

Social Semantics and Linked Data for improving access to financial data on the Web

Alexandre Passant¹, Michael Hausenblas¹, Sean O'Riain¹,
Kingsley Idehen³, John G. Breslin^{1,2}, Stefan Decker¹

¹ *DERI, NUI Galway, Ireland*

² *School of Engineering and Informatics, NUI Galway, Ireland*

³ *OpenLink Software, Burlington, MA, USA*

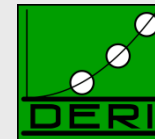
W3C Workshop on Improving Access to
Financial Data on the Web

October 2009, Arlington, Virginia USA



- Where we were coming from
- Financial Linked Data (FLD)
 - Applying LD principles to Financial Data
 - What's needed, enabling technologies
 - Linking Financial data with existing Linked Data
- Social Web and Financial Data
 - Current issues and proposal
 - Social Semantics and lightweight Ontologies
 - Benefits and use-cases for data transparency
- Discussion Topics

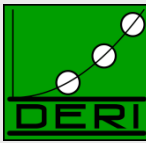
Where were coming from



- *“Toxic Assets Were Hidden Assets”*
 - Hernando De Soto, Wall Street Journal
 - <http://online.wsj.com/article/SB1237938113981320>
- A need to make financial information
 - More transparent, interoperable, discoverable
- **Linked Data**
 - Financial Linked Data to enable transparent assets
 - Social Semantics to enable transparent discussions

- Applying Linked Data principles to financial data
 - Use URIs as names for ‘things’
 - Financial items will be named and referencable by any application, making financial data interoperable
 - Use HTTP URIs to look up those names
 - Each piece of financial data will be accessible from the Web, without requiring additional protocols or tools
 - When the URI is looked up, provide useful information, using W3C standards (RDF, SPARQL)
 - Provide a uniform way to access the financial information using these standards
 - Include links to other URIs so that they can discover more ‘things’
 - Opening data silos by interlinking FLD on a global scale

Technologies for enabling FLD



- Ontologies for representing FLD
 - XBRL Ontology, RDFized microformats ...
- Exporters and wrappers
 - OpenLink Virtuoso Sponger, microformats + GRDDL ...

About:

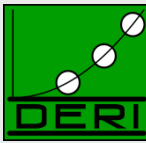
<http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0>

An Entity in Data Space: linkeddata.uriburner.com



Property	Value
dc:subject	<ul style="list-style-type: none">• http://www.rdfabout.com/rdf/usgov/secid/cik000038009
sioc:container_of	<ul style="list-style-type: none">• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/FinanceRevenue• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/InterestExpense• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/OtherNonOperatingIncome• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/OperatingExpenses• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/ProvisionLoanLeaseLosses
sioc:has_container	<ul style="list-style-type: none">• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml
opi-xbrl:endDate	<ul style="list-style-type: none">• 2008-12-31 (xsd:date)
opi-xbrl:identifier	<ul style="list-style-type: none">• 0000038009
opi-xbrl:scheme	<ul style="list-style-type: none">• http://www.sec.gov/CIK
opi-xbrl:startDate	<ul style="list-style-type: none">• 2008-01-01 (xsd:date)
rd:type	<ul style="list-style-type: none">• sioc:Container
is:sioc:container_of	<ul style="list-style-type: none">• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml
is:sioc:has_container_of	<ul style="list-style-type: none">• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/FinanceRevenue• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/InterestExpense• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/OtherNonOperatingIncome• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/OperatingExpenses• http://www.sec.gov/Archives/edgar/data/38009/000115752309002636/fmcc-20081231.xml#C0/ProvisionLoanLeaseLosses

