) | [JSON-LD](#) | [About](#)



OpenLink Virtuoso version 07.10.3207 as of Feb 5 2014, on Linux (x86_64-redhat-linux-gnu), Cluster Edition(4 server processes, 756 GB total memory)
 Data on this page belongs to its respective rights holders.
 Virtuoso Faceted Browser Copyright © 2009-2013 OpenLink Software





Class Equivalence Reasoning

schema.org Organization Class Description – No Inference rules

OPENLINK SOFTWARE

Facets (new session) Description Metadata Settings

About: Organization [Sponge](#) [Permalink](#)
 An Entity of Type : [owl:Thing](#), within Data Space : [lod.openlinksw.com](#) associated with source [dataset\(s\)](#)

Attributes	Values
is equivalentClass of	Organization Organization
is first of	nodeID://b2624214010 nodeID://b2624214028 nodeID://b2624214035 nodeID://b2624214038 nodeID://b2624214067 »more»

« First « Prev Next » Last » Page 1 of 3 1 Go



Faceted Search & Find service v1.13.53
 Alternative Linked Data Views: [iSPARQL](#) | [ODE](#) Raw Data in: [CXML](#) | [CSV](#) | [RDF \(N-Triples\)](#) | [N3/Turtle](#) | [JSON XML](#) | [OData \(Atom JSON\)](#) | [Microdata \(JSON HTML\)](#) | [JSON-LD](#) [About](#)



OpenLink Virtuoso version 07.10.3207 as of Feb 5 2014, on Linux (x86_64-redhat-linux-gnu), Cluster Edition(4 server processes, 756 GB total memory)

Data on this page belongs to its respective rights holders.

Virtuoso Faceted Browser Copyright © 2009-2013 OpenLink Software





Class Equivalence Reasoning

schema.org Organization instances -- Inference rules enabled

The screenshot displays the OpenLink Software interface. At the top, there is a navigation bar with tabs for "Facets (new session)", "Description", "Metadata", and "Settings". The "Description" tab is active. Below the navigation bar, the main content area shows the "About: Organization" section. It includes a "Sponge" icon, a "Permalink" link, and a description: "An Entity of Type : owl:Thing, within Data Space : lod.openlinksw.com associated with source dataset(s)".

Below the "About" section, there is a section titled "Attributes Values" with the label "is type of". This section lists numerous organizations, including:

- American Council on Renewable Energy
- European Biomass Association
- African Wind Energy Association
- Associação Portuguesa da Indústria Solar
- Portuguese Association of Independent Renewable Energy Electricity Producers
- American Wind Energy Association
- British Association for Biofuels and Oils
- German Renewable Energy Association
- Bioenergy Australia
- Renewable Energy Association
- German Solar Industry Association
- German Windenergy Association
- British Wind Energy Association
- Canadian Wind Energy Association
- Club for Rural Electrification
- China Renewable Energy Industries Association
- Chinese Wind Energy Association
- Vindmølleindustrien
- Danmarks Vindmølleforening
- European Photovoltaic Industry Association
- European Renewable Energy Council
- European Renewable Energies Federation
- European Small Hydropower Association
- European Solar Thermal Industry Federation
- European Biomass Industry Association
- European Insulation Manufacturers Association
- European Copper Institute
- Finnish Wind Power Association
- Geothermal Association (Geothermische Vereinigung)
- Global Wind Energy Council
- Association HESPUJ
- Austrian Wind Energy Association
- International Geothermal Association

On the right side of the interface, there is a "QRCode" window displaying a QR code.

Class Equivalence Reasoning

schema.org Offers instances -- Inference rules enabled

About: **Offering**

[Sponge](#) [Permalink](#)

An Entity of Type : [owl:Class](#), within Data Space : [lod.openlinksw.com](#) associated with source [dataset\(s\)](#)

[New Facets Session with This Class](#)

An offering represents the public, not necessarily binding, not necessarily exclusive, announcement by a `gr:BusinessEntity` to provide (or seek) a certain `gr:BusinessFunction` for a certain `gr:ProductOrService` to a specified target audience. An offering is specified by the type of product or service or bundle it refers to, what business function is being offered (sales, rental, ...), and a set of commercial properties. It can either refer to (1) a clearly specified instance (`gr:Individual`), (2) to a set of anonymous instances of a given type (`gr:SomeItems`), (3) a product model specification (`gr:ProductOrServiceModel`), see also section 3.3.3 of the GoodRelations Technical Report. An offering may be constrained in terms of the eligible type of business partner, countries, quantities, and other commercial properties. The definition of the commercial properties, the type of product offered, and the business function are explained in other parts of this vocabulary in more detail. Example: Peter Miller offers to repair TV sets made by Siemens, Volkswagen Innsbruck sells a particular instance of a Volkswagen Golf at \$10,000. Compatibility with schema.org: This class is a superclass to <http://schema.org/Offer>, since `gr:Offering` can also represent demand.

