propel

PROPELLING THE POTENTIAL OF
ENTERPRISE LINKED DATA IN AUSTRIA

(Enterprise) Linked Data: What's missing?

Axel Polleres, Sabrina Kirrane, Javier D. Fernández

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

***Disclaimer*: This is NOT a research talk…**

The PROPEL project:

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

Vienna Semantic Web Meetup

- 15 January 2015 year's Semantic Web meetup:
- **Business Semantics & Enterprise Linked Data**

- Various companies in Austria already making business & products with Linked Data, e.g.:

redlink · poolparty · Capsenta The Smart Data Company™ · CAMBRIDGE SEMANTICS

- **But**: Still a niche market... What's missing?
  - Awareness! https://www.semantic-web.at/news/linked-data-awareness-barometer-2015
  - Addressing Technology gaps (security, time, efficient interchange, standard tool chain)

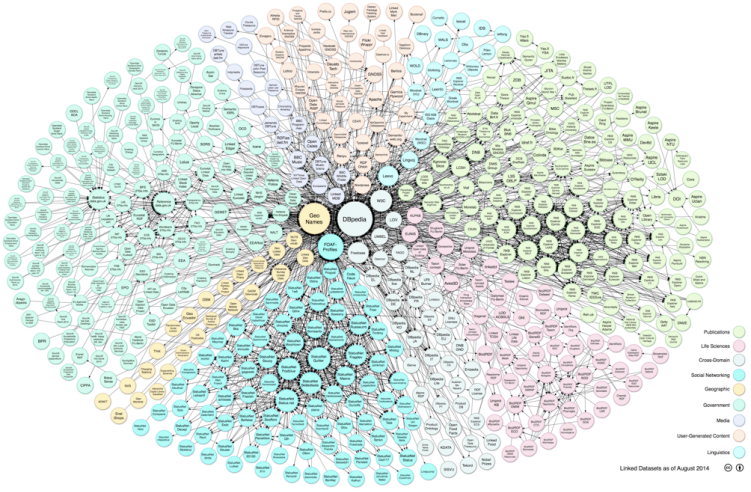
… Can we learn from industry? Shall we?

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

PROPEL

Propelling the Potential of Enterprise **Linked Data in Austria**

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA



Goals:

- How big is the market? Market Analysis, requirements and use cases
- How ready is Linked Data as an integration paradigm for the Enterprise?
- Research and Development Roadmap, Exploratory Study, Technology Gaps
- Eventually: Models for an **Enterprise** *Linked Data value chain*

IDC Analyze the Future · TU WIEN · WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

So, how to approach this?

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

- First, we better know what we can offer…
  - What is Enterprise Linked Data?
  - What is the Semantic Web?
  - Who are we?

IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY
PROPEL
bmvit
FFG

Some common misunderstandings...

What's the difference between Enterprise Linked Data and... ?

Linked Data vs. “The” Semantic Web

Overlaps:

- *”Linked Data is the Semantic Web done right"* (Tim Berners-Lee)
- The actual Semantic **Web** is made up of Linked Data.
- Linked Data is based on **Semantic web standards**.

Key Differences:

- Semantic Web was all about "semantifying" the Web, Linked Data is based on Web standards (URIs, http), but doesn't center around **Web** *pages*.
- LD is a more **pragmatic** "**bottom-up**" approach.

- *"Linked Data is mainly about publishing structured data in RDF using URIs rather than focusing on the ontological level or inference."*

M. Hausenblas "Exploiting Linked Data For Building Web Applications"
IEEE Internet Computing, 2009

Linked Data vs. Open Data

Overlaps:

- Openness is a core principle in the design of LD
- Many Linked Data sets published under an open license
  → Linked Open Data and LD are often used interchangeably

Key differences:

- Linked Data technologies can be used without publishing data – e.g., for internal and external data integration.
- Probably not all open data will ever be linked (the majority will remain in formats such as csv, txt, json,xml etc.)

Linked Data vs Big Data

Overlaps:

- LD as a whole is big (38.606.408.854 triples and counting! *)
- No rigid up-front (e.g., relational) data model
- Big Data technologies (e.g., Hadoop) are used to handle LD
- LD can represent knowledge extracted from big unstructured data

Key Differences:

- Individual linked data sets are typically not "big" per se (e.g., English DBpedia dump currently < 5 GB)
- LD is structured and semantically explicit, single format (RDF) "big data lakes" are typically neither - RDMBS, NoSQL/“Polyglott persistence”, non-core formats, unstructured textual+mmedia data…
- Big data based on distributed data infrastructures within an organization (e.g., Hadoop clusters), LD creates a decentralized, globally distributed data infrastructure

*) http://lodlaundromat.org as per 2016-05-10

ELD vs. LED

Enterprise Linked Data (intra-enterprise):

Internal use of LD technologies within organizations, e.g.,

- to integrate heterogeneous systems at the data level
- for advanced content/knowledge/… management
- as a basis for innovative products and services

Linked Enterprise Data (inter-entrprise):

- Cross-organizational data integration
- Data markets and data ecosystems
- Decentralized infrastructure for a networked economy

Let's take a step back…

- What are the expectations/requirements?
- What can we offer as a community?

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Let's take a step back…

- What are the expectations/requirements?
- What can we offer as a community?

Taking a business/application centric view

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Interviews

23 interviews:

- Domains
  - Consulting, Engineering, Environment, Finance and Insurance, Government, Healthcare, ICT, IT, Media, Pharmaceutical, Professional Services, Real Estate, Research, Startup, Tourism, Transports & Logistics
- Roles
  - Business Intelligence, CEO, Chief Engineer, Data and Systems Architect, Data Scientist, Director Information Management, Enterprise Architect, Founder, General Secretary, Governance, Risk & Compliance Manager, Head of Communications and Media, Head of Development, Head of HR, Head of R&D, Innovation Manager, Information Architect, IT Project Manager, Management, Managing director, Marketing Analyst, Principle System Analyst, Project Coordinator, Researcher, Technical Specialist

Technologies in need…

Analytics	Computational linguistics & NLP	Concept tagging & annotation	Data integration
Data management	Dynamic data / streaming	Extraction, data mining, text mining, entity extraction	Logic, formal languages & reasoning
Human-Computer Interaction & visualization	Knowledge representation	Machine learning	Ontology/thesaurus /taxonomy management
Quality & Provenance	Recommendations	Robustness, scalability, optimization and performance	Searching, browsing & exploration
	Security and privacy	System engineering	

We pretty much ended up in all areas that SW touches upon!

IDC Analyze the Future

TU WIEN

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

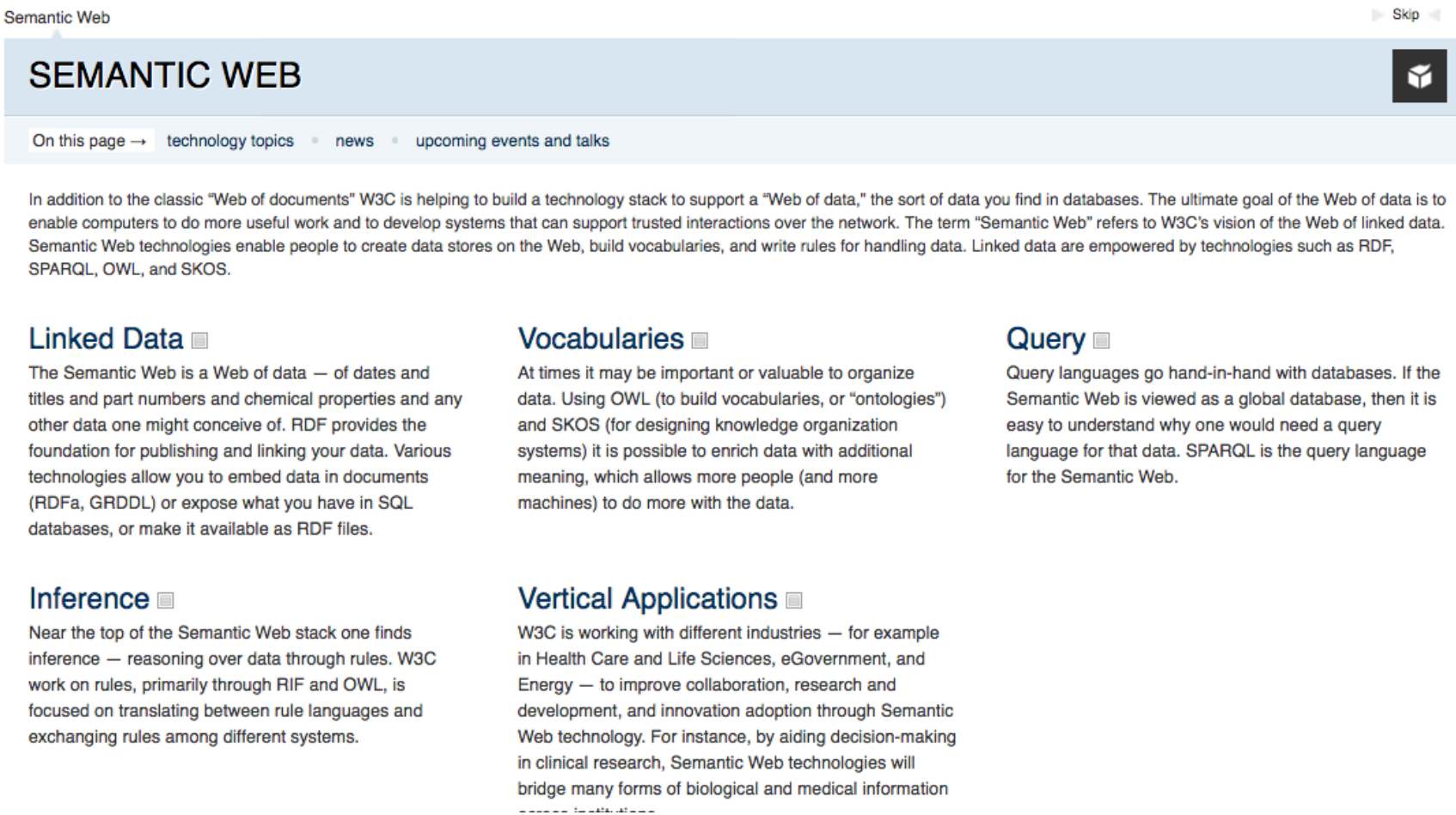
SEMANTIC WEB COMPANY PROPEL

bmvit

Standards

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

https://www.w3.org/standards/semanticweb/



Semantic Web

Skip

SEMANTIC WEB

On this page → technology topics • news • upcoming events and talks

In addition to the classic "Web of documents" W3C is helping to build a technology stack to support a "Web of data," the sort of data you find in databases. The ultimate goal of the Web of data is to enable computers to do more useful work and to develop systems that can support trusted interactions over the network. The term "Semantic Web" refers to W3C's vision of the Web of linked data. Semantic Web technologies enable people to create data stores on the Web, build vocabularies, and write rules for handling data. Linked data are empowered by technologies such as RDF, SPARQL, OWL, and SKOS.

Linked Data

The Semantic Web is a Web of data — of dates and titles and part numbers and chemical properties and any other data one might conceive of. RDF provides the foundation for publishing and linking your data. Various technologies allow you to embed data in documents (RDFa, GRDDL) or expose what you have in SQL databases, or make it available as RDF files.

Vocabularies

At times it may be important or valuable to organize data. Using OWL (to build vocabularies, or "ontologies") and SKOS (for designing knowledge organization systems) it is possible to enrich data with additional meaning, which allows more people (and more machines) to do more with the data.

Query

Query languages go hand-in-hand with databases. If the Semantic Web is viewed as a global database, then it is easy to understand why one would need a query language for that data. SPARQL is the query language for the Semantic Web.

Inference

Near the top of the Semantic Web stack one finds inference — reasoning over data through rules. W3C work on rules, primarily through RIF and OWL, is focused on translating between rule languages and exchanging rules among different systems.