FLD + Linking Open Data

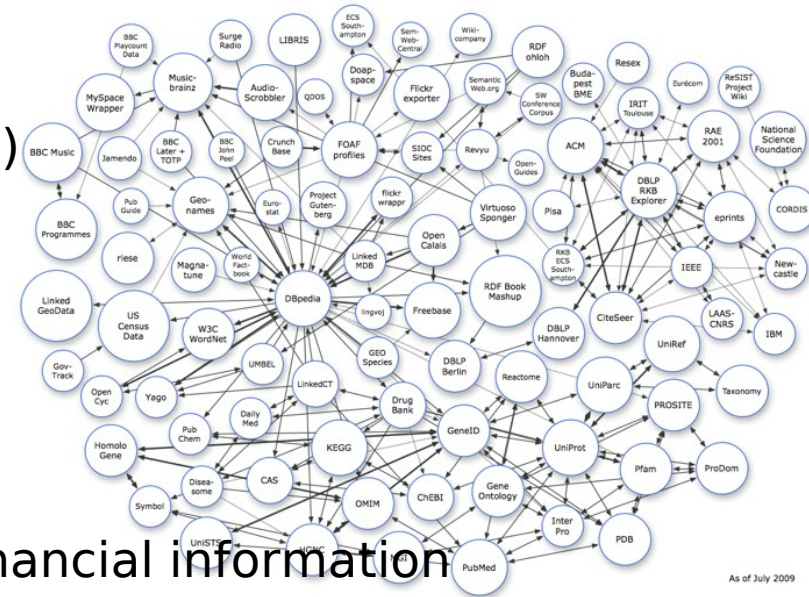


■ Interlinking FLD with existing resources from the LD cloud

- Location: Geonames
- Topics: Dbpedia
- Stats: EuroStat (riese; RDF enabled)

■ Advanced querying capabilities

- LD cloud linked with financial data and SPARQL
- Combining Geonames, riese and financial information



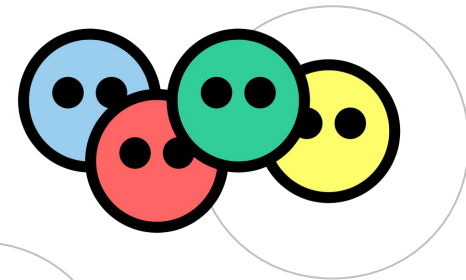
As of July 2009

Now possible to find: Comparison of budgets of Irish cities with a population between 70,000 and 100,000 inhabitants