Attributes Values

type	Class
equivalentClass	Offer Offer
disjointWith	Business entity Product or service Price specification Type and quantity node Quantitative value »more»
isDefinedBy	GoodRelations Ontology
label	
comment	
is type of	Pilsner Glasses Brushed Stainless Beverage Dispenser

QRCode





Ontology Problem Resolution

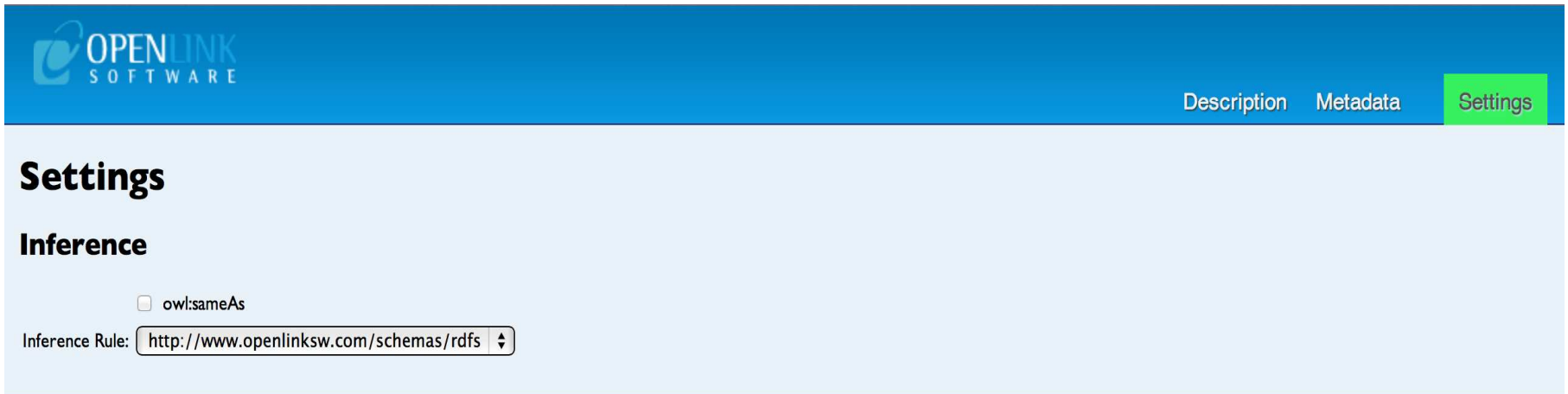
Use Inference Rules, conditionally





Using Inference Rules, Conditionally

Enabling or Disabling a specific Inference Rule



The screenshot shows the OpenLink Software interface. At the top left is the OpenLink Software logo. On the right, there are three tabs: "Description", "Metadata", and "Settings" (which is highlighted in green). Below the tabs, the page title is "Settings". Underneath, the section "Inference" is visible. There is a checkbox labeled "owl:sameAs" which is currently unchecked. Below this, there is a label "Inference Rule:" followed by a text input field containing the URL "http://www.openlinksw.com/schemas/rdfs" and a dropdown arrow icon.

Faceted Search & Find service v1.13.53

OpenLink Virtuoso version 07.10.3207 as of Feb 27 2014, on Linux (x86_64-redhat-linux-gnu), Cluster Edition(4 server processes, 756 GB total memory)
Copyright © 2009-2013 OpenLink Software | [About](#)





Additional Information Links

■ An [Glossary of terms](#) that includes term descriptions for:

- [Data](#)

- [Big Data](#)

- [Open Data](#)

- [Public Open Data](#)

- [Linked Data](#)

- [Linked Open Data](#)

- [Semantic Web](#)

- [Resource Description Framework \(RDF\)](#)

Additional Information

■ OpenLink Software:

- [OpenLink Software](#)
- [OpenLink Virtuoso](#)
- [Universal Data Access Drivers](#)

■ Social Media Data spaces

- <http://www.linkedin.com/in/kidehen> (LinkedIn)
- <https://plus.google.com/112399767740508618350/about> (Google+)
- <https://twitter.com/kidehen> (Twitter)
- Hashtag: #LinkedData (Anywhere) .