Vertical Applications

W3C is working with different industries — for example in Health Care and Life Sciences, eGovernment, and Energy — to improve collaboration, research and development, and innovation adoption through Semantic Web technology. For instance, by aiding decision-making in clinical research, Semantic Web technologies will bridge many forms of biological and medical information

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Standards Toolbox

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

- Constraint language → Shapes Constraint Language (SHACL)
- Query language → SPARQL
- Machine readable data → RDF Quads, RDF N-Triples Trig …
- Update language → SPARQL UPDATE
- Schema/ Ontology/ Concepts/ Taxonomy → RDF Schema OWL SKOS
- Rule language → RIF
- Meta data → PROV, PROV-O, PROV-N DCAT
- Trans-lation language → GRDDL
- Web Services Description → SA-WSDL
- Transfor-mation language → RDB2RDF Direct Mapping R2RML
- Multidim-ensional data → RDF Data Cube
- Vocab-ulary → The Organization Ontology Time ontology in OWL

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Standards Toolbox (incl. W3C member submissions)

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

Constraint language
Shapes Constraint Language (SHACL)

Query language
SPARQL

Machine readable data
RDF Quads,
RDF N-Triples
Trig …
HDT

Update language
SPARQL UPDATE

Schema/ Ontology/ Concepts/ Taxonomy
RDF Schema
OWL
SKOS

Rule language
RIF
N3
SWRL
WSML

Meta data
PROV,
PROV-O,
PROV-N
DCAT

Trans-lation language
GRDDL
XSPARQL

Web Services Description
SA-WSDL
WSMO
OWL-S
METEOR-S

Transfor-mation language
RDB2RDF Direct Mapping
R2RML

Multidim-ensional data
RDF Data Cube

Vocab-ulary
The Organization Ontology
Time ontology in OWL
SIOC
…

IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY PROPEL
bmvit

Technologies in need…

- Analytics
  - RDF Cube
- Computational linguistics & NLP
- Concept tagging & annotation
  - SKOS
- Data integration
  - RDB2RDF
  - GRDDL Semantic Annotations
- Data management
  - PROV DCAT
- Dynamic data / streaming
- Extraction, data mining, text mining, entity extraction
  - RDB2RDF
  - GRDDL Semantic Annotations
- Logic, formal languages & reasoning
  - RIF
  - RDFS OWL
- Human-Computer Interaction & visualization
- Knowledge representation
  - RDF
  - SHACL
  - The Org Ontology
- Machine learning
- Ontology/thesaurus/taxonomy management
  - RDFS OWL
- Quality & Provenance
  - PROV DCAT
- Recommendations
- Robustness, scalability, optimization and performance
- Searching, browsing & exploration
  - SPARQL
- Security and privacy
- System engineering
  - Web Service Descriptions

Technologies in need…

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

Analytics

RDF Cube

Computational linguistics & NLP

Concept tagging & annotation

SKOS

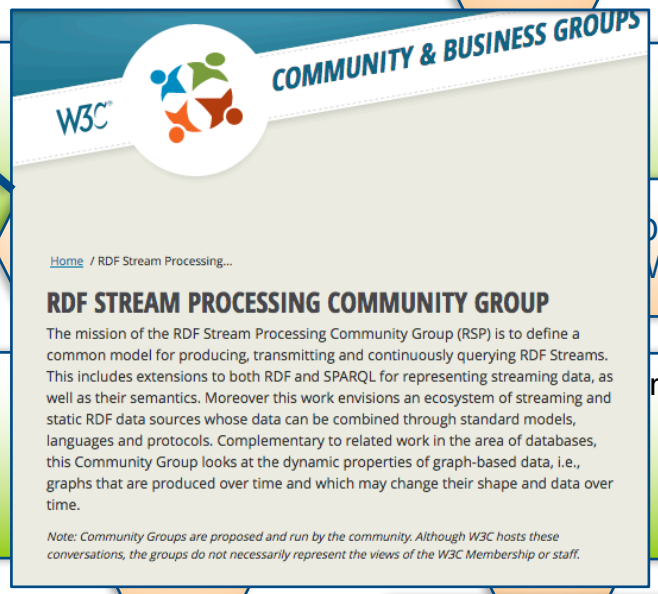
Dynamic data / streaming

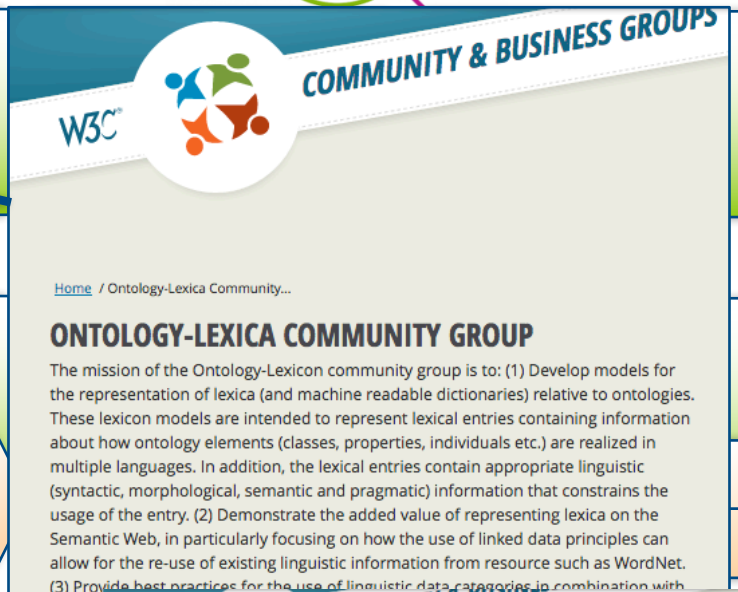
Machine learning

Security and privacy

browsing

ARQL





???? What could **standards** offer here…

THINKING OUT LOUD

Potentially interesting: exchange of ML models, e.g. learnt regression models, decision trees, etc. extension of PROV,…? **W3C CGs** offer a forum for these things!

Plus many more… not necessarily listed as SW standards, but definitly related!

Home / RDF and XML...

RDF AND XML INTEROPERABILITY COMMUNITY GROUP

The goal of this group is to 1) identify application areas in which the combined processing of XML and RDF data and tooling is beneficial; 2) identify issues that hinder the joint usage of the two technology stacks 3) formulate best practices to resolve the issues or propose standardization topics. The goal does not only take into account the data representation formats XML and RDF, but all related technologies (e.g. for XML: XSLT, XQuery; for RDF: RDF Schema, SPARQL) and selected XML (e.g. OData) or RDF vocabularies. The group should be driven by needs of industries that already deplo
also cover adjacent technologies like JS
this group. The outcome should focus
with XML and RDF, but on small buildi
standardization topics) that can be re-
scenarios.

Data integration

Page | Discussion

WebID

Contents [hide]

1 WebIDs and the WebID Protocol
1.1 What is a WebID?
1.1.1 Why should I get a WebID?
PublicKey WebID
WebID?
WebIDs for Users
1.2.7 Can I use
1.3 What is the WebID Auth
1.3.1 Some Protocol Details
1.3.2 Why is the WebID Protocol Viable?

Security and privacy

Special pages
Printable version
Permanent link
Page information
Browse properties

W3C Architecture domain

Nearby: XML Activity • *Documents:* XML Binary Characterization Use Cases, XML Binary Characterization Properties, XML Binary Characterization Measurement

XML Binary Characterization Working Group Public Page

Note: This Working Group ended successfully, and on schedule. The actual technical work to establish an internationally recognized standard for

The XML Binary Characterization Working Group is the public web page for the (now completed) XML Binary Characterization Working Group (details).

This is part of the W3C XML Activity.

- Introduction
- Charter and Participation
- IPR Disclosures
- Publications

Introduction

The XML Binary Characterization Working Group was created as a result of the Binary Interchange Workshop.

The XML Binary Characterization Working Group was tasked with gathering information about uses cases where the overhead of generating, pars application, characterizing the properties that XML provides as well as those that are required by the use cases, and establishing objective, shared r properties.

The Binary Interchange Workshop report can be found at http://www.w3.org/2003/08/binary-interchange-workshop/Report.html.

Robustness, scalability, optimization and performance

IDC Analyze the Future

FFG

User Stories

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

4 out of 60 user stories we collected in the interviews:

- Horizontals
  - Business processes (e.g. product logistics and supply chain management)
  - Human resources (e.g. expert and resource management)
- Verticals
  - Media & Publishing
  - Healthcare & Pharma

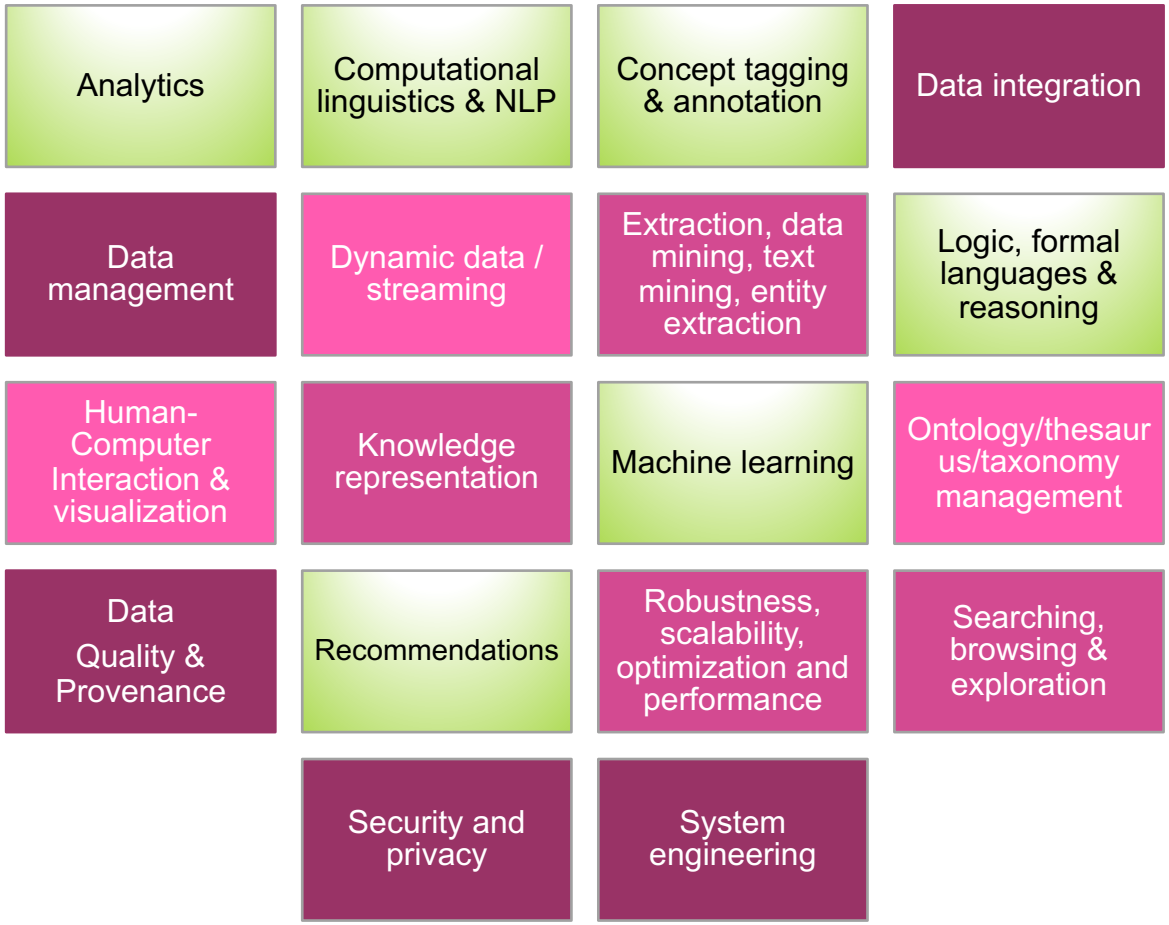
IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Business Processes

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

"I would like *to be able to exchange information and coordinate production and logistics with suppliers and customers…"*

