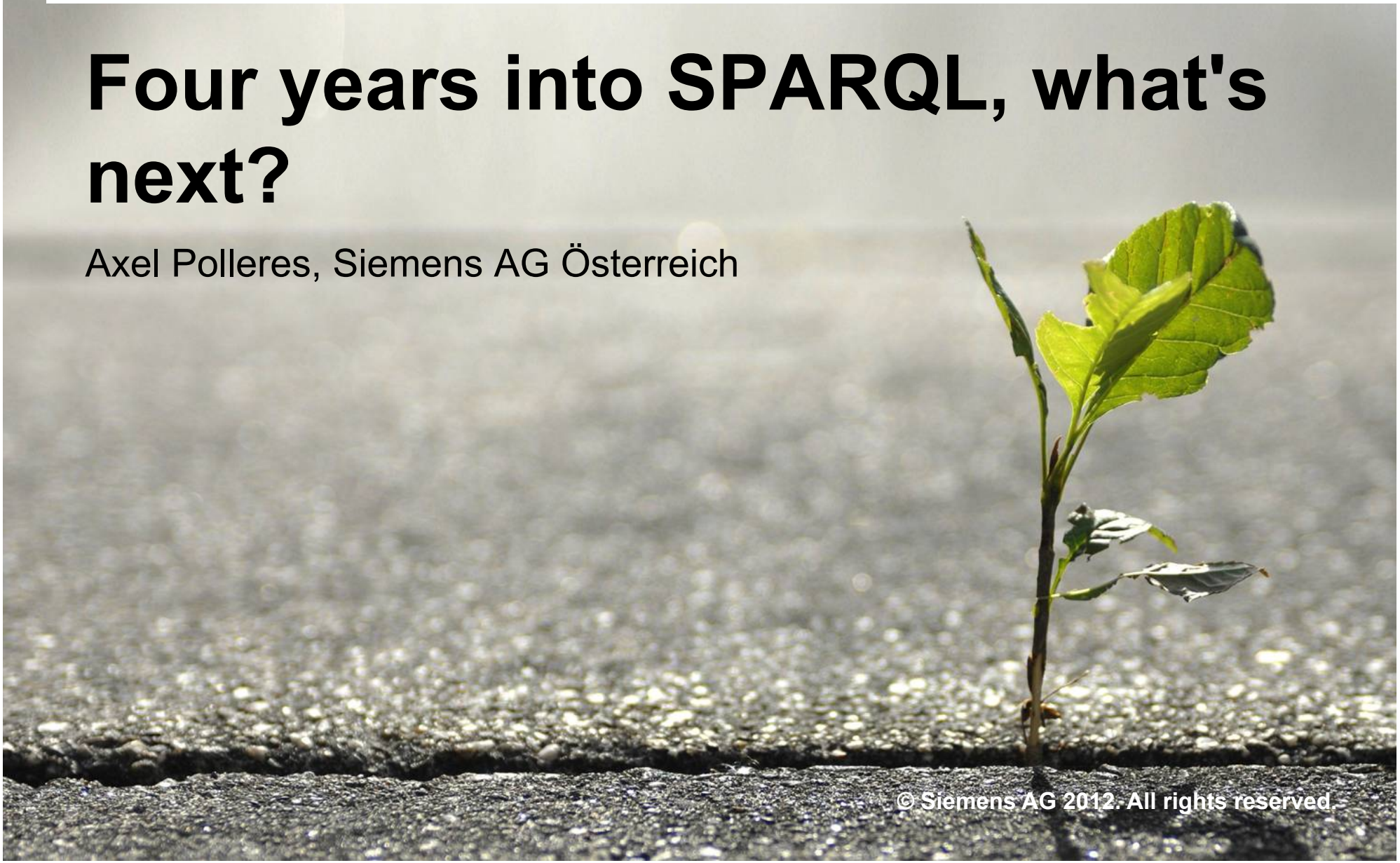
SIEMENS

PAAMS'12

Four years into SPARQL, what's next?

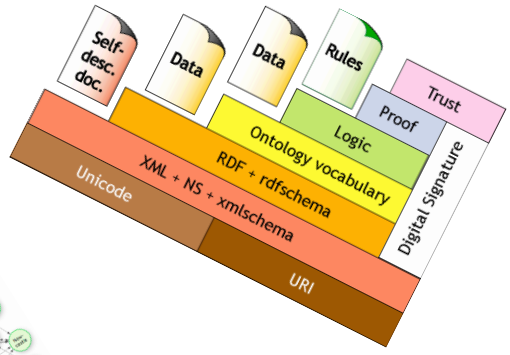
Axel Polleres, Siemens AG Österreich

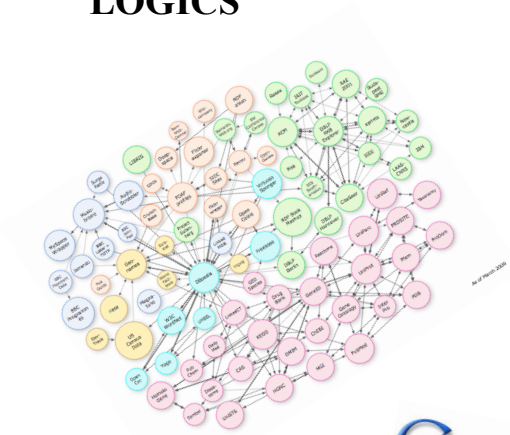


Introduction / Contents

What have You heard about
"Semantic Web"?

DESCRIPTION LOGICS





Google?

Outline

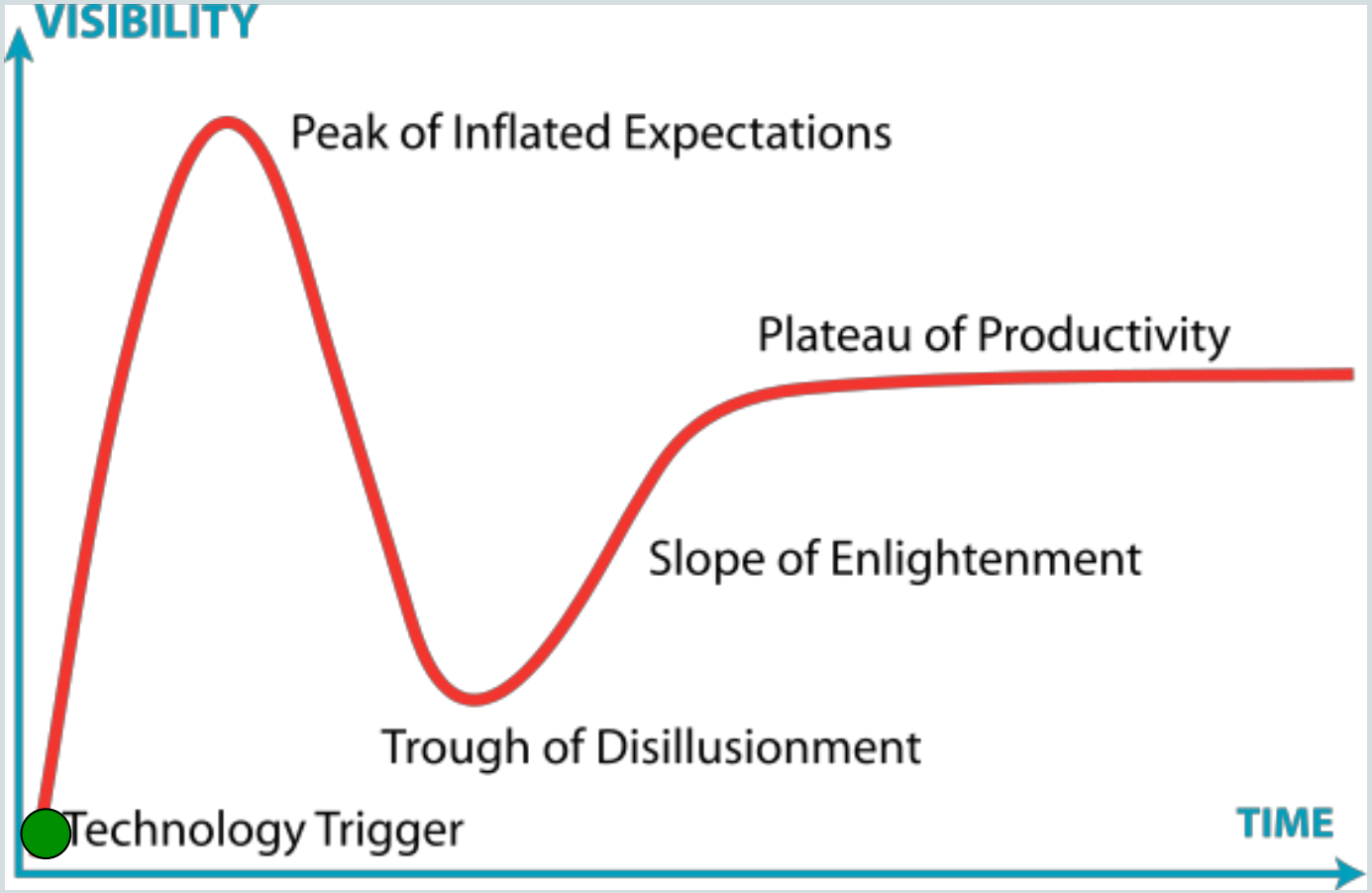
- Where is the "Semantic Web"?
- The role of Linked Data and SPARQL
- SPARQL1.1
- BTW, what about Ontologies?
- Why is this interesting for companies?
- Challenges ahead…

SIEMENS

Where is the Semantic Web?

SIEMENS

What happened over the last decade?



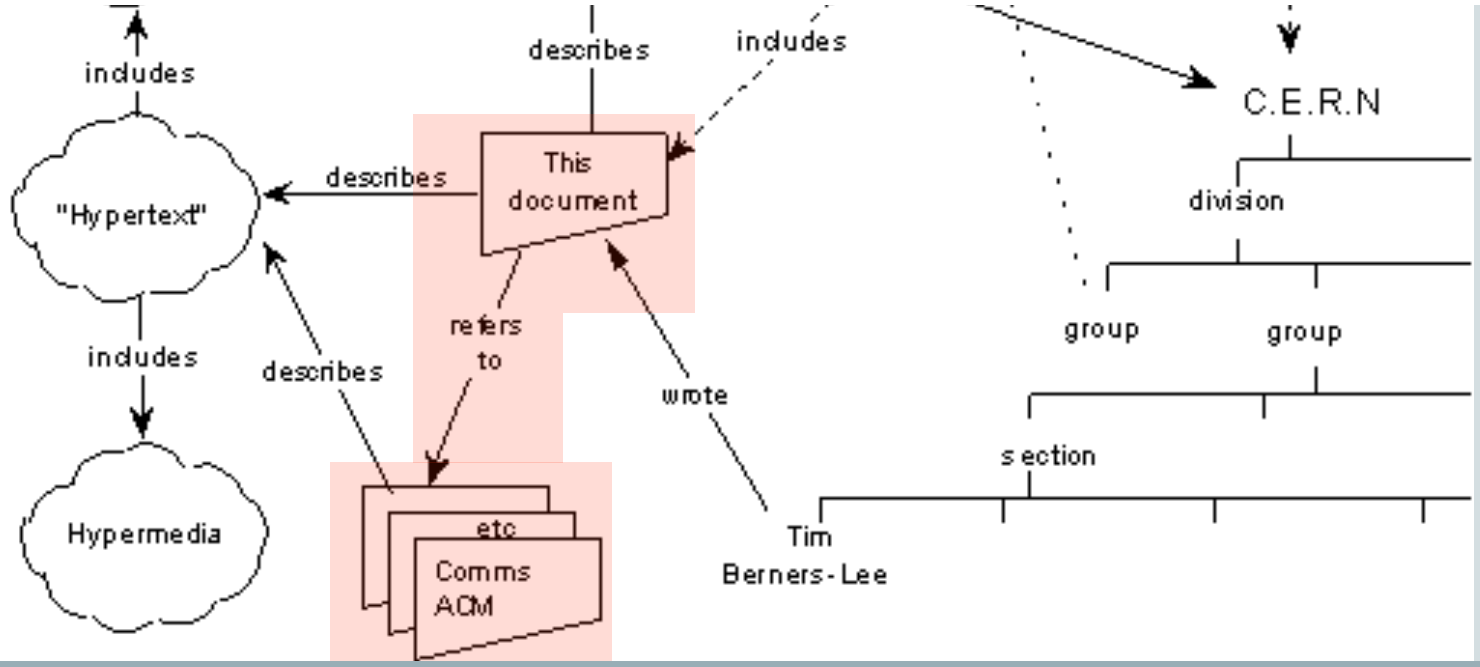
2001

- Tim Berners-Lee, James Hendler and Ora Lassila, "The Semantic Web", Scientific American, May 2001