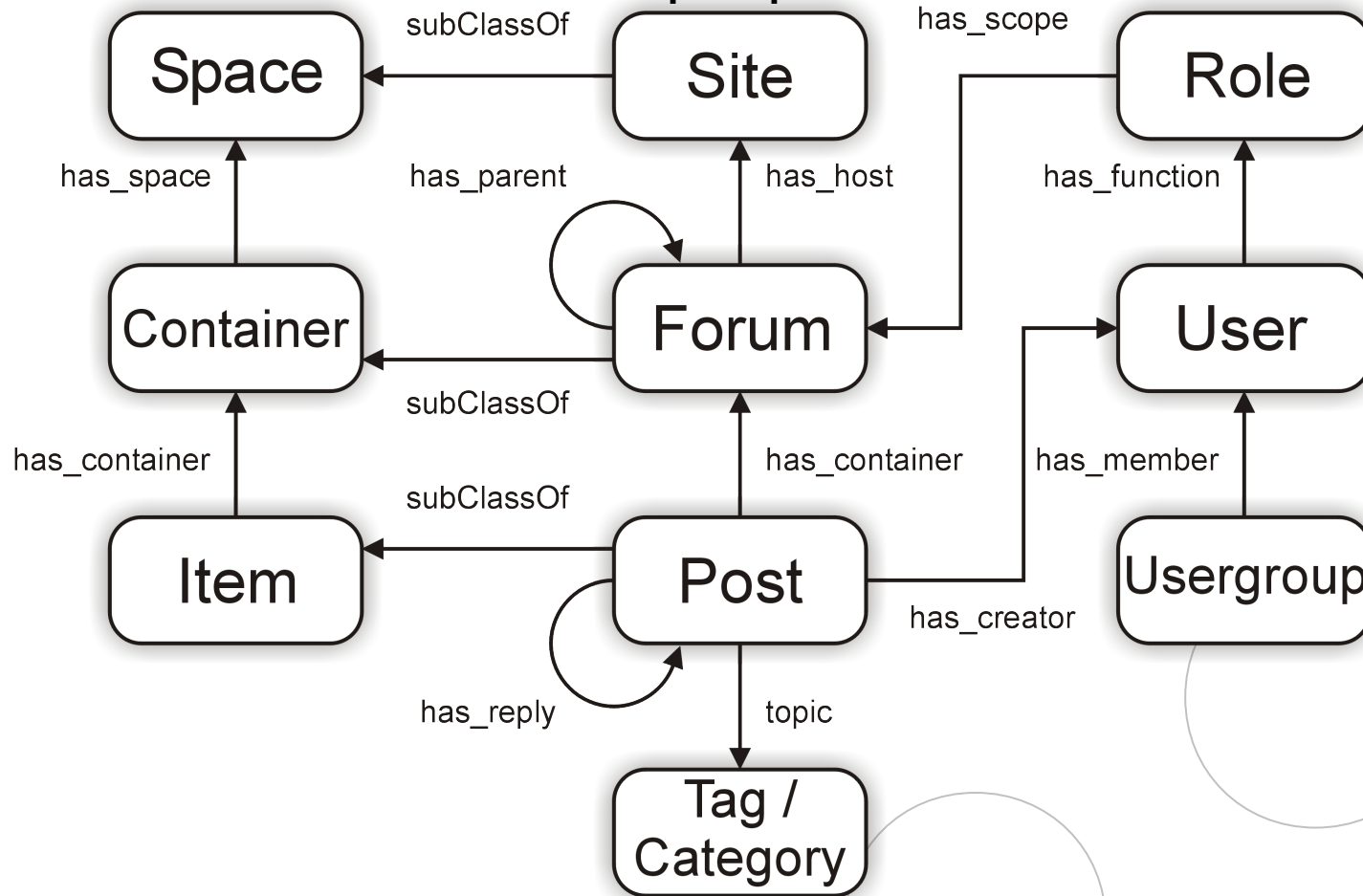
- **Social Web technologies (blogs, wikis, tweets ...)**
 - Can be used to gather discussions around financial data
 - Most remain as independent closed data silos
- **Opportunity for**
 - People to comment on and interact with the web FLD
 - e.g. Comments on financial data in grant proposals, budget reports
 - Applying LD principles to related discussions will allow it to be machine-readable
- **Providing**
 - Community insight into the underlying financial data being discussed