***"…so that I can** improve efficiency, effectiveness and flexibility of my inventory management and operations"*



IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Human Resources

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

"I would like *identify expertise within our large organisation and be able to pinpoint the relevant experts…"*

"…so that I can *I can identify top trends within the organisation and expertise for the organisation as a whole"*

Analytics	Computational linguistics & NLP	Concept tagging & annotation	Data integration
Data management	Dynamic data / streaming	Extraction, data mining, text mining, entity extraction	Logic, formal languages & reasoning
Human-Computer Interaction & visualization	Knowledge representation	Machine learning	Ontology/thesaurus/taxonomy management
Quality& Provenance	Recommendations	Robustness, scalability, optimization and performance	Searching, browsing & exploration
	Security and privacy	System engineering	

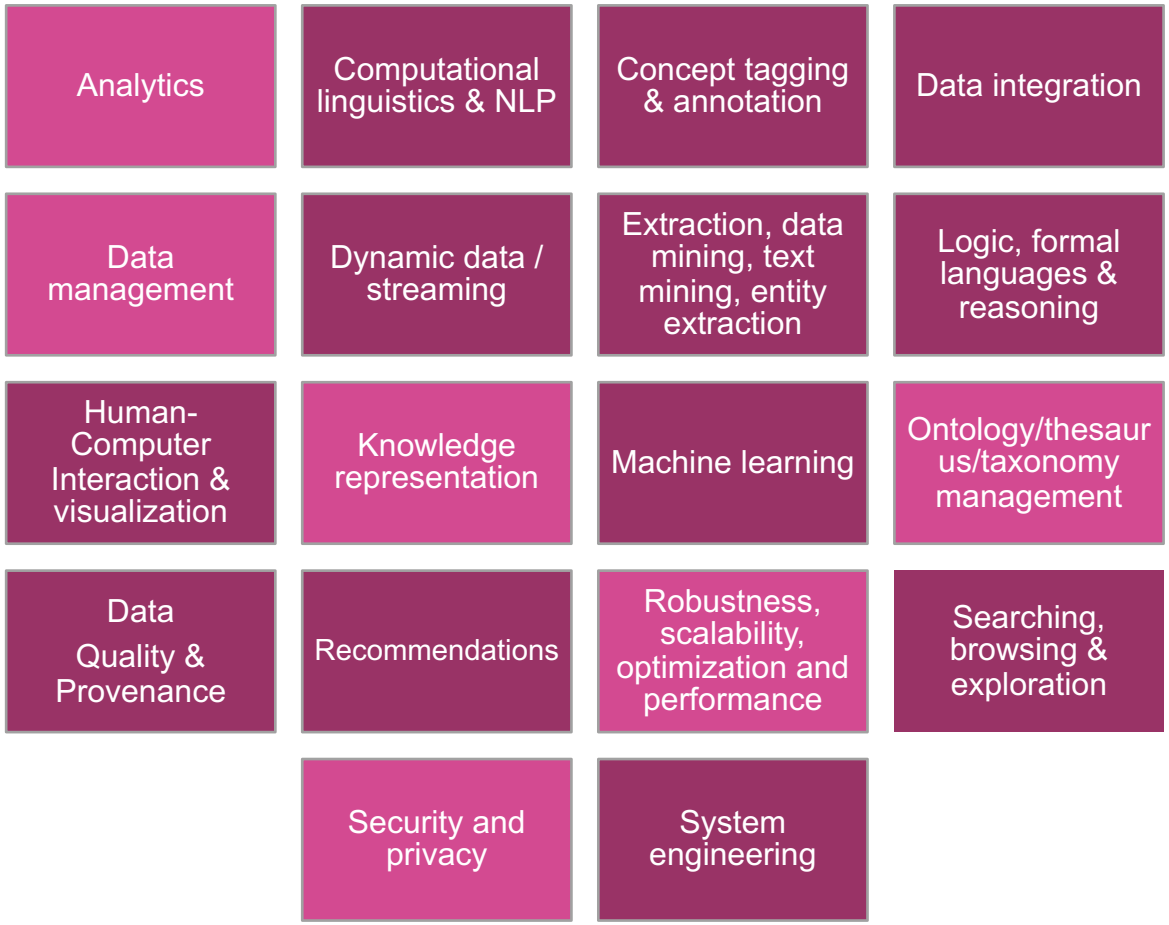
IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Media & Publishing

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

I would like to display personalized content as precise as possible

So that my readers stay as long as possible on my website.



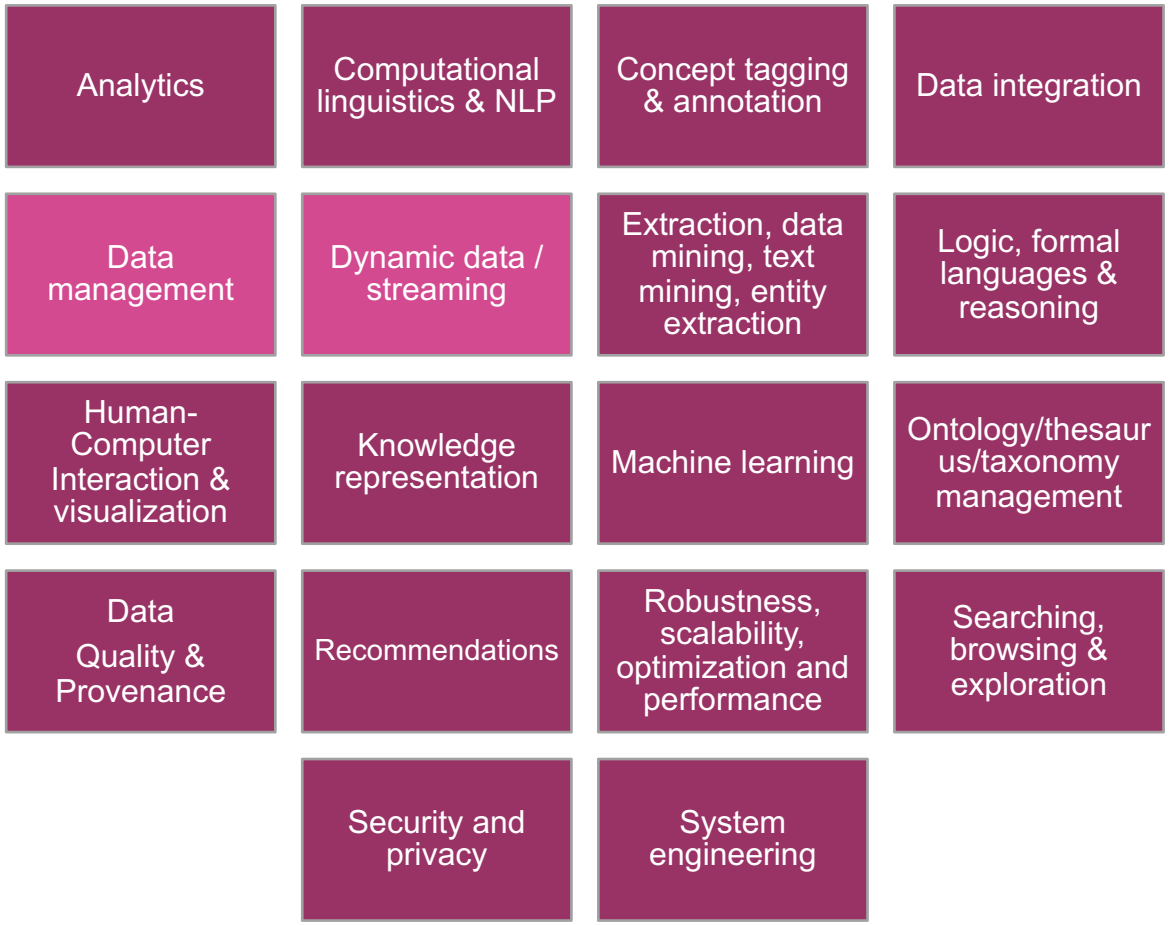
IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Healthcare & Pharma

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

I would like *to Integrate disparate systems that are:*
-Hard to integrate
-Widespread
-Contain the same data that contradicts each other

So that I can *gain insights from other clinical trials*



IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

User Stories – Bottomline…

4 out of 60 user stories we collected in the interviews:

- Horizontals
  - Business processes (e.g. product logistics and supply chain management)
  - Human resources (e.g. expert and resource management)
- Verticals
  - Media & Publishing
  - Healthcare & Pharma

According to our interviews best fit to what we can offer technologywise!

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · Semantic Web Company · PROPEL · bmvit · FFG

Let's take a step back…

- What are the expectations/requirements?

- What can we offer as a community?

What technologies and standards do we have available?

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Technology Toolbox

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY
PROPEL
bmvit
FFG

Let's take a step back…

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

- What can we offer as a community?

Taking an introspective view

IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY
PROPEL
bmvit
FFG

Let's check back with our research results…

- 3 seminal papers:

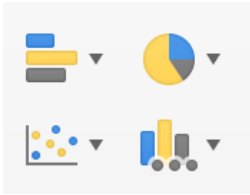
SCIENTIFIC AMERICAN.COM
May 17, 2001
The Semantic Web
A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities
By Tim Berners-Lee, James Hendler and Ora Lassila

SCIENTIFIC AMERICAN
Features - January 19, 2009
The Semantic Web in Action
Corporate applications are well under way, and consumer uses are emerging
By Lee Feigenbaum, Ivan Herman, Tonya Hongsermeier, Eric Neumann and Susie Stephens

viewpoints
DOI:10.1145/2890489 Abraham Bernstein, James Hendler, and Natalya Noy
Viewpoint
A New Look at the Semantic Web
Seeking to make Web data "smarter" by utilizing a new kind of semantics.

2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016

- Monitoring SW communitie's major venues:
  - ISWC (since 2006), ESWC (since 2006), SEMANTiCS (since 2007), JWS (since 2006), SWJ (since 2010)



Semantic Web/Linked Data over time…

SCIENTIFIC AMERICAN.COM

May 17, 2001

The Semantic Web

A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities

By Tim Berners-Lee, James Hendler and Ora Lassila

Subtopics:

Expressing Meaning

Knowledge Representation

Ontologies

Agents

Evolution of Knowledge

Knowledge Representation

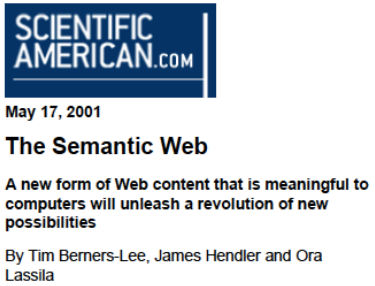
propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

"The challenge of the Semantic Web, therefore, is to provide a language that **expresses** both **data** and **rules for reasoning** about the data and that allows rules from any existing knowledge-representation system to be **exported** onto the Web."

"Adding **logic** to the Web—the means to use **rules** to make inferences, choose courses of action and answer questions"

"Ideally, the program must have a way to discover such common **meanings** for whatever databases it encounters."

How has knowledge representation based research evolved?