Expressing Meaning Evolution of Knowledge
Knowledge Representation Agents Ontologies

*"This proposal concerns the **management of general information** about accelerators and experiments at CERN*

*[…] based on a **distributed hypertext system**. "*



I studied here

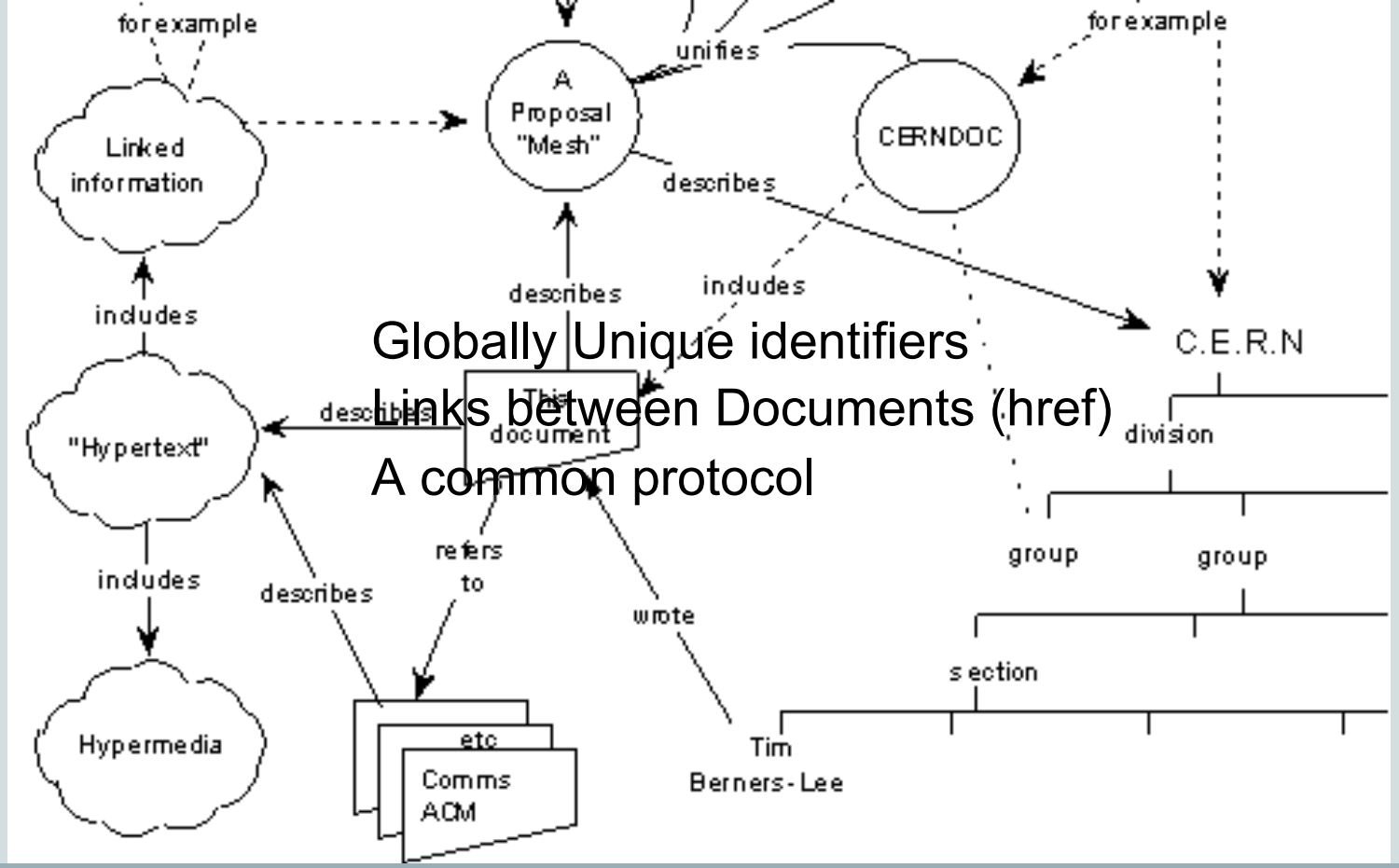
polleres.net#me → www.tuwien.ac.at

Document Document

Globally Unique identifiers — URIs

Links between Documents (href)

A common protocol — HTTP



Globally Unique identifiers

Links between Documents (href)

A common protocol

```
<p about="#me">I studied <a rel="foaf:schoolHomepage" href="http://www.tuwien.ac">here</a></p>
```

polleres.net#me —xmlns.com/foaf/0.1/schoolHomepage→ www.tuwien.ac.at

Person — Document

Globally Unique identifiers

Typed Links between **Entities**

A common protocol

URIs

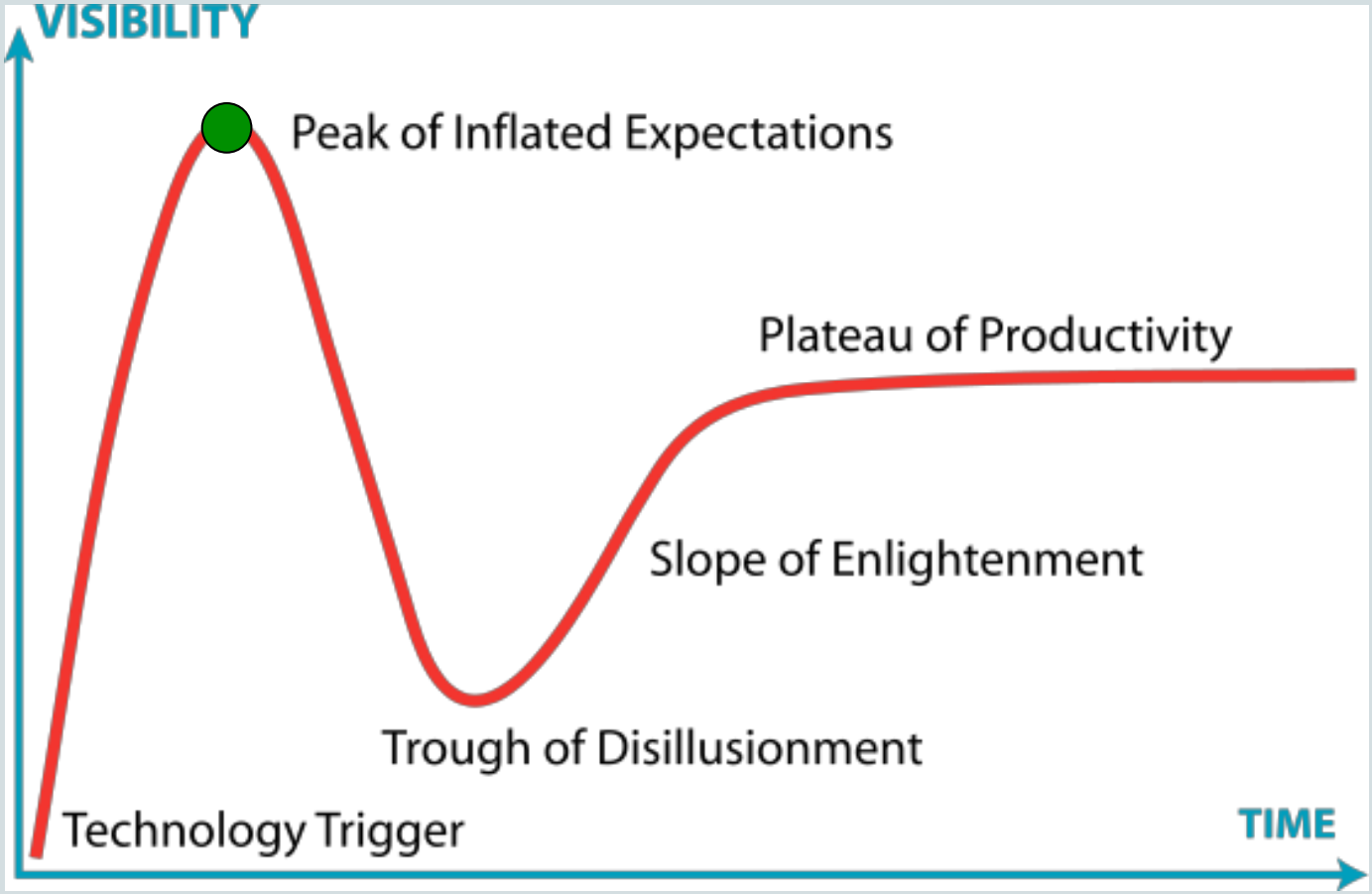
RDF

HTTP

Expressing Meaning

SIEMENS

What happened over the last decade?



2004-2005

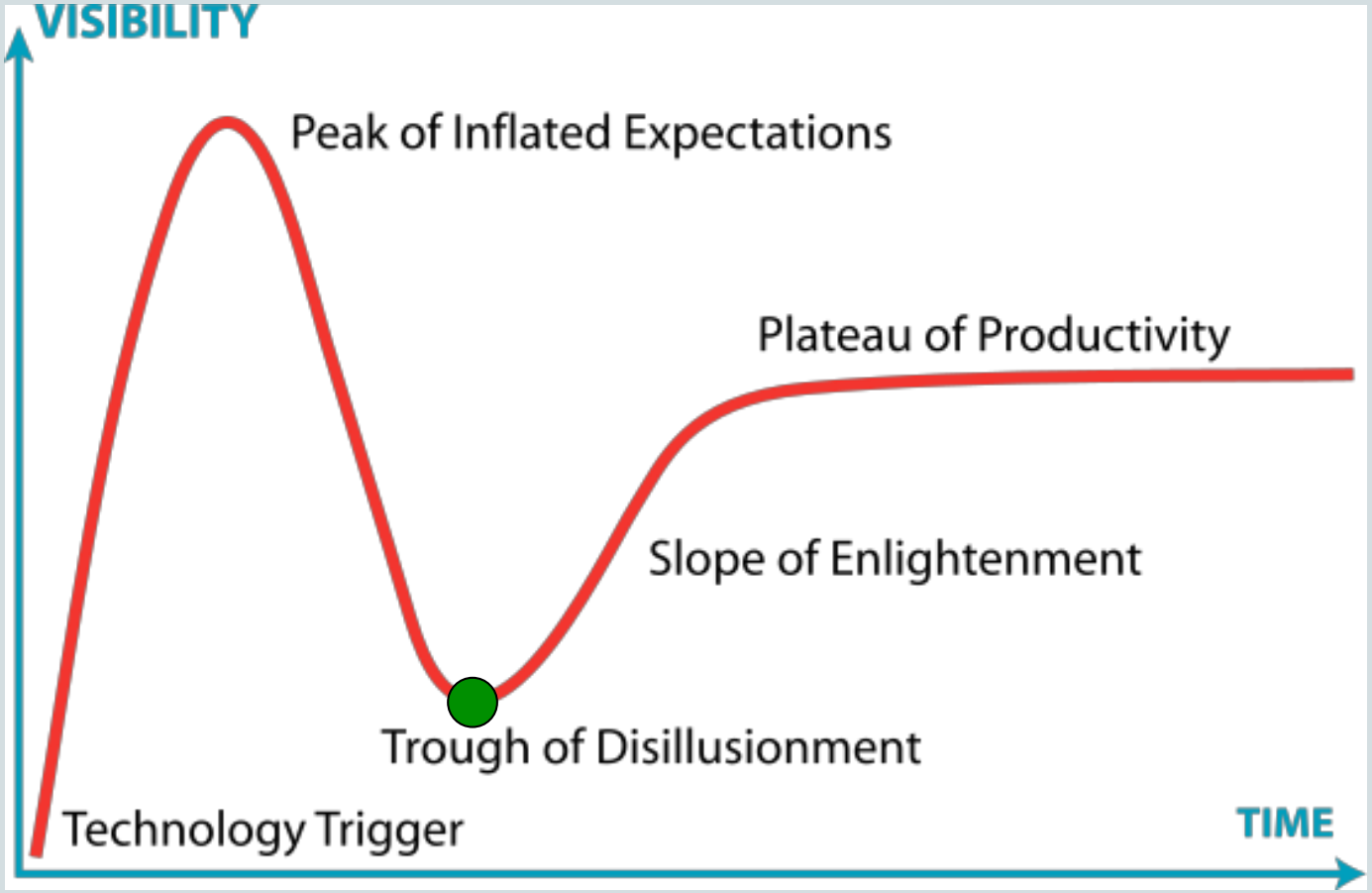
OWL (2004), Semantic Web Services: OWL-S (2004), WSMO (2005)

… two steps ahead of time?