- Existing vocabularies to enable Semantics in Social Web applications
 - FOAF, used to represent personal identity and some characteristics of social networks
 - SIOC, a unified model for representing authors' contributions (blog posts, wiki pages) about financial data

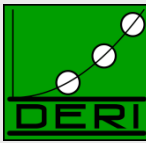
- Lightweight ontologies
 - Easy adoption and integration
 - Easy foster adoption on the Web
 - Community support via API's



- The main classes and properties are:



Sig.ma SIOC Widgets



From: SW Human Resources
To: Web careers (webcareers@..com)
Subject: Semantic Web job opportunity

We're currently looking for a Semantic Web consultant to help us developing the Sindice search engine.
Send your CV to the list and it will be reviewed ASAP by representatives.
Kind regards,
SW hiring team

From: Johns New Nickname
To: Web careers (webcareers@..com)
Subject: Recent activity by user "Cloud" [X]

I might

Posts

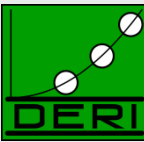
"It was a very good year..."
"SIOC tutorial accepted for WWW2008"
"boards.ie latest: hires and hiring; doubling visits since 2005"
"Videos of "Paddy's Valley" to VC pitches"
"Congrats to Eyal on his PhD defense"

Topics

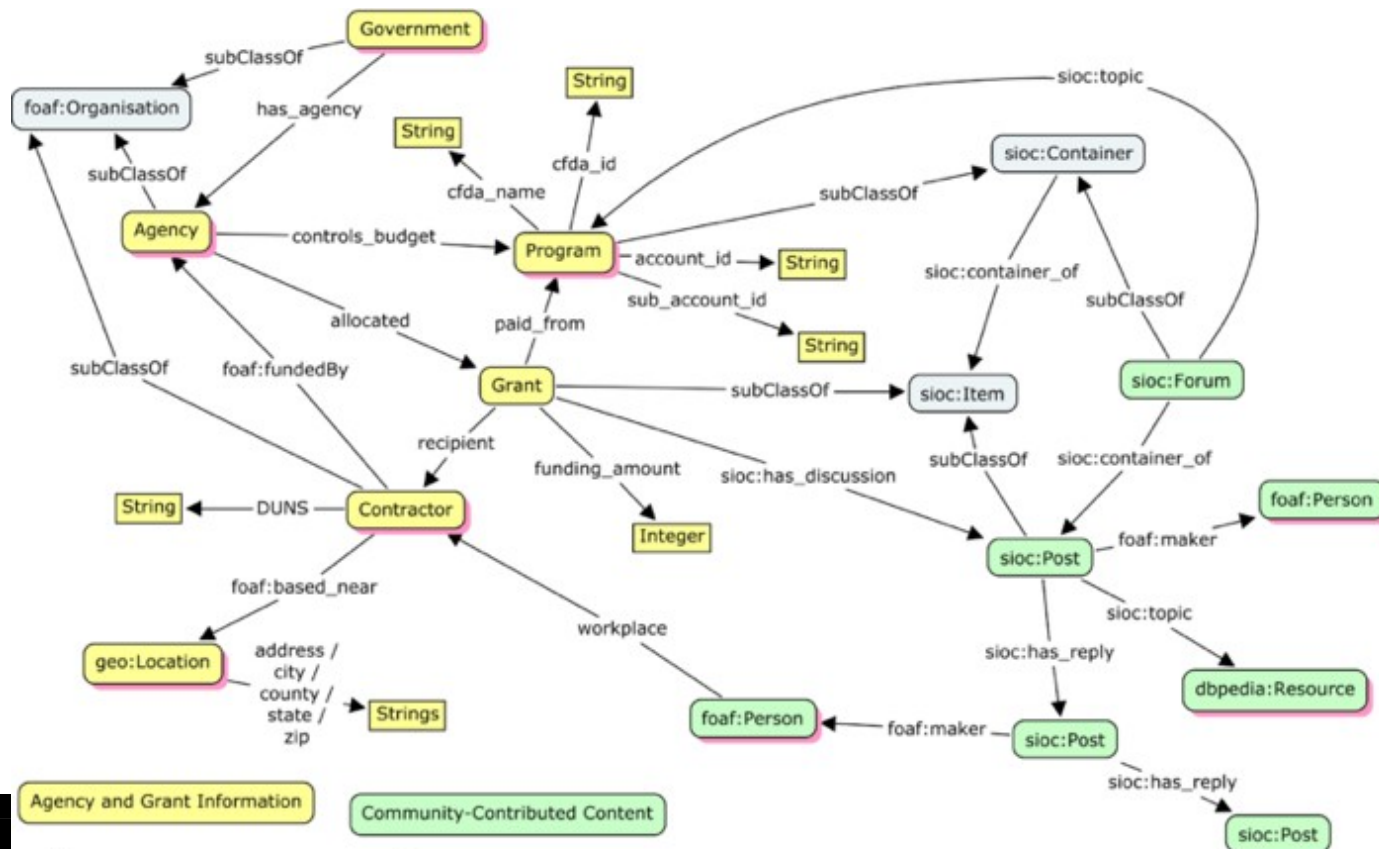
"Boards"
"DERI"
"NUI Galway"
"General"
"Semantic Web"
"SIOC"

From:
To: Web careers (webcareers@..com)
Subject:

Social Semantics and FLD



- Linking agencies, grants, etc. to related discussions using ...



Agency and Grant Information

Community-Contributed Content

Object with Potential for Linked Data

■ Identifying authors

- Identify particular contribution from authors
 - employee of a certain funding agency, authenticated on service X (using FOAF+SSL or OpenID)
- A first level of trust with respect to related financial discussions

■ Identifying contributions across communities

- Identify related discussion threads and people involved on the Web by opening the data silos structures
- Giving an idea of the "hot topics" in a particular domain.

Discussion Topics



- Exploiting the community effect
 - At the conversation level
- Authoring for non technical users
- Identifiers, privacy, security and trust
- Applicability of linking existing LD datasets to financial information
 - Start small and build out
 - Analytics
- Getting Government to share what it has