SCIENTIFIC AMERICAN.COM

May 17, 2001

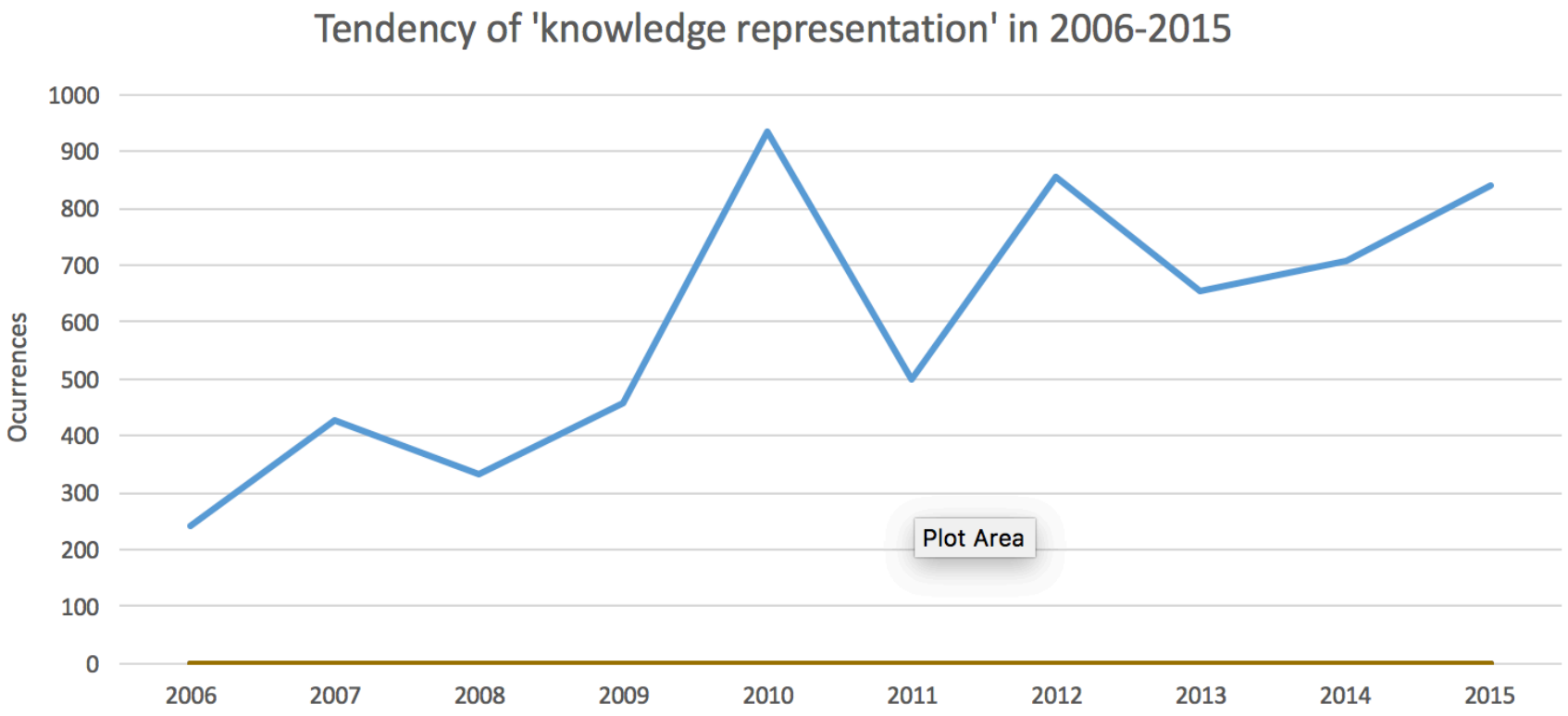
The Semantic Web

A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities

By Tim Berners-Lee, James Hendler and Ora Lassila

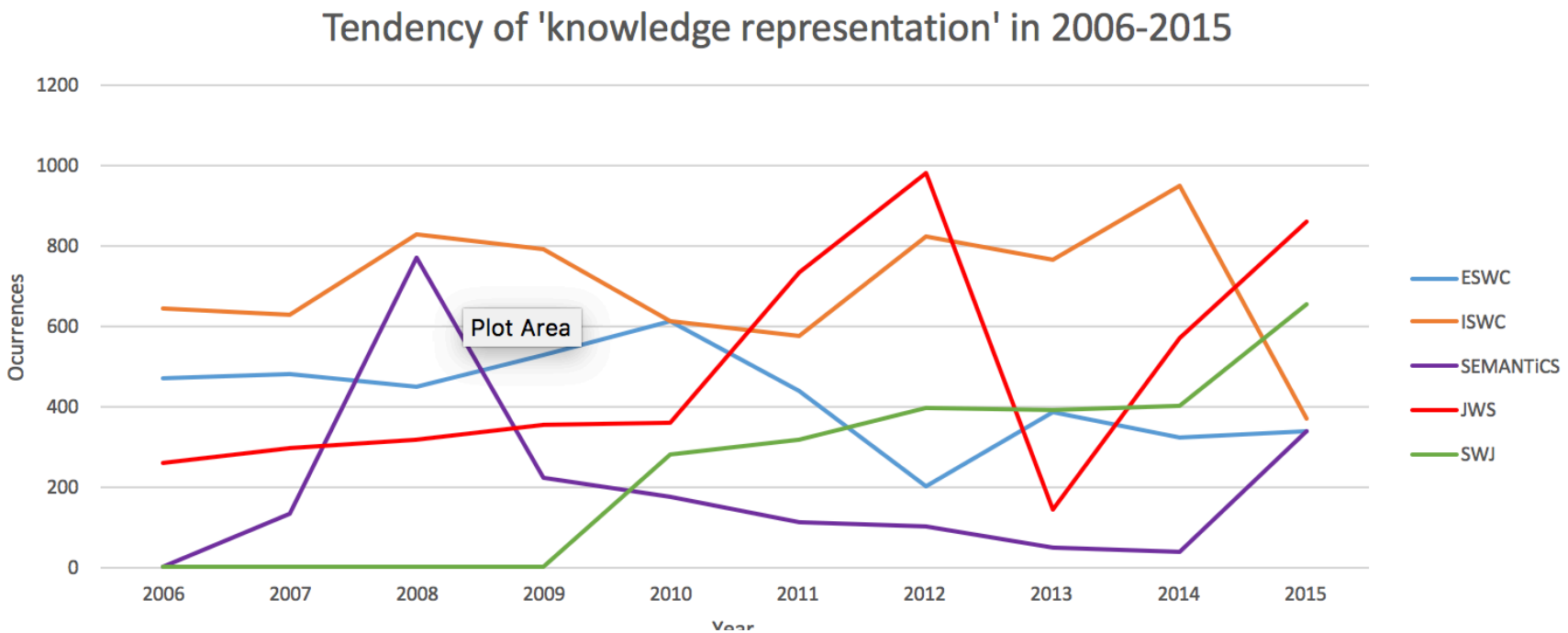
The Semantic Web T Berners-Lee, J Hendler, O Lassila Scientific American 284 (5), 34-43, 2001

The importance of Knowledge Representation & Reasoning:



The importance of Knowledge Representation & Reasoning:

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA



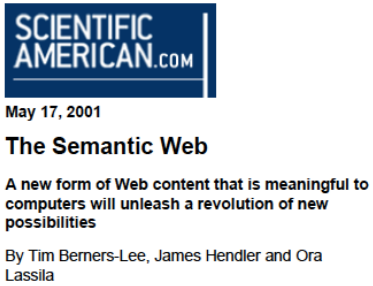
IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY PROPEL
bmvit
FFG

Agents

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

"The real power of the Semantic Web will be realized when people create many programs that collect Web content from diverse sources, process the information and exchange the results with other programs. The effectiveness of such **software agents will increase exponentially** as more machine-readable Web content and **automated services** (including other agents) become available"

Are agents still a hot topic?


SCIENTIFIC AMERICAN.COM

May 17, 2001

The Semantic Web

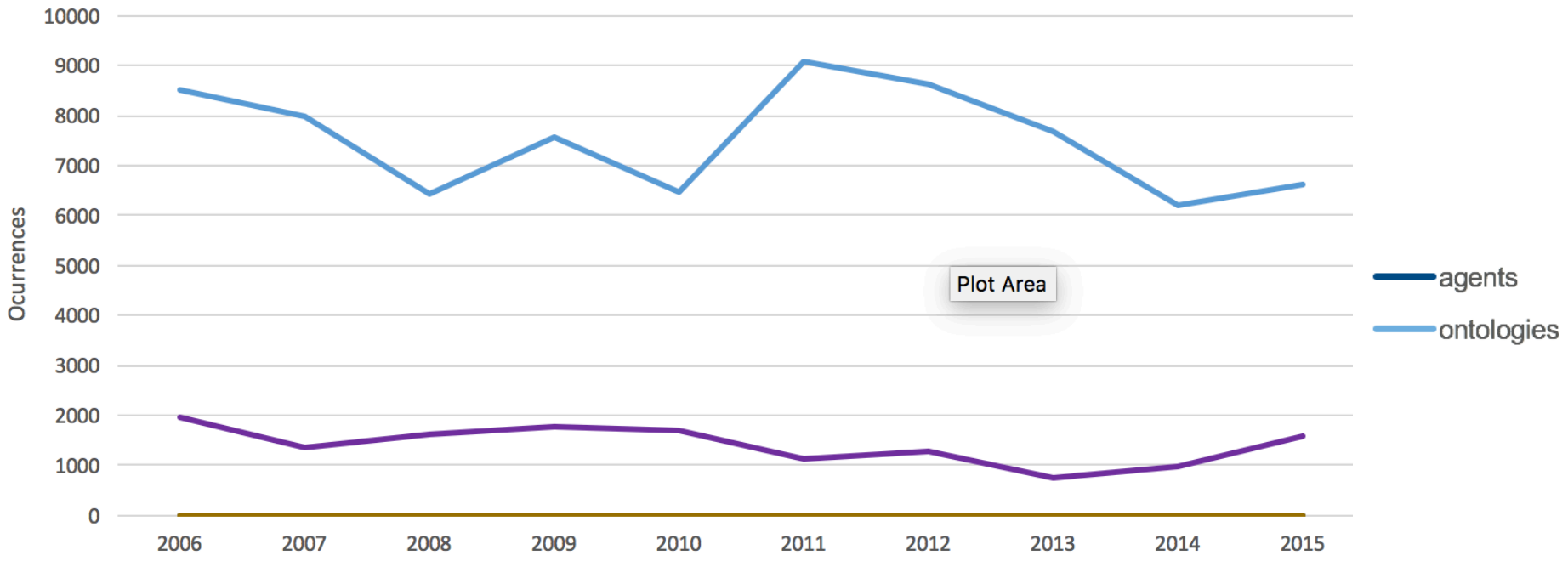
A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities

By Tim Berners-Lee, James Hendler and Ora Lassila

The Semantic Web T Berners-Lee, J Hendler, O Lassila Scientific American 284 (5), 34-43, 2001

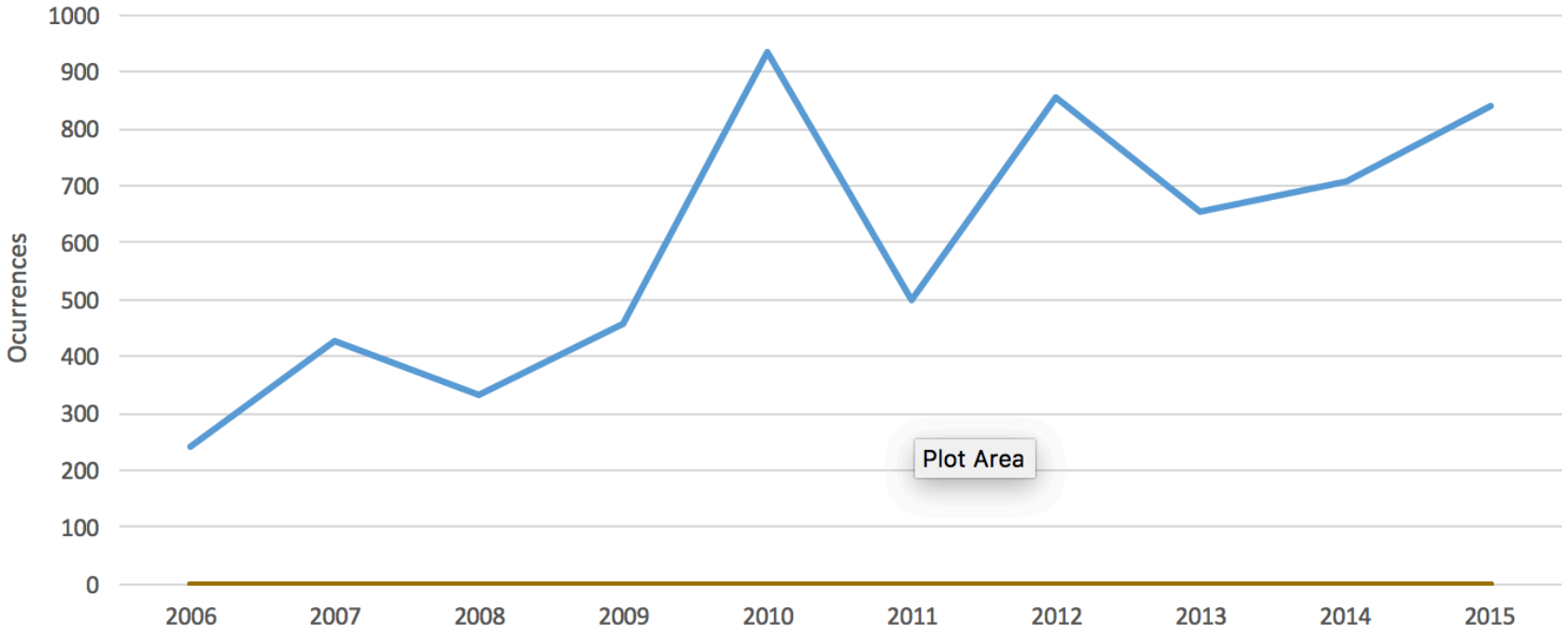
Agents vs Ontologies

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA



IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY
PROPEL
bmvit
FFG

Evolution of Knowledge



“temporal semantics/reasoning” “evolving/evolution”

… overall lower than than the other areas?