Agents Ontologies

SIEMENS

What happened over the last decade?

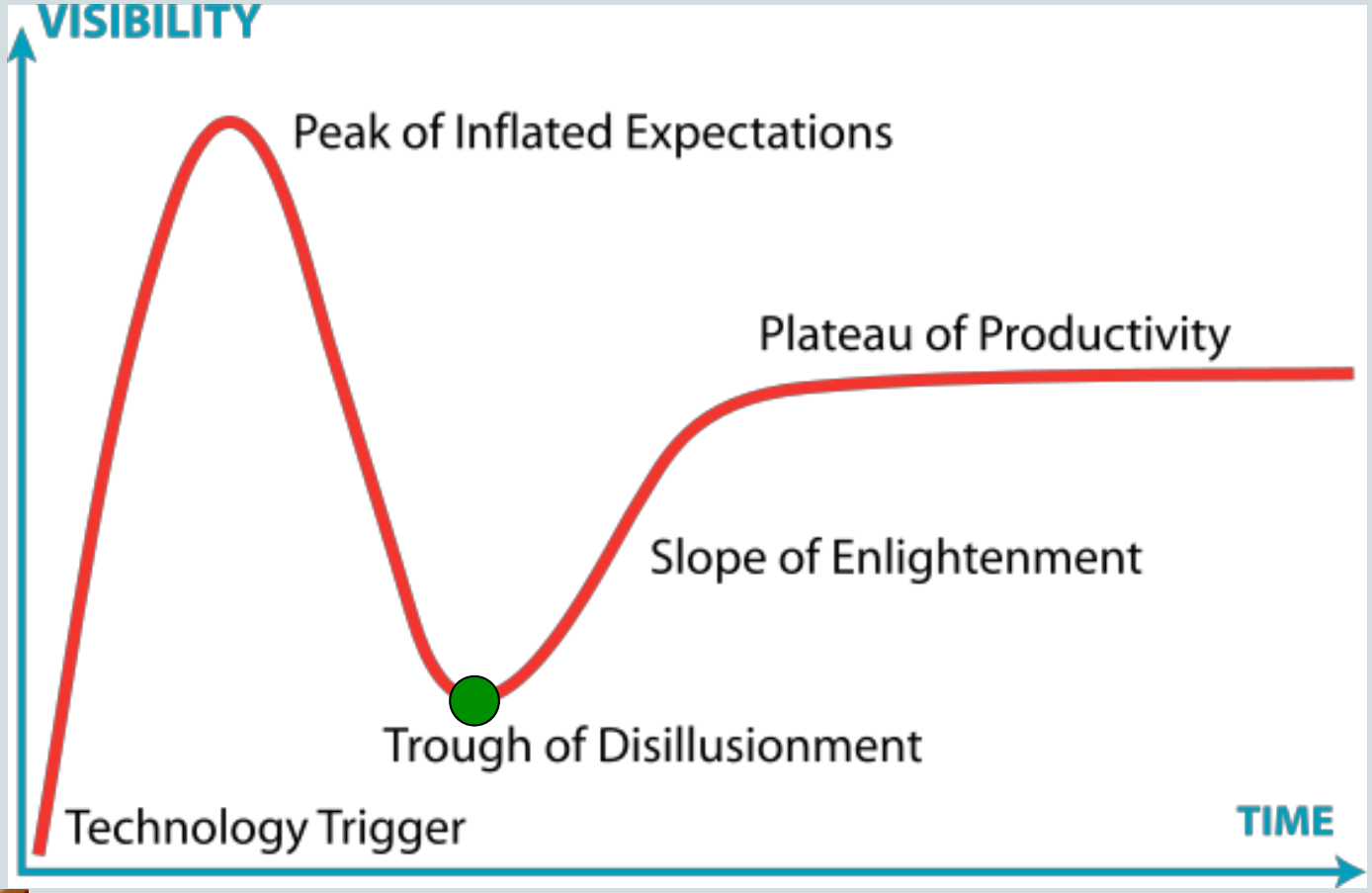


2006

- Problems:
  - 5 years into the idea, but no significant take-up
  - No agreed way to **publish** “Semantic Web” Data (apart from RDF/XML)
  - No query language to **consume** “Semantic Web” Data

SIEMENS

What happened over the last decade?



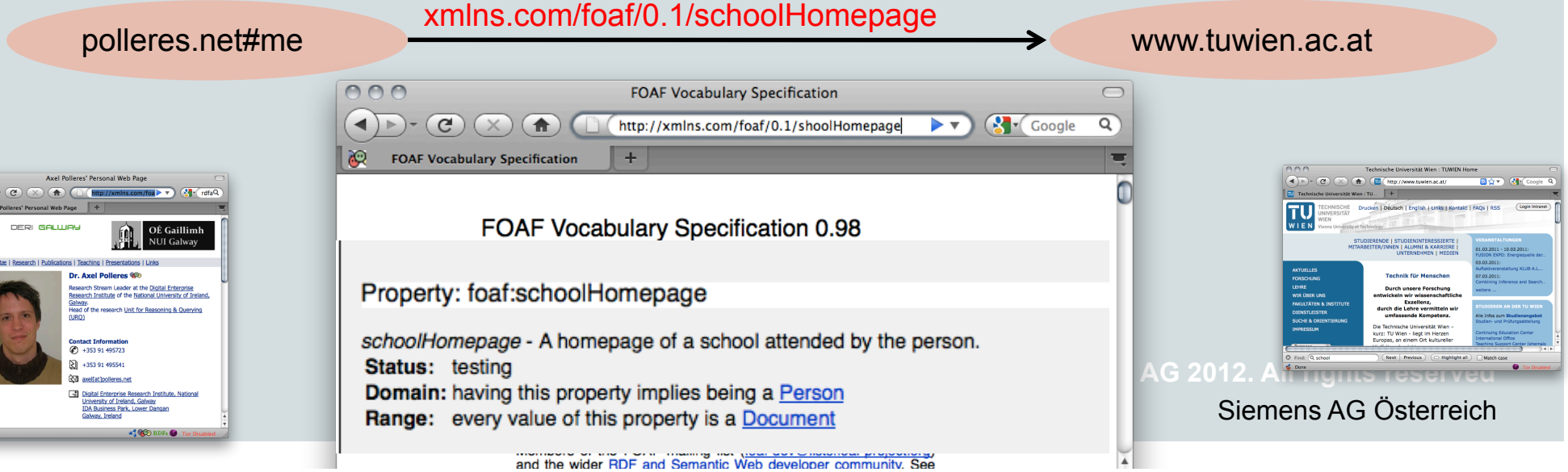
2006

http://www.w3.org/DesignIssues/LinkedData.html

SIEMENS

Linked Data Principles

1. Everything gets a URI (papers, people, talks, organizations, topics…)
2. These URIs are linked via RDF describing relations
3. Relations are URIs again (e.g. :name)
4. **When I dereference the URIs, I should find more information about them, defining them.**

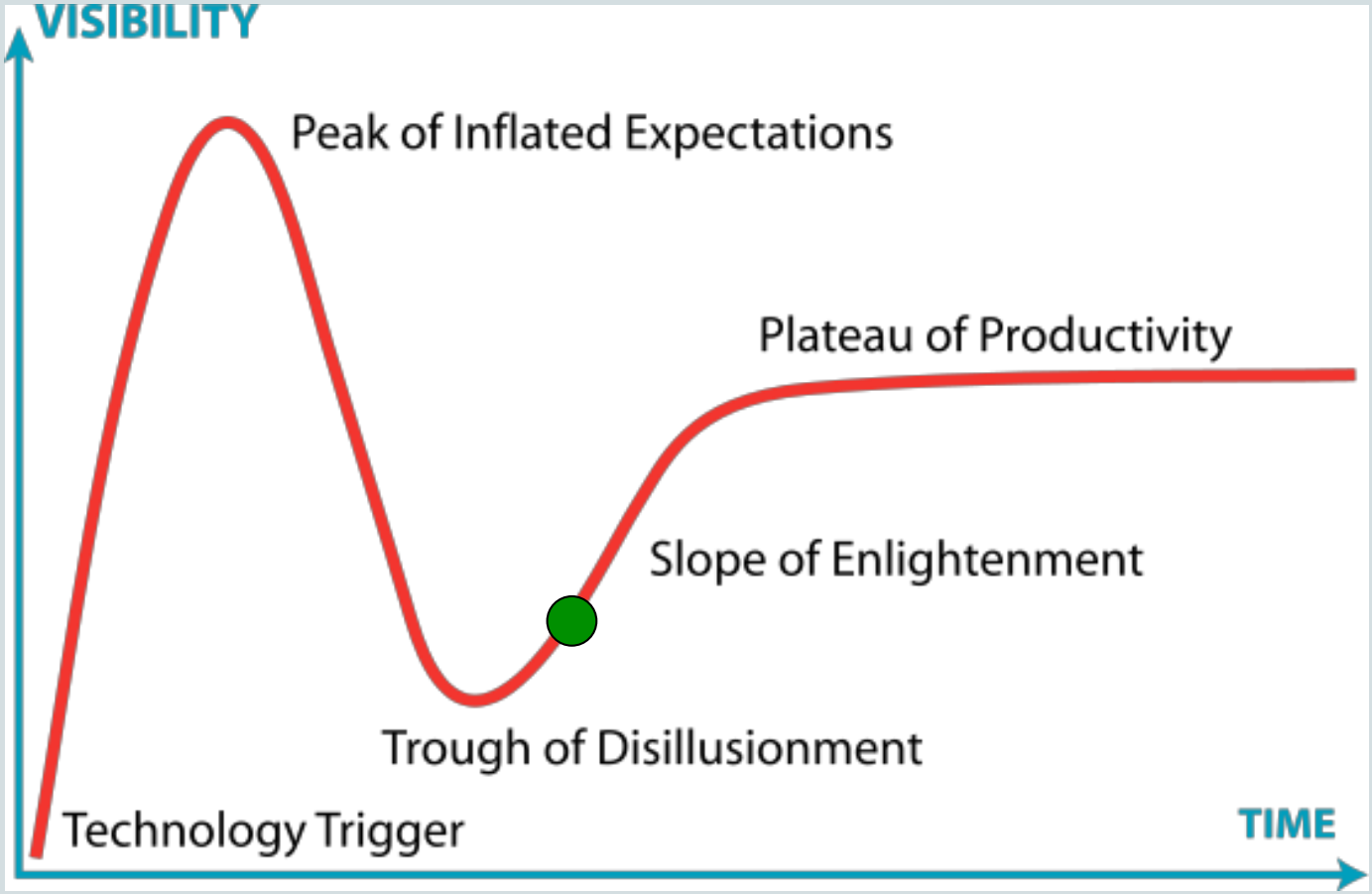


Linked Data – Adoption (since 2007)