Semantic Web/Linked Data over time…

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

SCIENTIFIC AMERICAN

Features - January 19, 2009

The Semantic Web in Action

Corporate applications are well under way, and consumer uses are emerging

By Lee Feigenbaum, Ivan Herman, Tonya Hongsermeier, Eric Neumann and Susie Stephens

Early adopters:
- MITRE
- Chevron
- **British Telecom**
- Boeing
- **Ordnance Survey**
- Eli Lily
- **Pfizer**
- Agfa
- Food and Drug Administration
- **National Institutes of Health**

Software adopters/products:
- Oracle
- Adobe
- Altova
- OpenLink
- TopQuadrant
- Software AG
- Aduna Software
- Protége
- SAPHIRE

IDC Analyze the Future

TU WIEN

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

SEMANTIC WEB COMPANY PROPEL

bmvit

FFG

Companies

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

"Other **companies** are improving the back-end operations of consumer services."

Did companies sustainably adopt SW technologies?
Which verticals/domains?
Who sponsors us?
Which conference sponsors also appear in papers?

SCIENTIFIC AMERICAN

The semantic web in action L Feigenbaum, I Herman, T Hongsermeier, E Neumann, S Stephens Scientific American 297 (6), 90-97, 2007

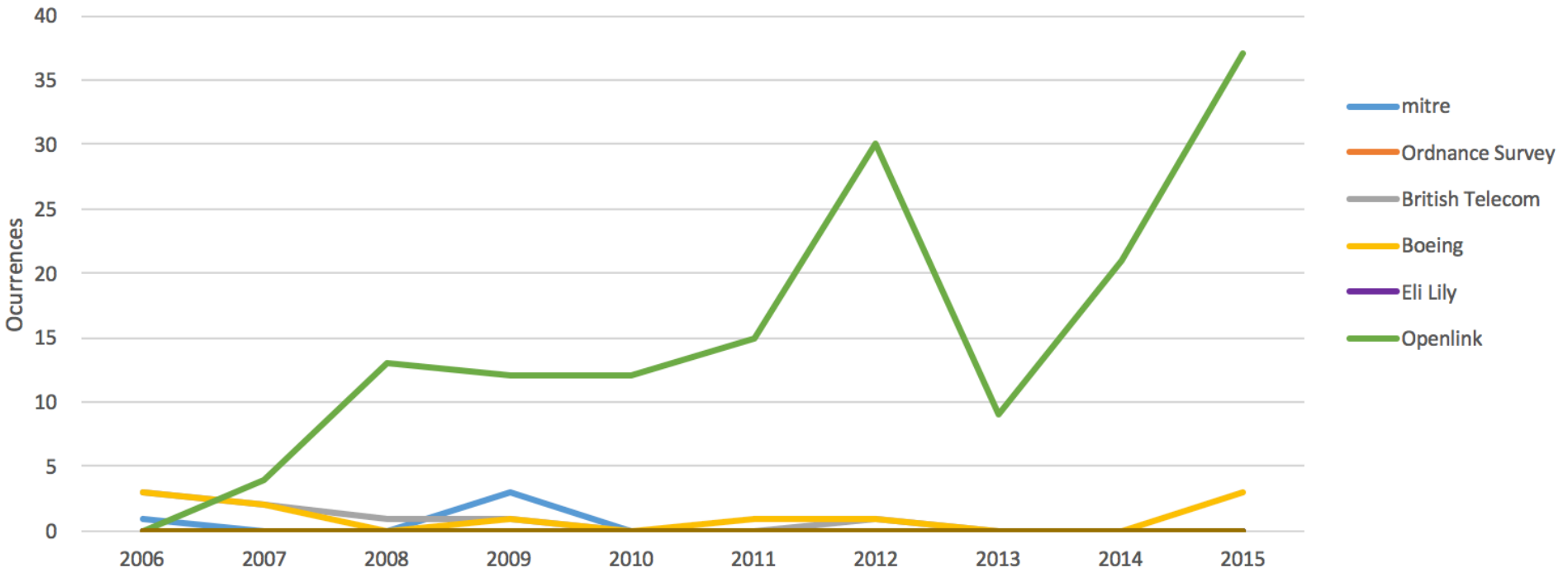
Features - January 19, 2009

The Semantic Web in Action

Corporate applications are well under way, and consumer uses are emerging

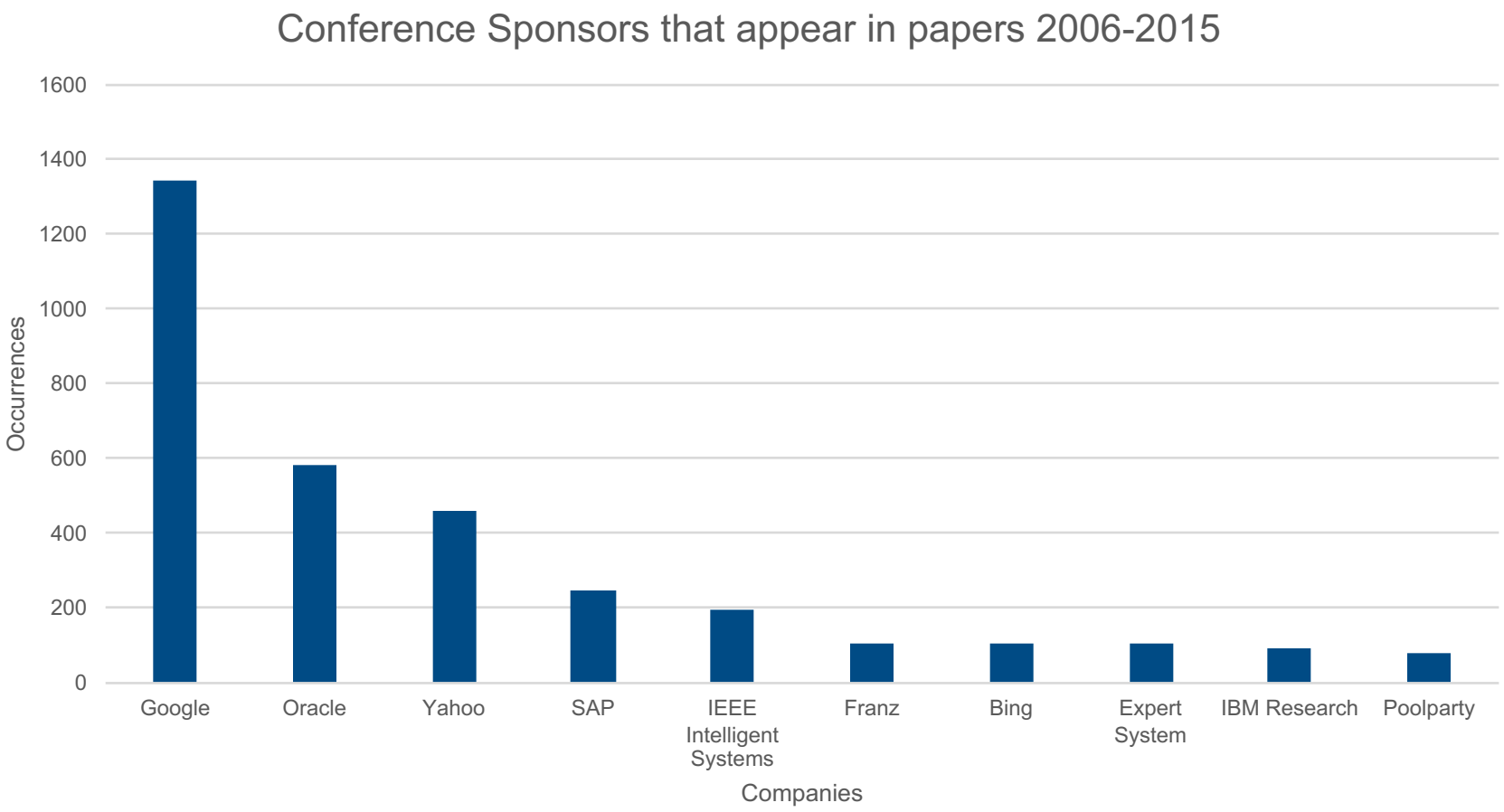
By Lee Feigenbaum, Ivan Herman, Tonya Hongsermeier, Eric Neumann and Susie Stephens

Companies


40
35
30
25
20
15
10
5
0
Ocurrences
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
mitre
Ordnance Survey
British Telecom
Boeing
Eli Lily
Openlink

Companies

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA



IDC Analyze the Future

TU WIEN

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

SEMANTIC WEB COMPANY PROPEL

bmvit

FFG

Domains

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

"some of the most advanced progress is taking place in the **life sciences** and **health care** fields"

What are the primary research domains within our community?

SCIENTIFIC AMERICAN

The semantic web in action L Feigenbaum, I Herman, T Hongsermeier, E Neumann, S Stephens Scientific American 297 (6), 90-97, 2007