SIEMENS

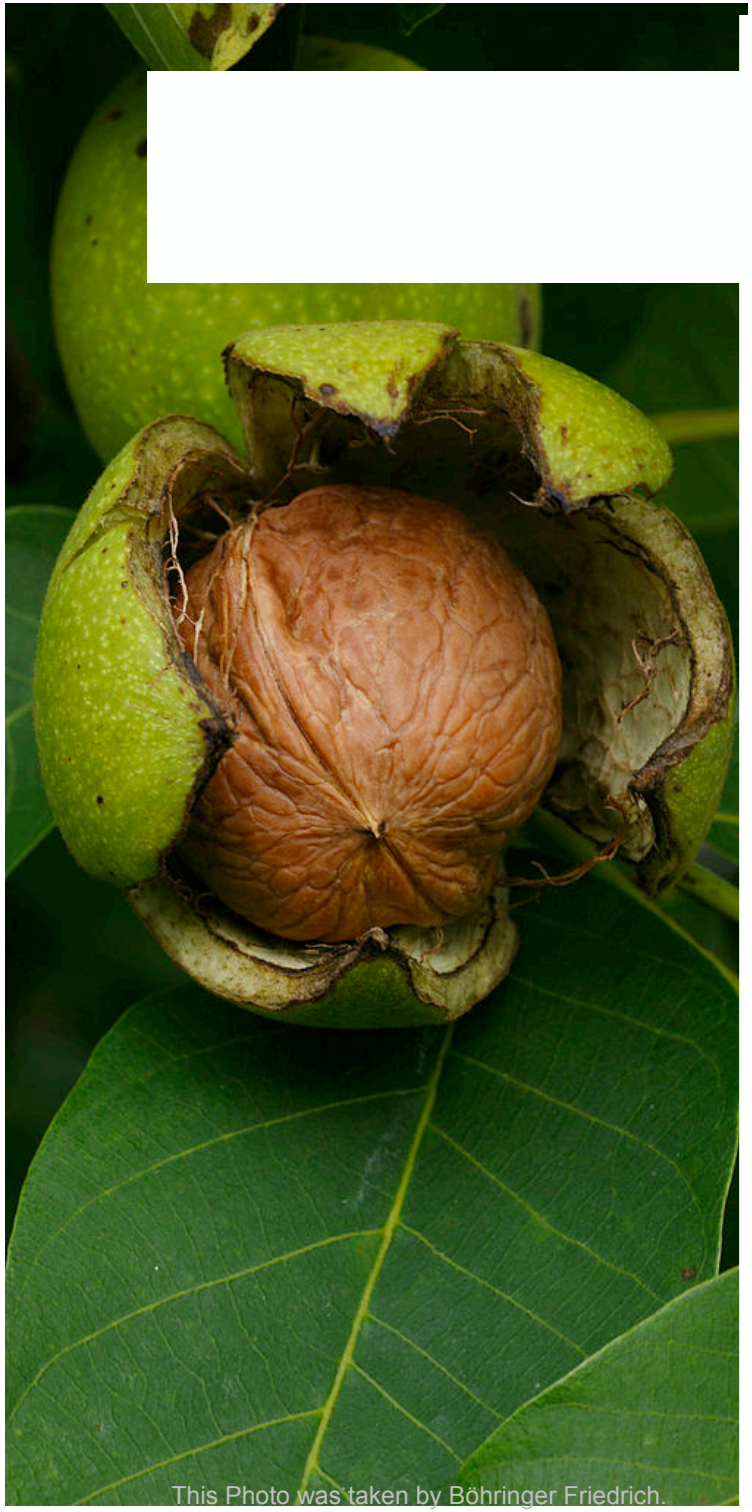
SIEMENS

What happened over the last decade?



2008

- Problems:
  - ~~5 years into the idea, but no significant take-up~~ → Linked Data
  - ~~No agreed way to **publish** "Semantic Web" Data~~ → RDFa (2008), Turtle (2008)
  - ~~No query language to **consume** "Semantic Web" Data~~ → **SPARQL** (2008)



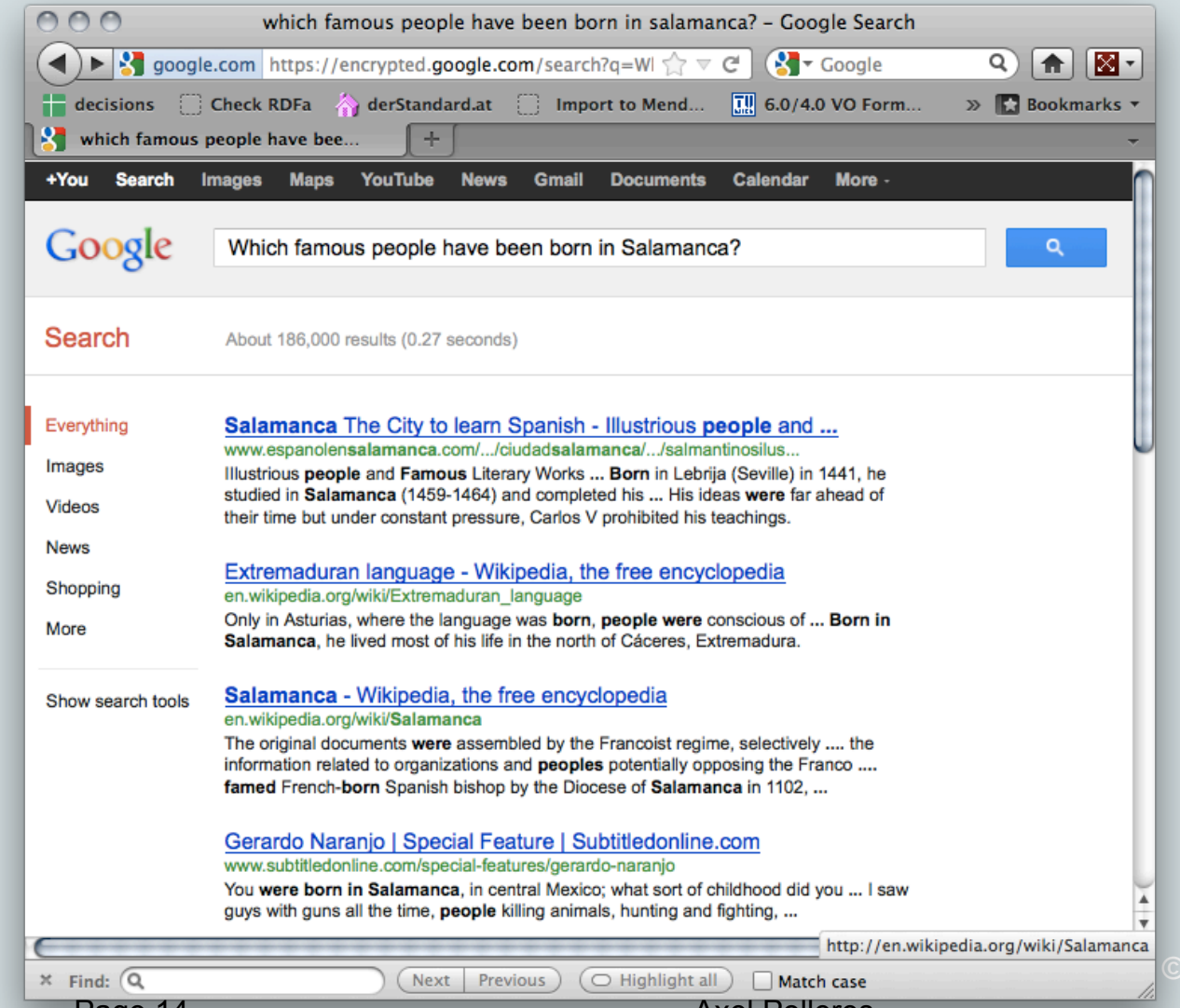
SIEMENS

SPARQL in a Nutshell...

SPARQL + Linked Data give you Semantic search almost “for free”

SIEMENS

Which famous people have been born in Salamanca?



SPARQL + Linked Data give you Semantic search almost “for free”

SIEMENS

Which famous people have been born in Salamanca?

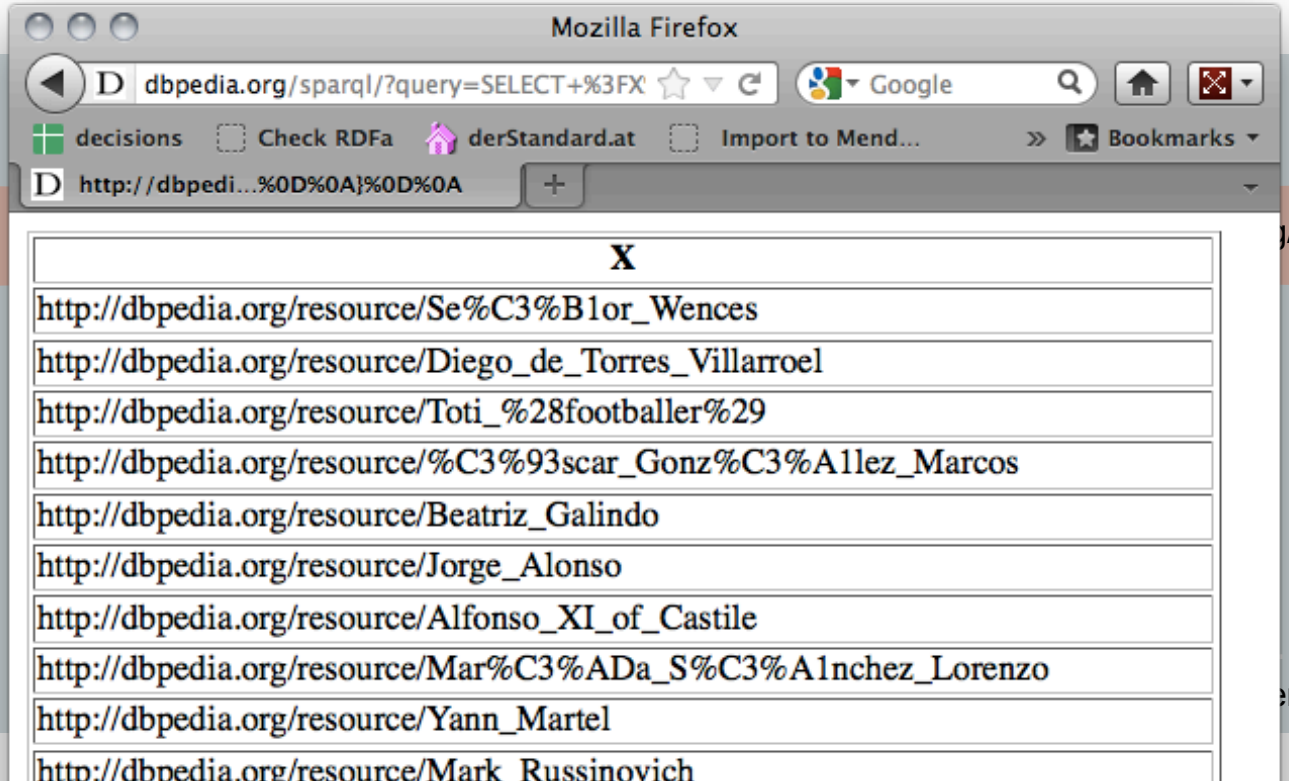
```
SELECT ?X
WHERE
{
 ?X  <http://dbpedia.org/property/birthPlace> <http://dbpedia.org/resource/Salamanca>
}
```

SPARQL + Linked Data give you Semantic search almost “for free”

SIEMENS

Which famous people have been born in Salamanca?

```
SELECT ?X
WHERE
{
 ?X  <http://dbpedia.org/property/birthPlace> <http://dbpedia.org/resource/Salamanca>
}
```



?X /Salamanca

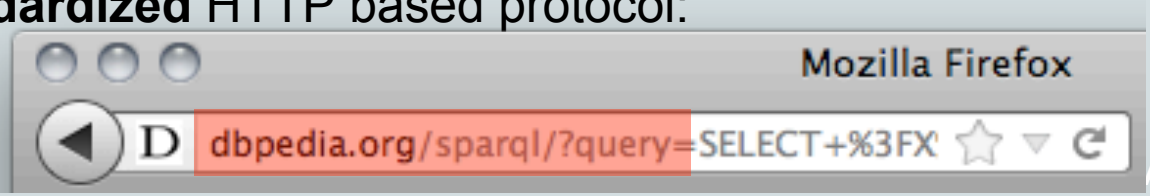
X
http://dbpedia.org/resource/Se%C3%B1or_Wences
http://dbpedia.org/resource/Diego_de_Torres_Villarroel
http://dbpedia.org/resource/Toti_%28footballer%29
http://dbpedia.org/resource/%C3%93scar_Gonz%C3%A1lez_Marcos
http://dbpedia.org/resource/Beatriz_Galindo
http://dbpedia.org/resource/Jorge_Alonso
http://dbpedia.org/resource/Alfonso_XI_of_Castile
http://dbpedia.org/resource/Mar%C3%ADa_S%C3%A1nchez_Lorenzo
http://dbpedia.org/resource/Yann_Martel
http://dbpedia.org/resource/Mark_Russinovich

SPARQL – Standard RDF Query Language and Protocol

SPARQL (2008):

```
SELECT ?X
WHERE
{
 ?X  <http://dbpedia.org/property/birthPlace> <http://dbpedia.org/resource/Salamanca>
}
```

- SQL “Look-and-feel” for the Web
- Essentially “graph matching” by *triple patterns*
- Allows conjunction (.) , disjunction (UNION), optional (OPTIONAL) patterns and filters (FILTER)
- Construct new RDF from existing RDF
- Solution modifiers (DISTINCT, ORDER BY, LIMIT, …)
- A **standardized** HTTP based protocol:



Conjunction (.) , disjunction (UNION), optional (OPTIONAL) patterns and filters (FILTER)

SIEMENS

***Names** of **soccer players** born **in Salamanca**?*

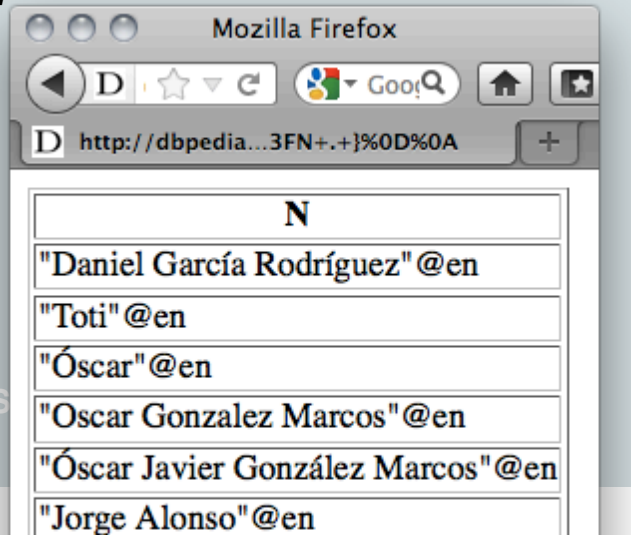
```
PREFIX : <http://dbpedia.org/resource/>
PREFIX dbprop: <http://dbpedia.org/property/>
PREFIX dbont: <http://dbpedia.org/ontology/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SELECT ?N

WHERE

{ ?X  dbprop:birthPlace :Salamanca ; a dbont:SoccerPlayer ; foaf:name ?N . }
```

- *Shortcuts for namespace prefixes and to group triple patterns*



Conjunction (.) , **disjunction (UNION)**, optional (OPTIONAL) patterns and filters (FILTER)

SIEMENS

***Names** of **soccer players** or **writers** born **in Salamanca**?*

```
PREFIX : <http://dbpedia.org/resource/>
PREFIX dbprop: <http://dbpedia.org/property/>
PREFIX dbont: <http://dbpedia.org/ontology/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
SELECT ?N

WHERE {

{ { ?X a dbont:SoccerPlayer } UNION  {?X a dbont:Writer } }

  ?X  dbprop:birthPlace :Salamanca ; ?X foaf:name ?N . }
```

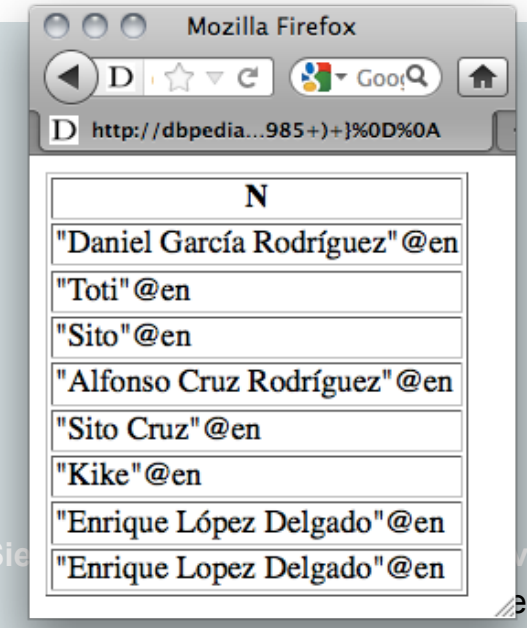
Conjunction (.) , disjunction (UNION), optional (OPTIONAL) patterns and **filters (FILTER)**

SIEMENS

***Names** of **soccer players** or born **after 1970**?*

```
PREFIX : <http://dbpedia.org/resource/>
PREFIX dbprop: <http://dbpedia.org/property/>
PREFIX dbont: <http://dbpedia.org/ontology/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SELECT  ?N
WHERE
{ ?X  dbprop:birthPlace :Salamanca ; a dbont:SoccerPlayer ; foaf:name ?N ;
  dbprop:birthDate ?D . FILTER( YEAR(?D) > 1985 ) }
```



Why we aren't done yet (and why SPARQL1.1 was needed)

SIEMENS

Based on implementation experience, in 2009 new W3C SPARQL WG founded to address common feature requirements requested urgently by the community: ***http://www.w3.org/2009/sparql/wiki/Main_Page***

1. **Negation**
2. **Assignment/Project Expressions**
3. **Aggregate functions (SUM, AVG, MIN, MAX, COUNT, …)**
4. **Subqueries**
5. **Property paths**
6. **Updates**
7. **Entailment Regimes**

- Other issues for wider usability:
  - Result formats (JSON, CSV, TSV),
  - Graph Store Protocol (REST operations on graph stores)
  - ***Goal: SPARQL 1.1 W3C Recommendation this summer***

SIEMENS

1. Negation: MINUS and NOT EXISTS

Select Persons without a homepage:

```
SELECT ?X
WHERE{ ?X rdf:type foaf:Person
       OPTIONAL { ?X foaf:homepage ?H }
       FILTER( !bound( ?H ) ) }
```

***Negation as failure** in SPARQL1.0 is "ugly":*
***SPARQL1.1** has two alternatives to do the same*

SIEMENS

1. Negation: MINUS and NOT EXISTS

Select Persons without a homepage:

```
SELECT ?X
WHERE{ ?X rdf:type foaf:Person
         FILTER ( NOT EXISTS { ?X foaf:homepage ?H } ) }
```

***Negation as failure** in SPARQL1.0 is "ugly":*
***SPARQL1.1** has two alternatives to do the same*

- *NOT EXISTS in FILTERs*
  - *detect non-existence*

SIEMENS

1. Negation: MINUS and NOT EXISTS

Select Persons without a homepage:

```
SELECT ?X
WHERE{ ?X rdf:type foaf:Person
       MINUS { ?X foaf:homepage ?H } ) }
```

***Negation as failure** in SPARQL1.0 is "ugly":*
***SPARQL1.1** has two alternatives to do the same*

- *NOT EXISTS in FILTERs*
  - *detect non-existence*
- *(P1 MINUS P2) as a new binary operator*
  - *"Remove rows with matching bindings"*
  - *only effective when P1 and P2 share variables*

SIEMENS

2. Assignment/Project Expressions

Assignments, Creating new values… not possible in SPARQL 1.0

```
PREFIX ex: <http://example.org/>
SELECT ?Item ?NewP
WHERE { ?Item ex:price ?Pr FILTER (?NewP = ?Pr * 1.1) }
```

Data:

```
@prefix ex: <http://example.org/> .

ex:lemonade1      ex:price 3 .
ex:beer1          ex:price 3 .
ex:wine1          ex:price 3.50 .
```

Results:

?Item	?NewP

2. Assignment/Project Expressions

Assignments, Creating new values…now works in SPARQL 1.1:

```
PREFIX ex: <http://example.org/>
SELECT ?Item (?Pr * 1.1 AS ?NewP )
WHERE { ?Item ex:price ?Pr }
```

Data:

```
@prefix ex: <http://example.org/> .

ex:lemonade1      ex:price 3 .
ex:beer1          ex:price 3 .
ex:wine1          ex:price 3.50 .
```

Results:

?Item	?NewP
lemonade1	3.3
beer1	3.3
wine1	3.85

SIEMENS

3. Aggregates

"Count items per categories"

```
PREFIX ex: <http://example.org/>
SELECT ?T (Count(?Item) AS ?C)
WHERE { ?Item rdf:type ?T }
GROUP BY ?T
```

Data:

```
@prefix ex: <http://example.org/> .

ex:lemonade1    ex:price 3 ;
                rdf:type ex:Softdrink.
ex:beer1        ex:price 3;
                rdf:type ex:Beer.
ex:wine1        ex:price 3.50 ;
                rdf:type ex:Wine.
ex:wine2        ex:price 4 .
                rdf:type ex:Wine.
ex:wine3        ex:price "n/a";
                rdf:type ex:Wine.
```

Results:

?T	?C
Softdrink	1
Beer	1
Wine	3

SIEMENS

4. Subqueries

- How to create new triples that concatenate first name and last name?
- Possible with SELECT sub-queries or BIND

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX fn: <http://www.w3.org/2005/xpath-functions#>

CONSTRUCT{ ?P foaf:name ?FullName }
WHERE {
  SELECT ?P ( fn:concat(?F, " ", ?L) AS ?FullName )
  WHERE { ?P foaf:firstName ?F ; foaf:lastName ?L. }
}
```

SIEMENS

4. Subqueries

- How to create new triples that concatenate first name and last name?
- Possible with SELECT sub-queries or BIND

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

PREFIX fn: <http://www.w3.org/2005/xpath-functions#>



CONSTRUCT{ ?P foaf:name ?FullName }

WHERE {

 ?P foaf:firstName ?F ; foaf:lastName ?L.

 BIND ( fn:concat(?F, " ", ?L) AS ?FullName )

}
```

SIEMENS

5. Property Path expressions

Arbitrary Length paths, Concatenate property paths, etc.
E.g. names of people Tim Berners-Lee transitively co-authored papers with…

```
SELECT DISTINCT ?N
 WHERE {<http://dblp…/Tim_Berners-Lee>
        (^foaf:maker/foaf:maker)+/foaf:name ?N
       }
```

SIEMENS

Path expressions full list of operators

- elt … Path Element

Syntax Form	Matches
uri	A URI or a prefixed name. A path of length one.
^elt	Inverse path (object to subject).
!uri or *!(uri$_1$/ .../uri$_n$)*	Negated property set. A URI which is not one of *uri$_i$*
!^uri and *!(uri$_1$/ .../uri$_j$/^uri$_{j+1}$/ .../^uri$_n$)*	Negated property set. A URI which is not one of *uri$_i$*, nor uri$_{j+1}$...^uri$_n$ as reverse paths
(elt)	A group path *elt*, brackets control precedence.
elt1 / elt2	A sequence path of *elt1*, followed by *elt2*
elt1 | elt2	A alternative path of *elt1*, or *elt2* (all possibilities are tried).
*elt**	A path of zero or more occurrences of *elt*.
elt+	A path of one or more occurrences of *elt*.
elt?	A path of zero or one *elt*.

- Recent discussion about semantics (counting vs. non-counting) see also [Arenas, Conca, Pérez, WWW2012] and [Losemann, Martens, PODS2012] (forthcoming)

SIEMENS

6. Updates

SQL has not only a query language, but also a Data manipulation language.
→SPARQL Update to fill this gap:

```
PREFIX ex: <http://example.org/>
DELETE { ?Item ex:price ?Pr }
INSERT { ?Item ex:price ?NewPr }
WHERE { ?Item ex:price ?Pr
        BIND (?Pr * 1.1 AS ?NewPr ) }
```

→ Allows to change/update an RDF Store from outside, again via standard HTTP protocol.

SIEMENS

Implementations of SPARQL 1.1:

Some current (partial) SPARQL1.1 implementations:

ARQ

- http://sourceforge.net/projects/jena/
- http://sparql.org/sparql.html

OpenAnzo

- http://www.openanzo.org/

Perl RDF

- http://github.com/kasei/perlrdf/

Corese

- http://www-sop.inria.fr/teams/edelweiss/wiki/wakka.php?wiki=CoreseDownloads

etc.