Features - January 19, 2009

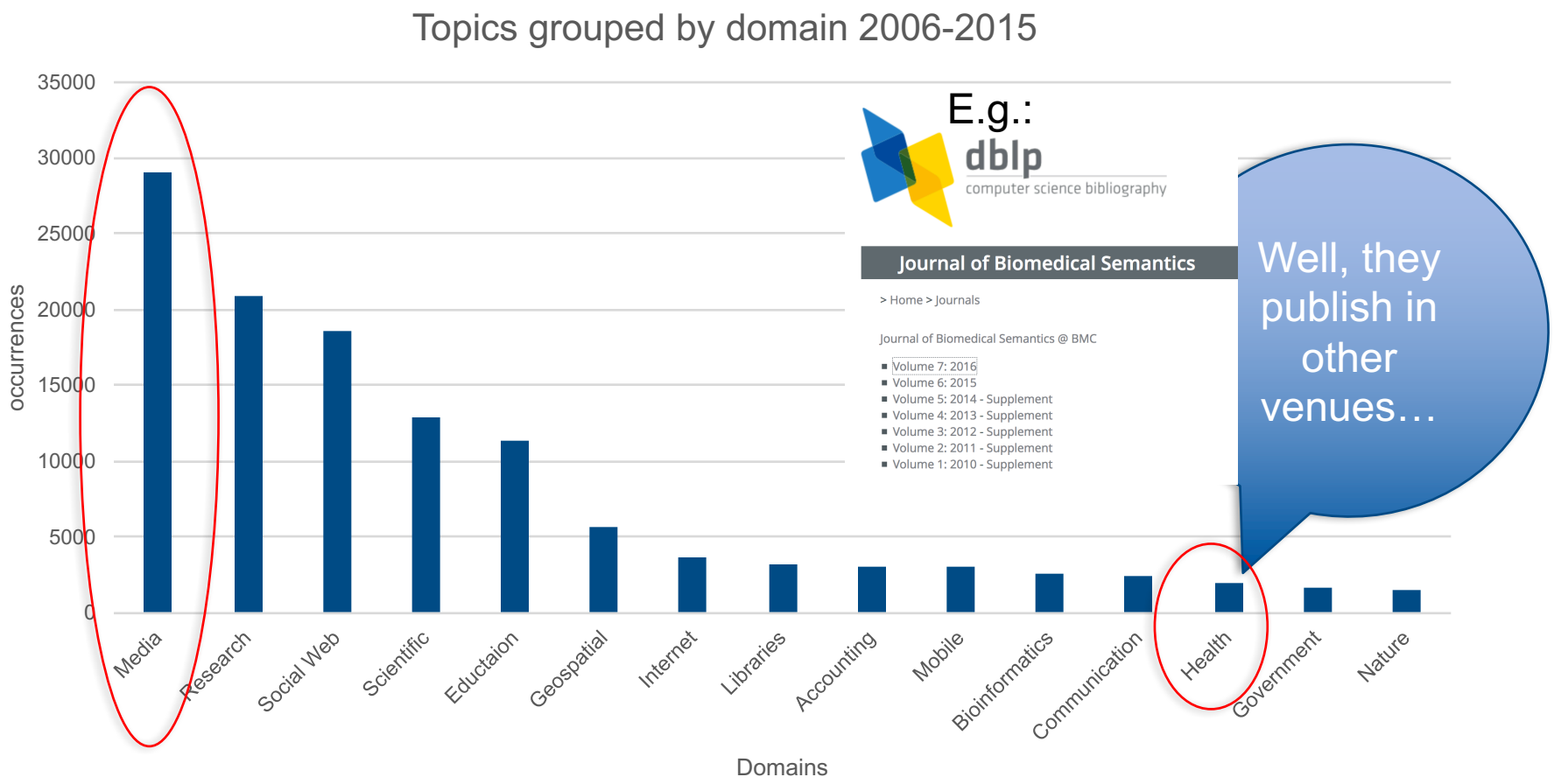
The Semantic Web in Action

Corporate applications are well under way, and consumer uses are emerging

By Lee Feigenbaum, Ivan Herman, Tonya Hongsermeier, Eric Neumann and Susie Stephens

Domains

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA



IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · Semantic Web Company · PROPEL · bmvit · FFG

Applications

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

"And like an iceberg, the tip of this large body of work is emerging in **direct consumer applications**, too."

Can we find evidence of this in our conference and journal paper corpus?

SCIENTIFIC AMERICAN

The semantic web in action L Feigenbaum, I Herman, T Hongsermeier, E Neumann, S Stephens Scientific American 297 (6), 90-97, 2007

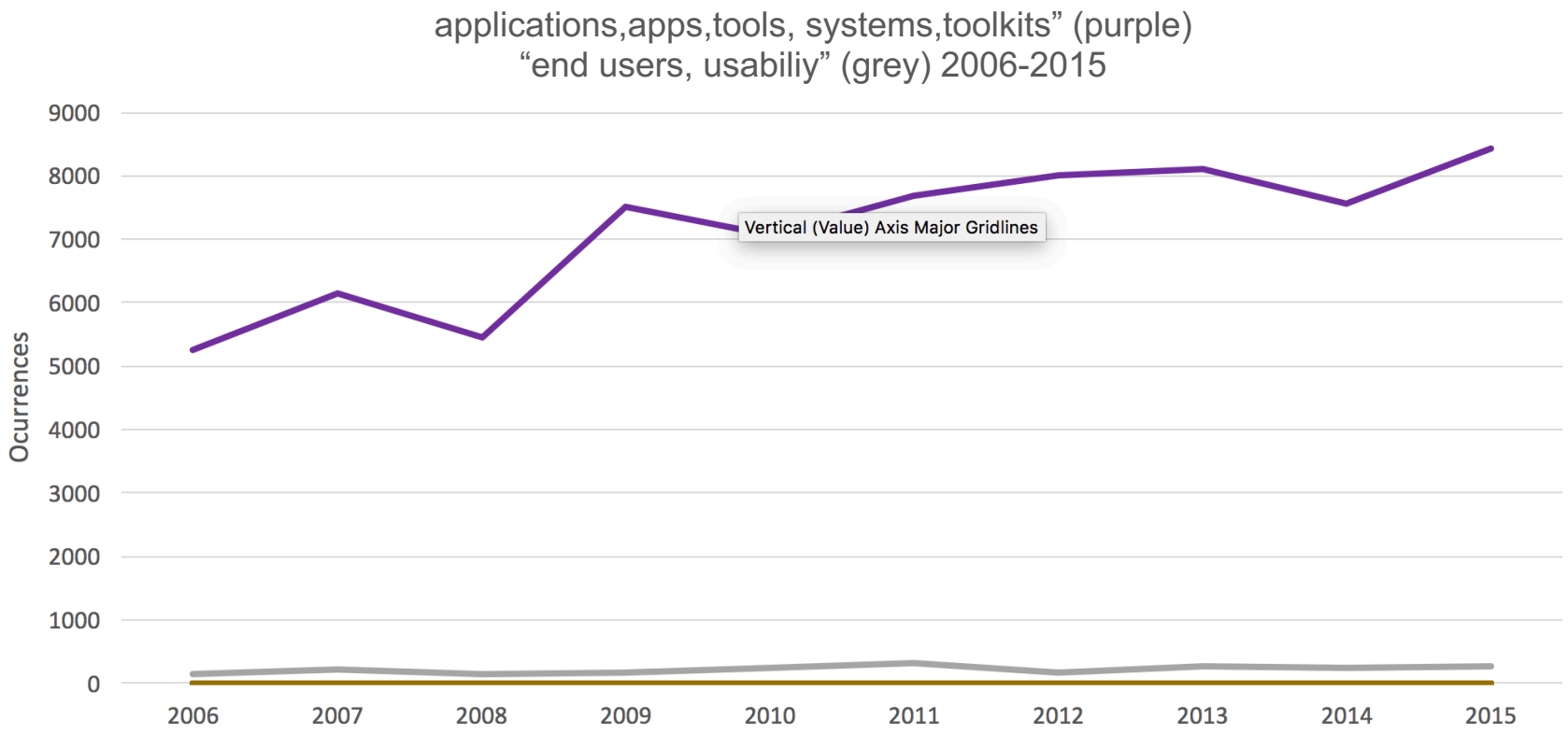
Features - January 19, 2009

The Semantic Web in Action

Corporate applications are well under way, and consumer uses are emerging

By Lee Feigenbaum, Ivan Herman, Tonya Hongsermeier, Eric Neumann and Susie Stephens

End user applications?



Semantic Web/Linked Data over time…

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

V viewpoints

DOI:10.1145/2890489
Abraham Bernstein, James Hendler, and Natalya Noy

Viewpoint

A New Look at the Semantic Web

Seeking to make Web data "smarter" by utilizing a new kind of semantics.

IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY PROPEL
bmvit
FFG

Research in Transition

propel

PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

*"As the early research has transitioned into these **larger**, **more applied** systems, today's Semantic Web research is changing: It builds on the earlier foundations but it has generated a more diverse set of pursuits."*

*"the representations that they used became **less formal** and **precise** than many early Semantic Web researchers had envisioned."*

*"As the semantics, in a sense, becomes more "**shallow**," it could be more widely applicable"*

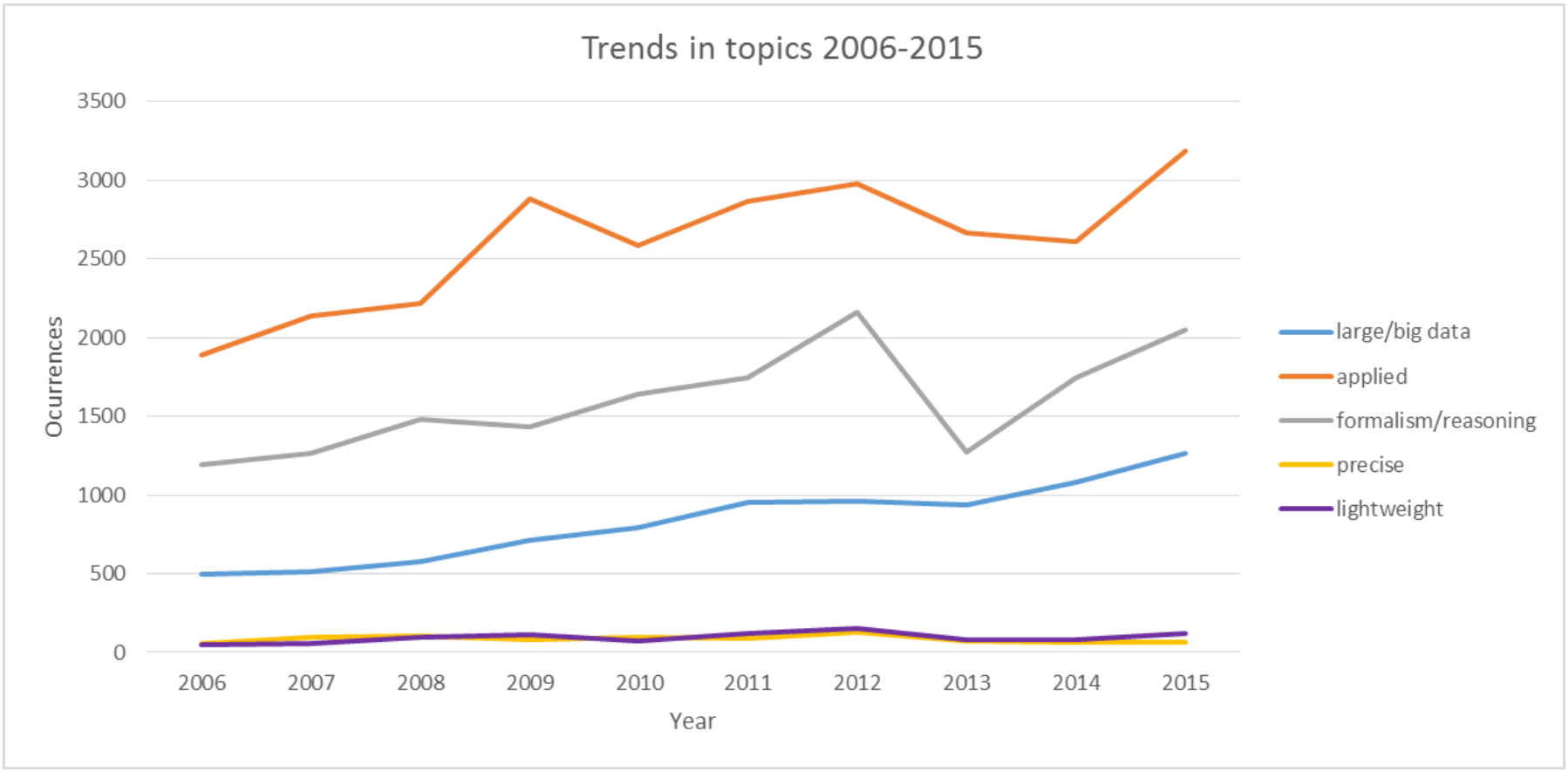
viewpoints

A new look at the semantic web A Bernstein, J Hendler, N Noy
Communications of the ACM 59 (9), 35-37

DOI:10.1145/2890489
Abraham Bernstein, James Hendler, and Natalya Noy

Viewpoint
A New Look at the Semantic Web
Seeking to make Web data "smarter" by utilizing a new kind of semantics.

More applied/lightweight/big data?