Others probably forthcoming…

Many SPARQL1.0 endpoints around

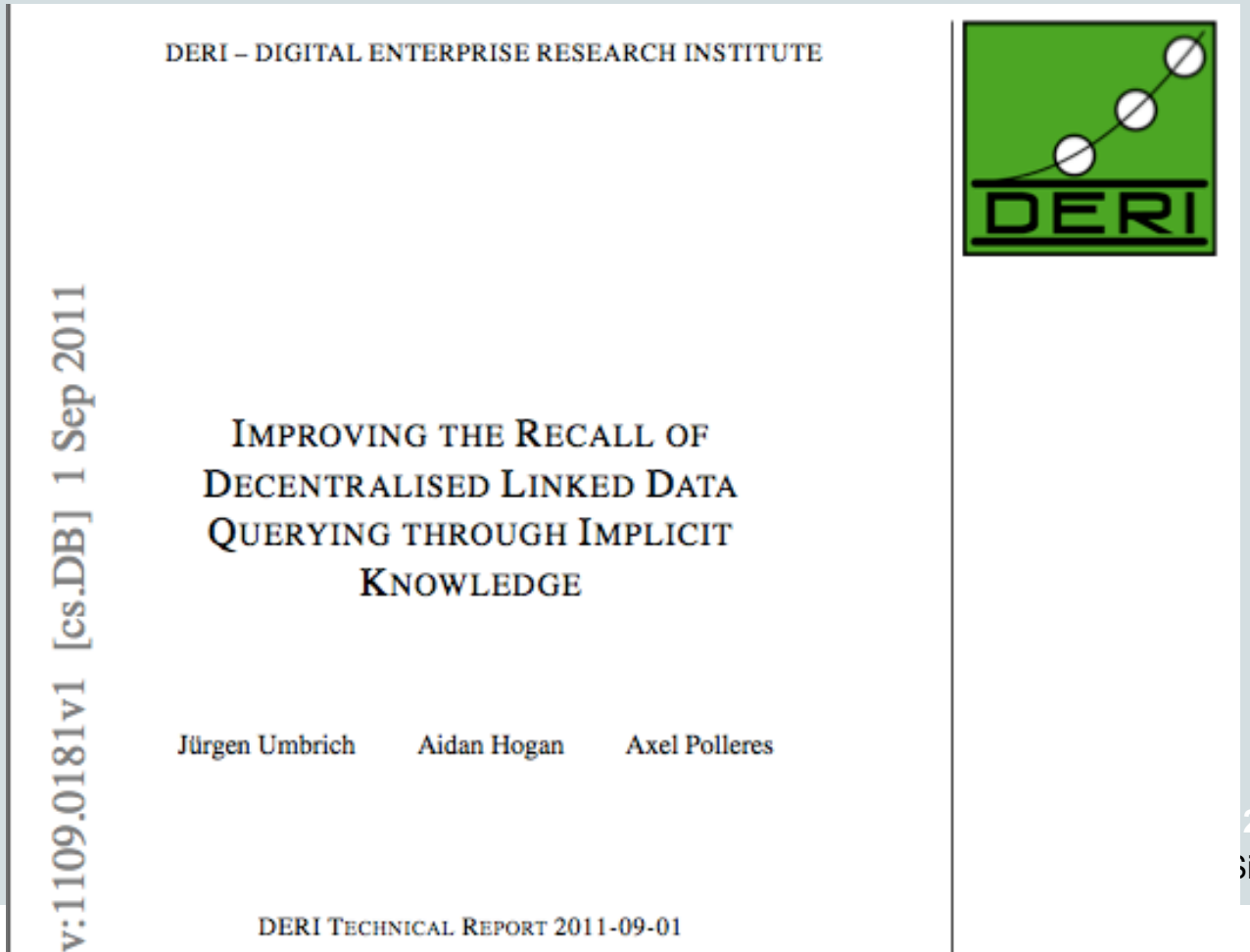
- Dbpedia: http://dbpedia.org/snorql/
- DBLP: http://dblp.l3s.de/d2r/snorql/
- Etc.

SIEMENS

BTW, what about Ontologies?

Ontological Reasoning is needed for Linked Data…

Light weight RDFS + OWL Reasoning
significantly improves Query answering in Linked Data!



DERI – DIGITAL ENTERPRISE RESEARCH INSTITUTE

IMPROVING THE RECALL OF DECENTRALISED LINKED DATA QUERYING THROUGH IMPLICIT KNOWLEDGE

Jürgen Umbrich Aidan Hogan Axel Polleres

DERI TECHNICAL REPORT 2011-09-01

v:1109.0181v1 [cs.DB] 1 Sep 2011

SIEMENS

Example where Reasoning is needed

Give me all facts about Tim Berners-Lee from DBPEdia and DBLP?

```
SELECT ?P ?O
WHERE { <http://dbpedia.org/resource/Tim_Berners-Lee>  ?P ?O }
```

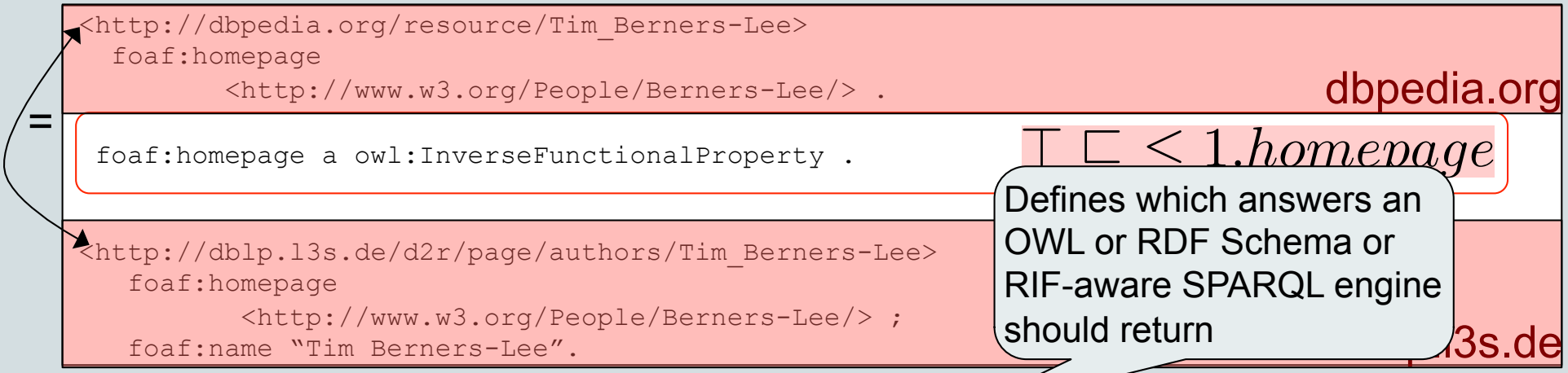
If I ask this query to DBPedia, I get quite some results…
… but not if I ask the same query to DBLP.

Because:
DBLP does not “know” that
`http://dbpedia.org/resource/Tim_Berners-Lee`
=
`http://dblp.l3s.de/d2r/page/authors/Tim_Berners-Lee`

SIEMENS

SPARQL1.1 and OWL/RDF Schema:

SPARQL 1.1 “understands” OWL:



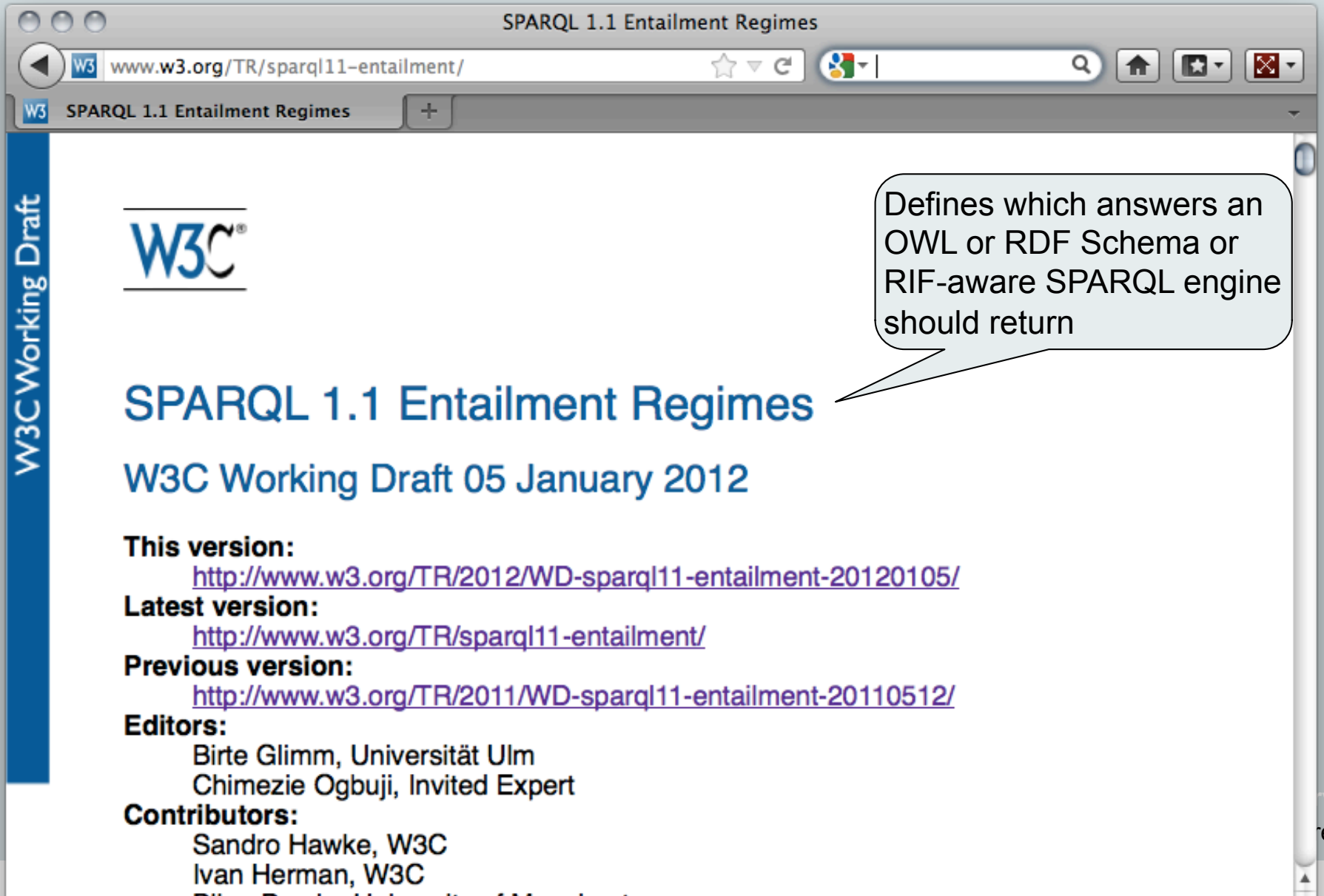

```
SELECT ?P ?O
WHERE { <http://dbpedia.org/resource/Tim_Berners-Lee>  foaf:name ?O }
```

?O
“Tim Berners-Lee”