The Next 10 Years

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

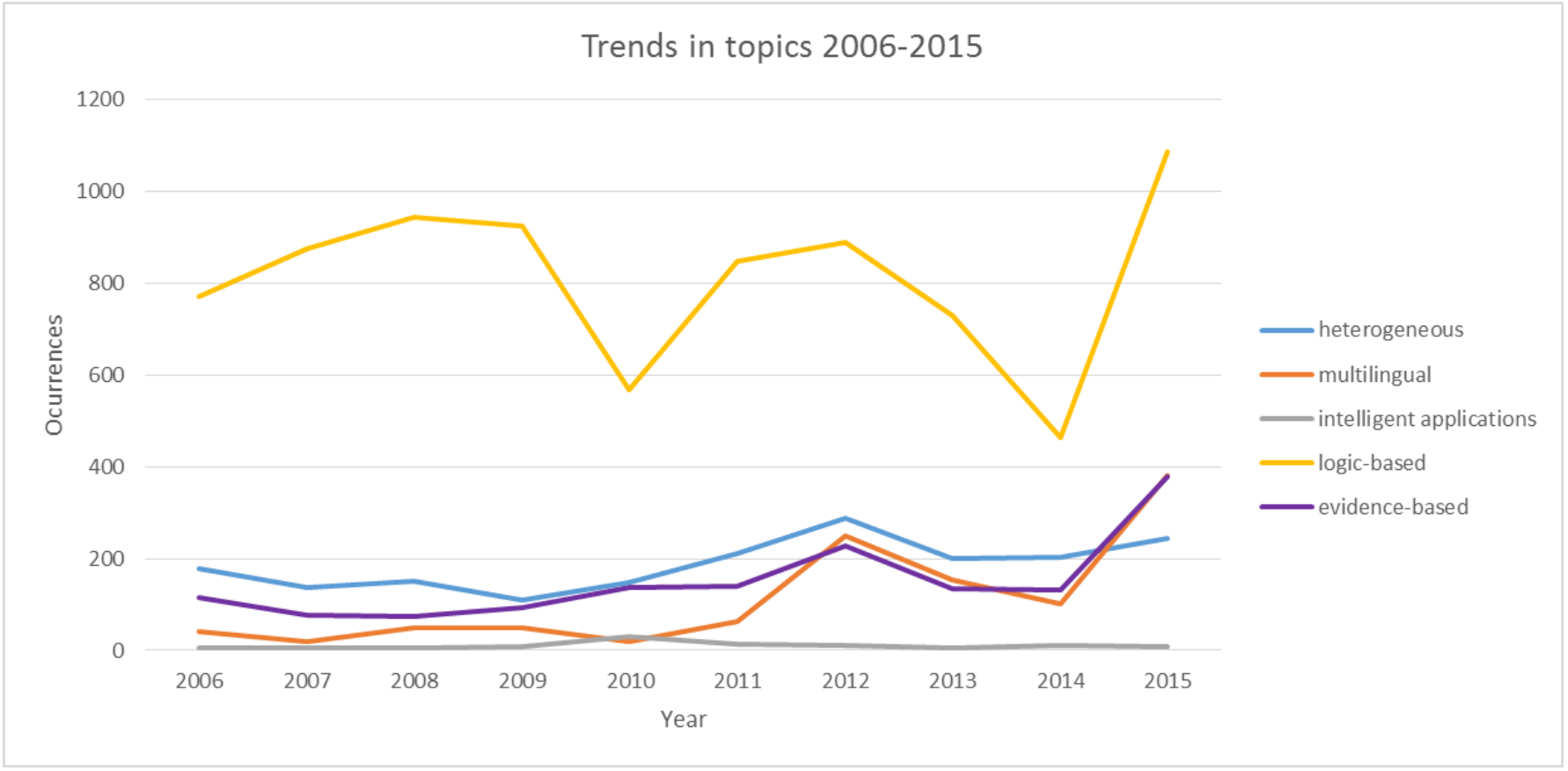
"We believe the objective of the next decade of Semantic Web research is to make this vast **heterogeneous multilingual** data provide the fuel for truly **intelligent applications**."

"relies less on **logic-based** approaches and more on **evidence-based** ones."

A new look at the semantic web A Bernstein, J Hendler, N Noy Communications of the ACM 59 (9), 35-37

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

Less logic-based, more evidence-based?



IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG

Representation and lightweight semantics

- How do we leverage these **diverse representations**?
- How do we coordinate the **diverse components** of structured knowledge that are defined by various parties and that must interact in order to achieve increasingly intelligent behavior?
- How do we define **lightweight**, **needs-based**, "**pay-as-you-go**" approaches for describing knowledge?
- What are the languages and architectures that will provide this knowledge to the increasingly **mobile** and **application-based** Web?

A new look at the semantic web A Bernstein, J Hendler, N Noy
Communications of the ACM 59 (9), 35-37

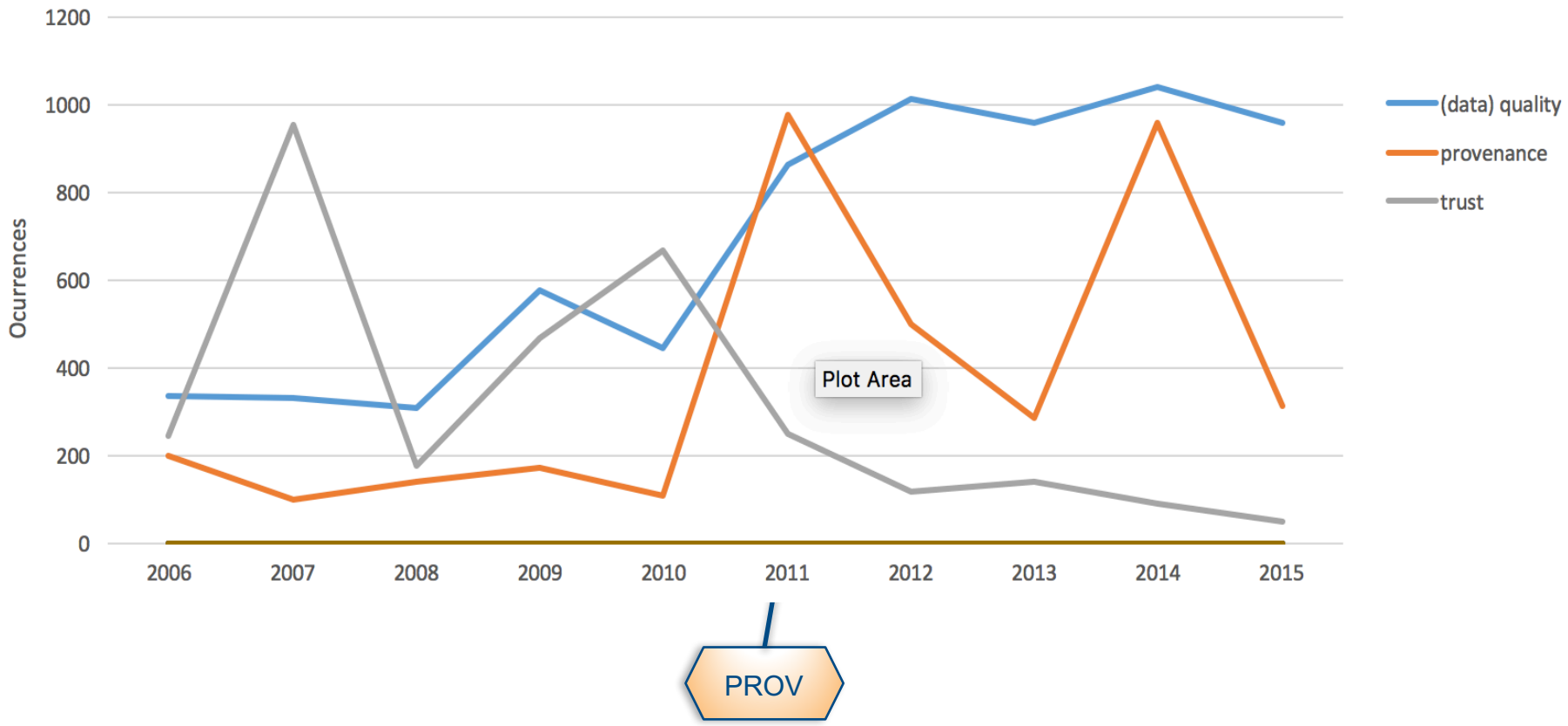
Heterogeneity, quality, and provenance

- How do we integrate **heterogeneous data** and particularly how can we understand which data can be integrated to what degree?
- How can we represent and assess **quality** and **provenance** of the data?
- How do we evaluate whether the quality of a particular source is sufficient for a given task?

A new look at the semantic web A Bernstein, J Hendler, N Noy Communications of the ACM 59 (9), 35-37

Didn’t really find a trend for more lightweight/shallow approaches in the data yet…

Didn't really find a trend for more lightweight/shallow approaches in the data yet… how about data quality & provenance topics?



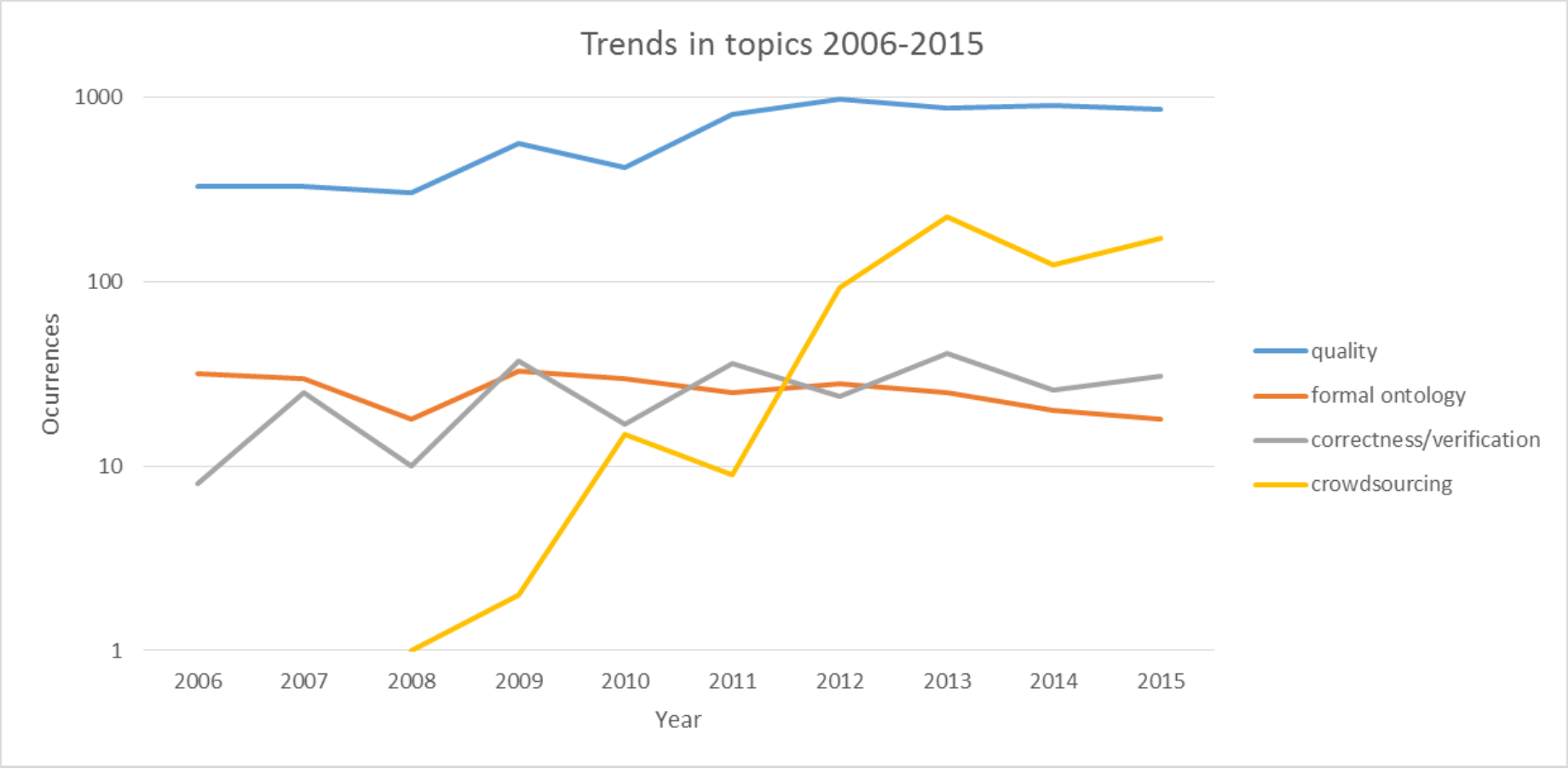
Latent semantics

PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

- How much of the semantics can we learn automatically and what is the **quality** of the resulting knowledge?
- As ontologies are learned or enhanced automatically, what is the very meaning of "**formal ontologies**"?
- How do we develop some notion of approximate **correctness**?
- Do similar or different reasoning mechanisms apply to the ontologies that are extracted in this way?
- How do **crowdsourcing** approaches allow us to capture semantics that may be less precise but more reflective of the collective wisdom?