SIEMENS

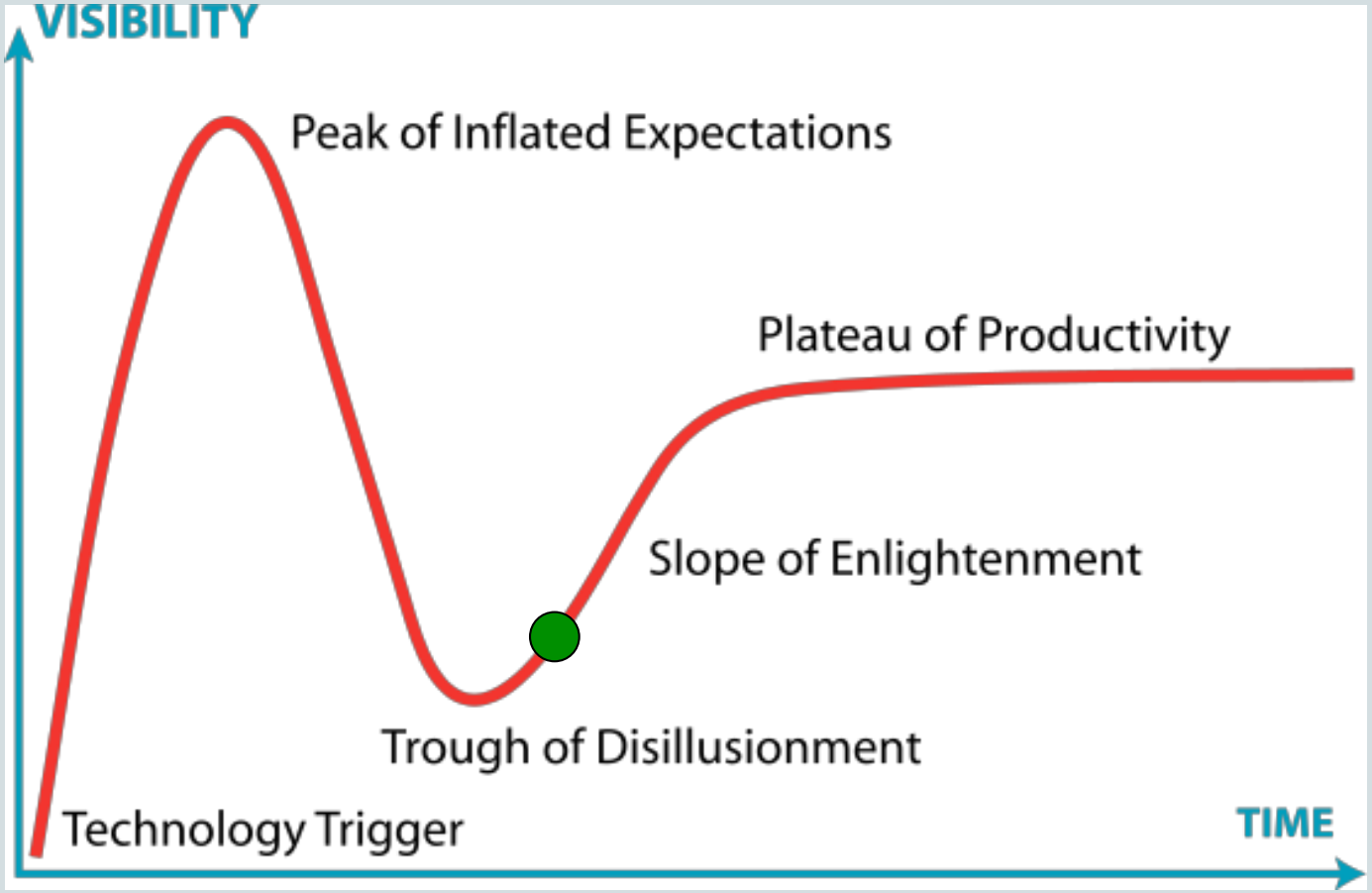
SPARQL1.1 and OWL/RDF Schema:

SPARQL 1.1 “understands” OWL:



SIEMENS

The slope of Enlightment is being climbed…



2012

SPARQL1.1 hopefully a W3C recommendation by late summer!

Review the specs at: www.w3.org/2009/sparql/wiki/

SIEMENS

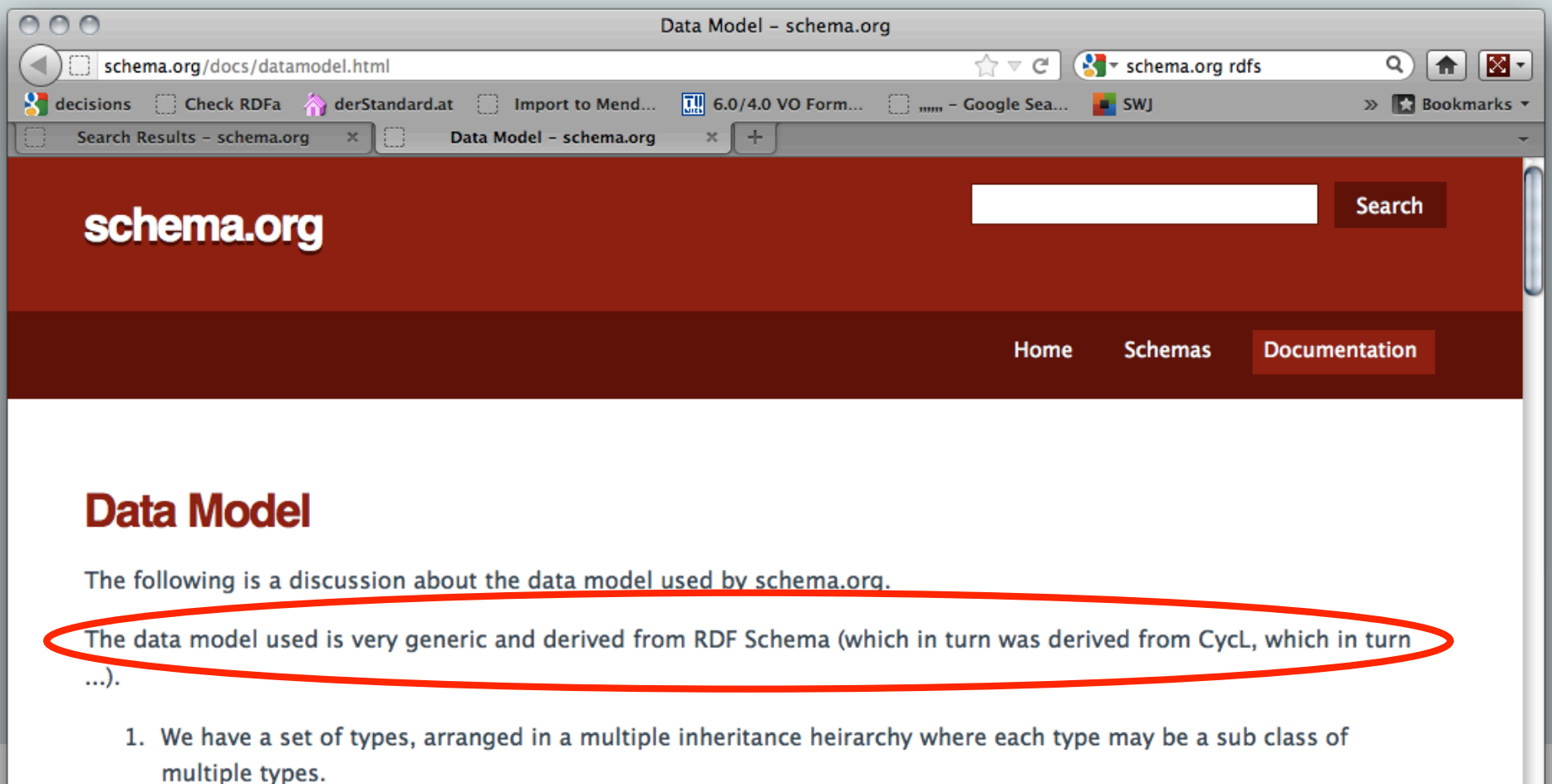
Why is this all interesting for companies?

SIEMENS

Why is this interesting for companies?

Linked Data and Open Data (apart from Linked Open Data) are both emerging paradigms:

- **Linked Data apart from the "LOD cloud":**
  - Enterprise Linked Data (for Knowledge Management within the Enterprise
  - Online companies (eCommerce, Search) start to leverage and support Linked Data

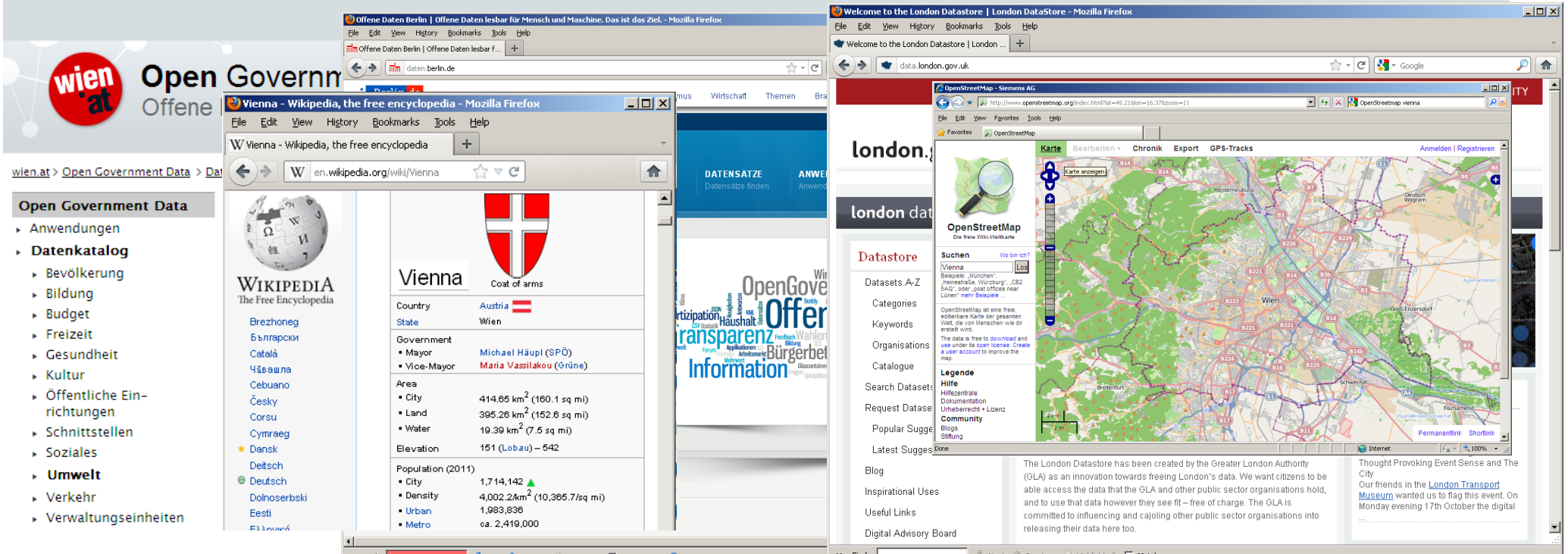


SIEMENS

Why is this interesting for companies?

Linked Data and Open Data (apart from Linked Open Data) are both emerging paradigms:

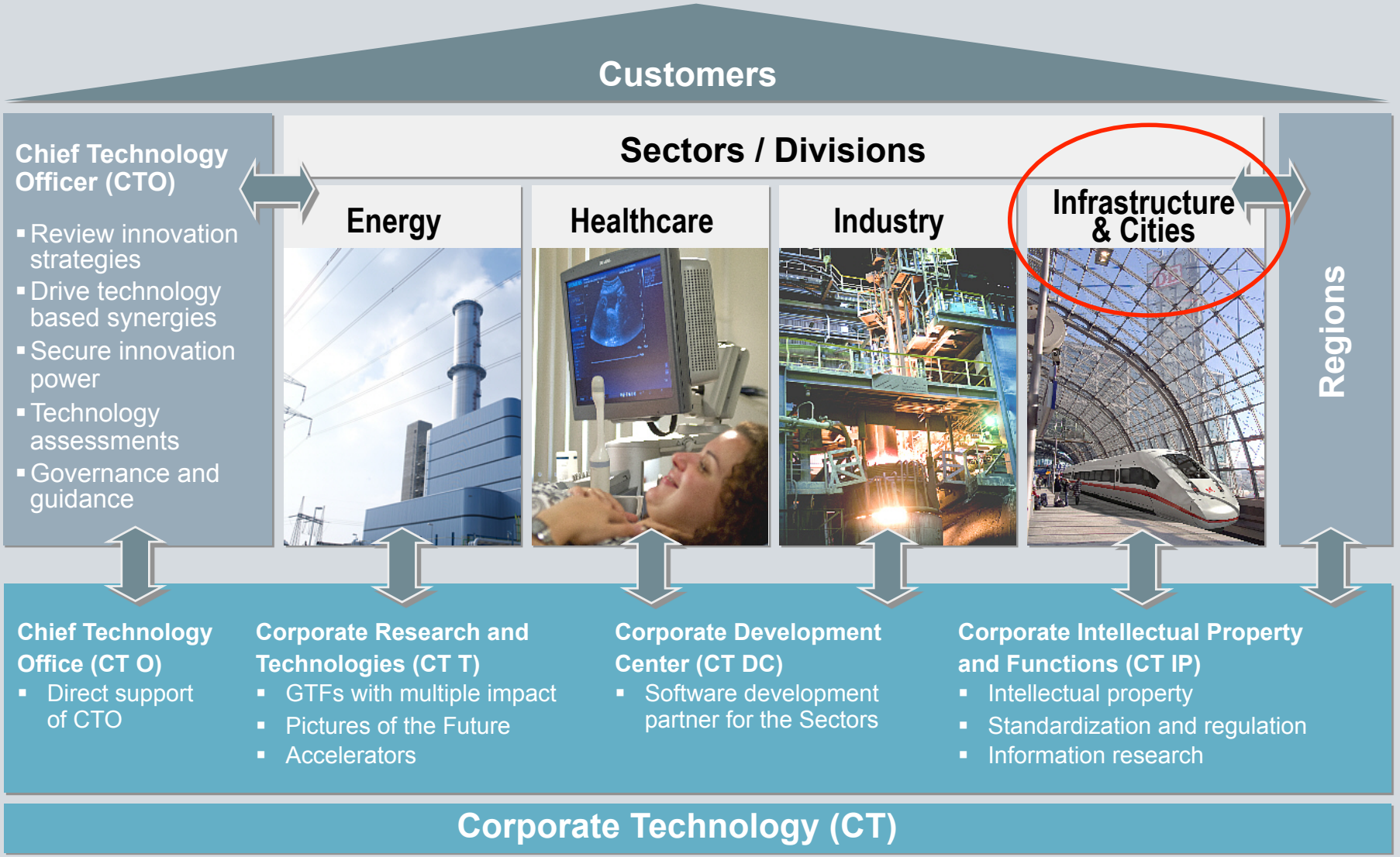
- **Open Data:**
  - Open Data is a trend towards transparency for Governments
  - More Publically available Data leverages new Business Models (not only for SMEs!)
  - Many Governments realize that Opening Data brings more revenue than selling it
  - (EU) regulations force Cities and Governments to publish Data
  - Trend towards harmonization (nationally, at European level, etc.)