A new look at the semantic web A Bernstein, J Hendler, N Noy
Communications of the ACM 59 (9), 35-37

More emphasis on data quality, less formal ontologies, more crowdsourcing?



- **CrowdSourcing** becoming hugely popular! (logscale)

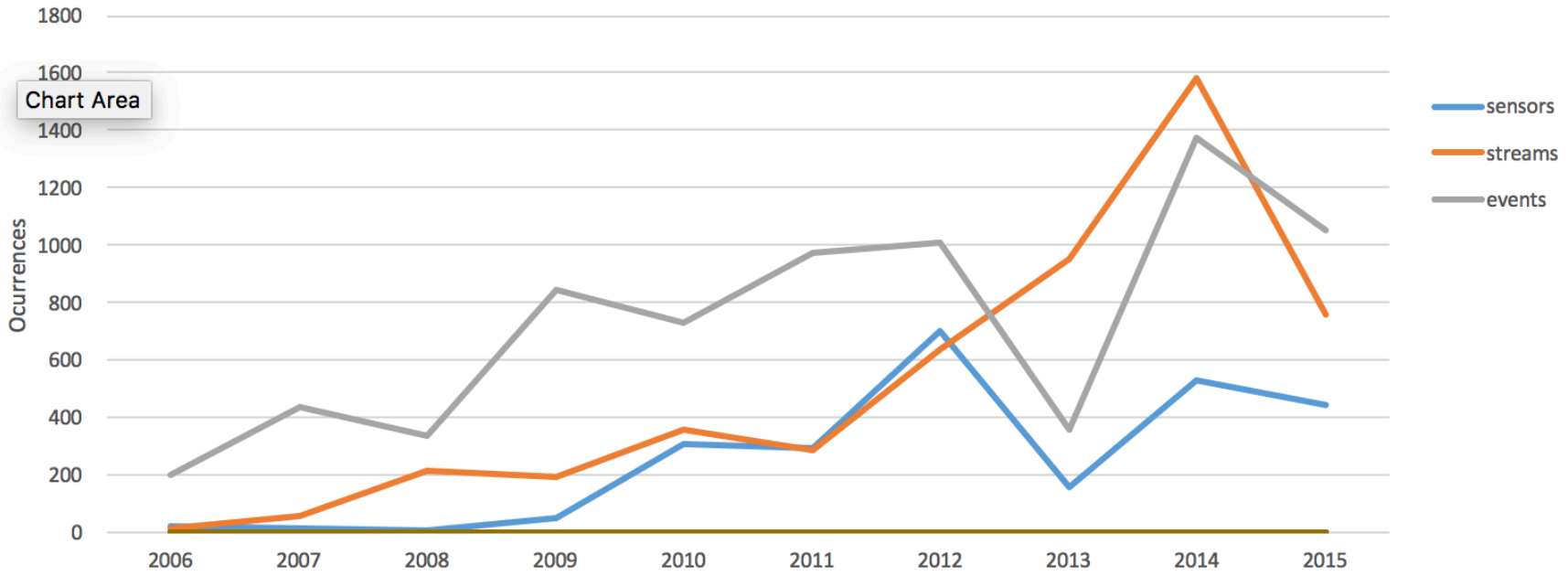
High volume and velocity data

- How do we triage the **data in motion** to determine what to keep and what we may choose, or need, to allow to be lost?
- How can our applications integrate constantly changing **sensor data** with fixed data of long duration and high quality semantic provenance?

A new look at the semantic web A Bernstein, J Hendler, N Noy
Communications of the ACM 59 (9), 35-37

Streams & Sensors

propel
PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA



IDC Analyze the Future
TU WIEN
WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS
SEMANTIC WEB COMPANY
PROPEL
bmvit
FFG

The Roadmap

Roadmap for Enterprise SW/LD?

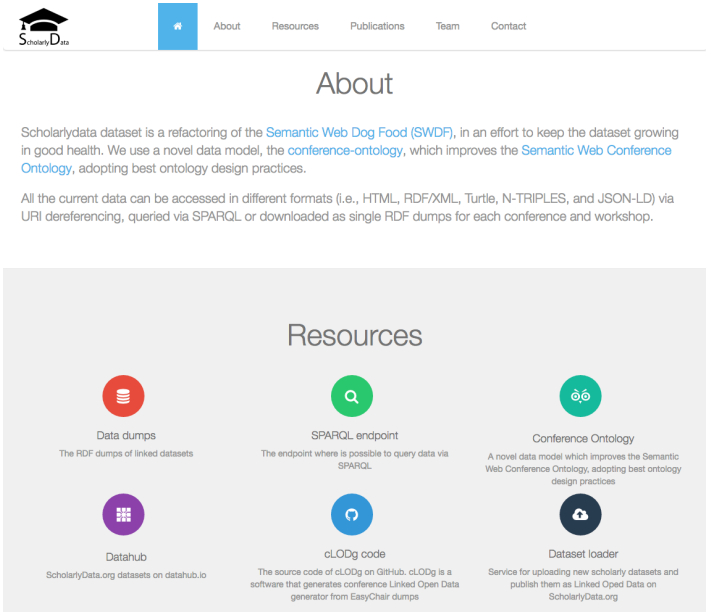
- hmm, I think that I don't have enough data for that as of yet…
- … but gives me a chance to throw in my own personal taste/preferences!!! ;-)

- What I believe to be hot/interesting SW topics (the no-brainers, if you want...)
  1. **Combining Open** and **Closed** Data (Data Security and Privacy)
     - ***Privacy & Policies*** *are the big thing in the age of "Big Linked Data"*
  2. **Archiving** and storage of **temporal data**, efficient indexing and efficient updates (Data Management)
  3. Move into "**not-quite-so-structured**" data, don't focus on just structured RDFdata and non-structured data ...
     - Embrace and deal with de facto standards and formats: *e.g.* ***schema.org, CSV, JSON****, ...*
     - Syntax doesn't matter, we've defined enough languages and syntaxes!
     - *Embrace a "**Cognitive computing**" approach (bridge between purely symbolic and subsymbolic AI*
  4. *Analysing our own research and its impact (...properly)!*

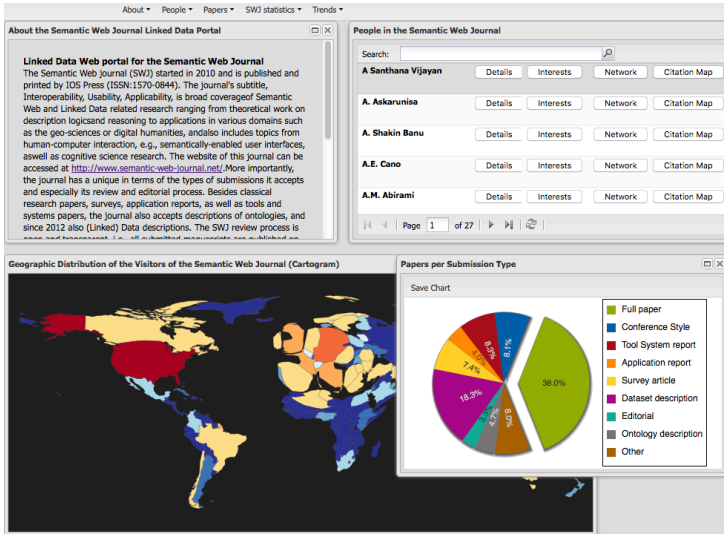
Giving back to the community

propel PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

Where to find data about our community (for example…):



http://www.scholarlydata.org/



http://semantic-web-journal.com/SWJPortal/

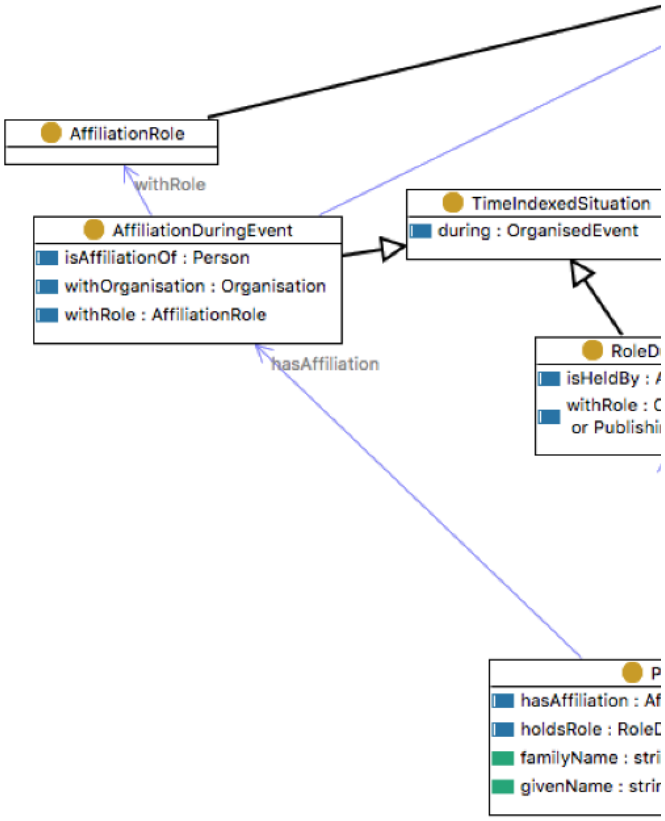
- Thanks to:
- Abraham Bernstein (JWS)
- Pascal Hitzler (SWJ)
- Krzystof Janowisz (SWJ)
- Annalisa Gentile, Andrea Nuzzolese (scholarlydata)
- Francesco Osborne (Rexplore)
- Knud Möller (data.semanticweb.org …)
- ...

http://technologies.kmi.open.ac.uk/rexplore/

We plan to play back our data into http://scholarlydata.org …

propel — PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

1. Classes
2. Object Properties
3. Data Properties
4. Named Individuals
5. Annotation Properties
6. Namespace Declarations



- What we plan* to add:
  - complete/link missing data from all major events
  - keyphrase extraction from fulltexts
  - keyphrase taxonomy
  - Company/Sponsorship data

poolparty

- Still a lot ***TO DO***:
  - respect **skews** such as:
    - higher importance of abstract/title keyphrases
    - No of papers per year (journals)
  - a curated SW keyphrase taxonomy
  - Doing the “usual NLP stuff”
    - improve stemming
    - disambiguate wordsenses

*) *will try, licensing etc. e.g. needs to be clarified!*

Key take-home:

PROPELLING THE POTENTIAL OF ENTERPRISE LINKED DATA IN AUSTRIA

- Stay tuned: PROPEL results will be collected at:
  - https://www.linked-data.at/
- Don't think our **standards** are known by industry ;-)
  - Nor that they cover "Semantic Linked" data…
  - … but standardization is still useful
- We could need some more research in understanding what we actually do and sell our successes!
  - Listen to the "wise old elves"!
  - … but maybe there is a more data-driven, data-analytics-driven way for this
  - Feel free to go ahead and play with the data we collected (soon on http://*scholarlydata.org*)...
  - ...improve our quick attempt!) ... and let me know!!!! ☺

- We shouldn't try to be

IDC Analyze the Future · TU WIEN · WU Wirtschaftsuniversität Wien Vienna University of Economics and Business · SEMANTIC WEB COMPANY PROPEL · bmvit · FFG