Siemens Corporate Technology (CT)

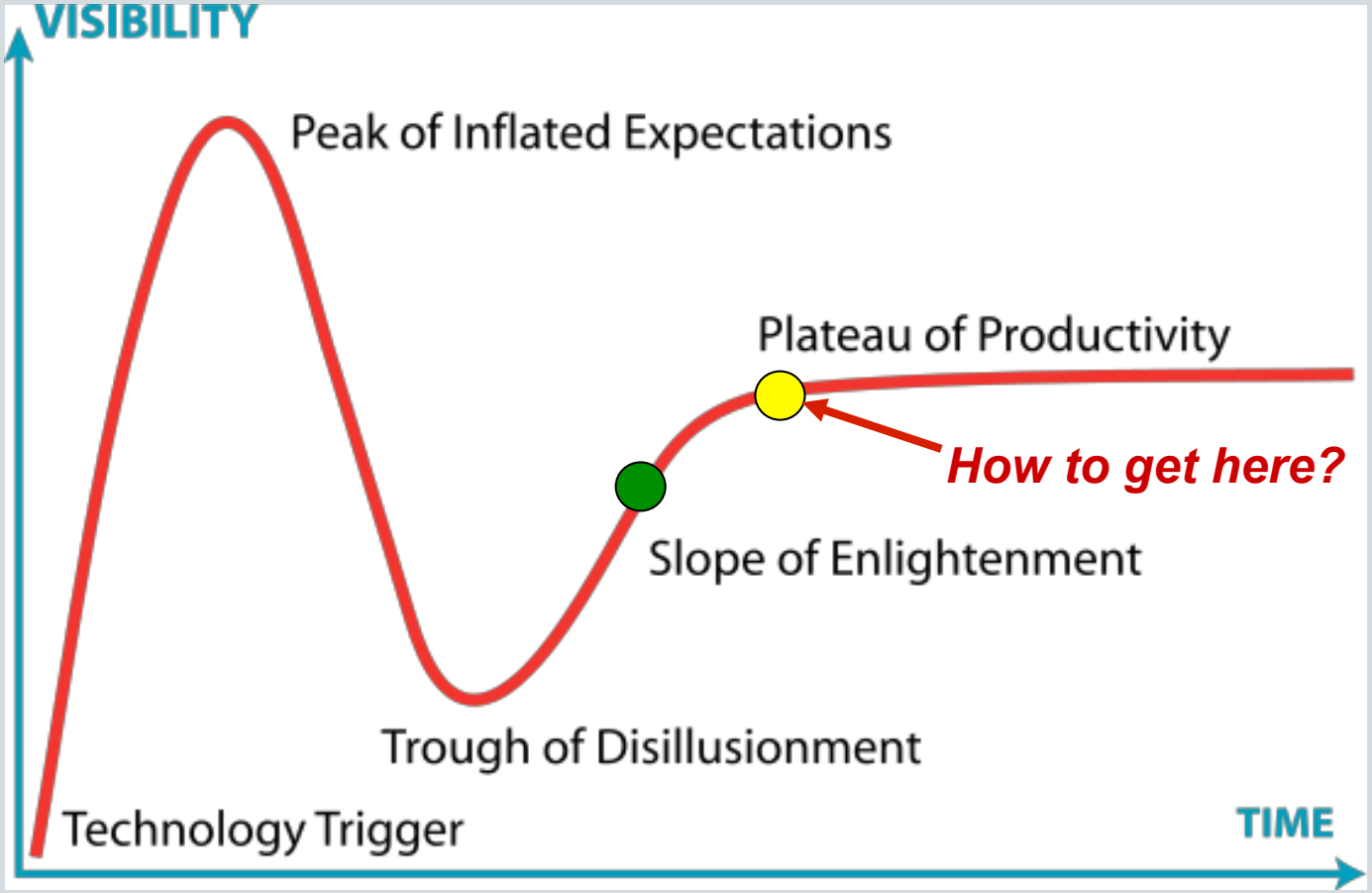
Networking the integrated technology company

SIEMENS



SIEMENS

The slope of Enlightment is being climbed…



- 2012

SPARQL1.1 hopefully a W3C recommendation by late summer!

Review the specs at: *www.w3.org/2009/sparql/wiki/*

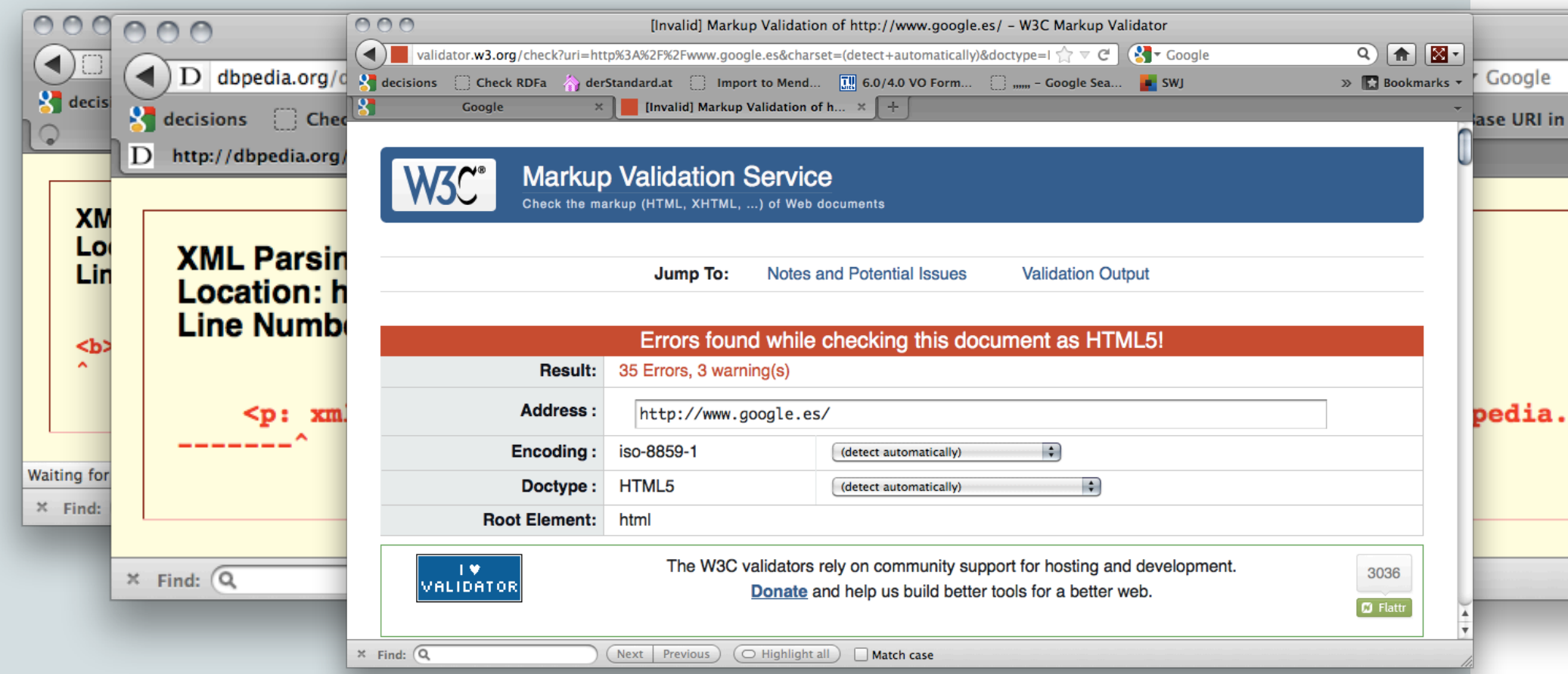
SIEMENS

Challenges ahead...

SIEMENS

Challenges/Problems

The Linked Data Web is “brittle”…



Just like the normal Web is (did you ever try to run an HTML validator on google.es?

SIEMENS

How good/bad is published Linked Data?

ISWC2010

"Almost all infrastructural connectivity on the WoD is mediated by 3 servers, xmlns.com, dbpedia.org and purl.org, making the system very brittle."

Finding the Achilles Heel of the Web of Data: using network analysis for link-recommendation

Christophe Guéret, Paul Groth, Frank van Harmelen, Stefan Schlobach

{cgueret,pgroth,Frank.van.Harmelen,schlobac}@few.vu.nl
VU University Amsterdam
De Boelelaan 1081a, 1081 HV, Amsterdam, The Netherlands

Journal of Web Semantics (forthcoming)

An empirical survey of Linked Data conformance

Aidan Hogan [a], Jürgen Umbrich [a], Andreas Harth [b], Richard Cyganiak [a], Axel Polleres [c], Stefan Decker [a]

[a] *Digital Enterprise Research Institute, National University of Ireland, Galway*
[b] *AIFB, Karlsruhe Institute of Technology, Germany*
[c] *Siemens AG Österreich, Siemensstrasse 90, 1210 Vienna, Austria*

"conformance of data providers varies significantly for the different Linked Data guide- lines highlighted, which in turn may have implications for ad hoc consumers operating over the Web of Data."

How much OWL is on the Web of Data? What's missing for using Linked Data?

SIEMENS

LDOW workshop @ WWW2012

OWL: Yet to arrive on the Web of Data?

Birte Glimm
Ulm University, Institute of Artificial Intelligence, 89069 Ulm, Germany

Aidan Hogan
Digital Enterprise Research Institute, National University of Ireland Galway, Ireland

Markus Krötzsch
University of Oxford, Department of Computer Science, OX1 3QD Oxford, United Kingdom

Axel Polleres
Siemens AG Österreich, Siemensstrasse 90, 1210 Vienna, Austria

"Single-triple expressible OWL RL axioms are most prominent on the Web."

DESWEB workshop @ICDE2012

"indexes for Linked Data in the Web are often incomplete and outdated."
→ Needs rethinking in terms of applying traditional Database techniques.

Linked Data and Live Querying for Enabling Support Platforms for Web Dataspaces

Jürgen Umbrich[1], Marcel Karnstedt[1], Josiane Xavier Parreira[1], Axel Polleres[2], Manfred Hauswirth[1]

[1]*Digital Enterprise Research Institute, National University of Ireland, Galway, Ireland*
[2]*Siemens AG Österreich, Siemensstraße 90, 1210 Vienna, Austria*
{firstname.lastname}@[1]deri.org/[2]siemens.com

SIEMENS

Challenges/Outlook

Last, but not least:

- A lot (most?) published Open Data is not yet RDF (but CSV, XML, GML, …
- Within Linked Data, not only malformatting, but also (logical) inconsistencies are an issue (see papers mentioned.)
- Taxonomic Reasoning (OWL/DL) Reasoning is not enough: Configuration, Planning, Trust/Provenance Models & Reasoning, etc.

→ Using the rich lode of Semantic Web data is all about Practical Applications of Agents … lots of work to be done for PAAMS

More details: 3hr SPARQL1.1 Tutorial on my WebPage http://www.polleres.net