

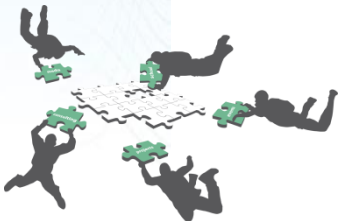
# Auf der Suche nach der besseren Suche

Wie semantische Suche und Semantic Web  
unsere Arbeitsweisen verändert

CON●ECT  
INFORMUNITY

*22. April 2010*

Mag. Andreas Blumauer  
*Semantic Web Company*



# Agenda

1. Wozu Semantische Suche? Prinzipien Semantischer Technologien & des Semantic Web

2. Anwendungen semantischer Technologien

3. Ganz kurze Einführung in das Semantic Web

4. Trends und Conclusio



# Was heißt „Suchen“?

Suchen

Lernen/Navigieren

## (1) Punkt-Genau Suche (Faktenwissen)

z.B. „Eyjafjallajökull“

## (2) Sammeln und Recherchieren

z.B. „volcanoes in Iceland“

## (3) Suche in Assoziationsräumen / Erkennen von Zusammenhängen

z.B. Vulkanausbrüche und Klimaveränderungen

## (4) Fragen & Antworten

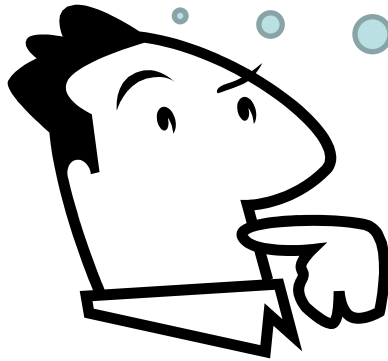
z.B. „Which volcanoes in Iceland erupted in 2010?“

## (5) Serendipity

Benutzer weiß nicht, was er/sie wissen muss → Erkennen neuer Möglichkeiten



# Wozu Semantische Suche?



Eyfafyell.....

isländischer vulkan



## [Vulkane in Island – Wikipedia](#)

Die zahlreichen **Vulkane** Islands haben die Insel im Nordatlantik geschaffen und nachhaltig geprägt. Inhaltsverzeichnis. 1 Geologische Bedingungen; 2 Die ...  
[de.wikipedia.org/wiki/Vulkane\\_in\\_Island](http://de.wikipedia.org/wiki/Vulkane_in_Island) - [Im Cache](#) - [Ähnlich](#)

## [Kreuzworträtsel Hilfe & Lösungen Isländischer -> Isländisches ...](#)

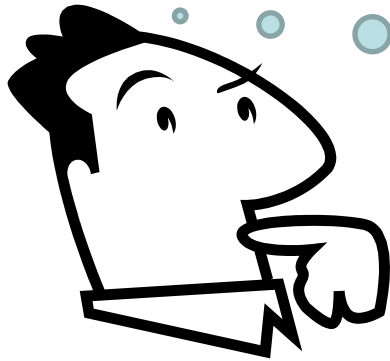
Isländischer. ELDGJA ELDGJA. **isländischer Vulkan isländischer Vulkan**. GULLFOSS GULLFOSS. isländischer Wasserfall isländischer Wasserfall. LAKI LAKI ...  
[www.raetsel-hilfe.de/n-5-isl.html](http://www.raetsel-hilfe.de/n-5-isl.html) - [Ähnlich](#)

## [Isländischer Vulkan spuckt unverändert Asche aus > Kleine Zeitung](#)

16. Apr. 2010 ... Der Vulkanexperte Armann Hoskuldsson von der **isländischen** Universität äußerte die Erwartung, dass der **Vulkan** noch zwei oder mehr Tage aktiv ...  
[www.kleinezeitung.at/.../vulkan/.../islaendischer-vulkan-spuckt-unveraendert-asche.story](http://www.kleinezeitung.at/.../vulkan/.../islaendischer-vulkan-spuckt-unveraendert-asche.story) - [Im Cache](#)



# Wozu Semantische Suche??



Eyfafyell.....

höchste Erhebung – Zentralvulkan ☒	Höhe in m ☒	Gebirgsmassiv – Vulkansystem ☒
<a href="#">Hvannadalshnúkur</a>	2110	<a href="#">Öræfajökull</a>
<a href="#">Bárðarbunga</a>	2010	<a href="#">Bárðarbunga</a> , <a href="#">Vatnajökull</a>
<a href="#">Austur-Kverkfjöll</a>	1933	<a href="#">Kverkfjöll</a>
<a href="#">Snæfell (Ostisland)</a>	1833	<a href="#">Snæfell</a>
<a href="#">Hofsjökull</a>	1765	<a href="#">Hofsjökull</a>
<a href="#">Grímsvötn</a>	1720	<a href="#">Grímsvötn</a> , <a href="#">Vatnajökull</a>
<a href="#">Eyjafjallajökull (siehe auch: <a href="#">Ausbruch des Eyjafjallajökull 2010</a>)</a>	1666	<a href="#">Eyjafjallajökull</a>
<a href="#">Tungnafellsjökull</a>	1535	<a href="#">Tungnafellsjökull</a>
<a href="#">Esja, Esjufjöll</a>	1522	<a href="#">Esjufjöll</a> , <a href="#">Vatnajökull</a>



# Semantische Suche!

The screenshot shows a Bing search results page for the query "icelandic volcano". The search bar at the top contains the text "icelandic volcano" and a magnifying glass icon. Below the search bar, there are navigation links for "Web", "Images", "Videos", "Shopping", "News", "Maps", "More", "MSN", and "Hotmail". The search results are categorized under "Reference" and "Articles from Wikipedia".

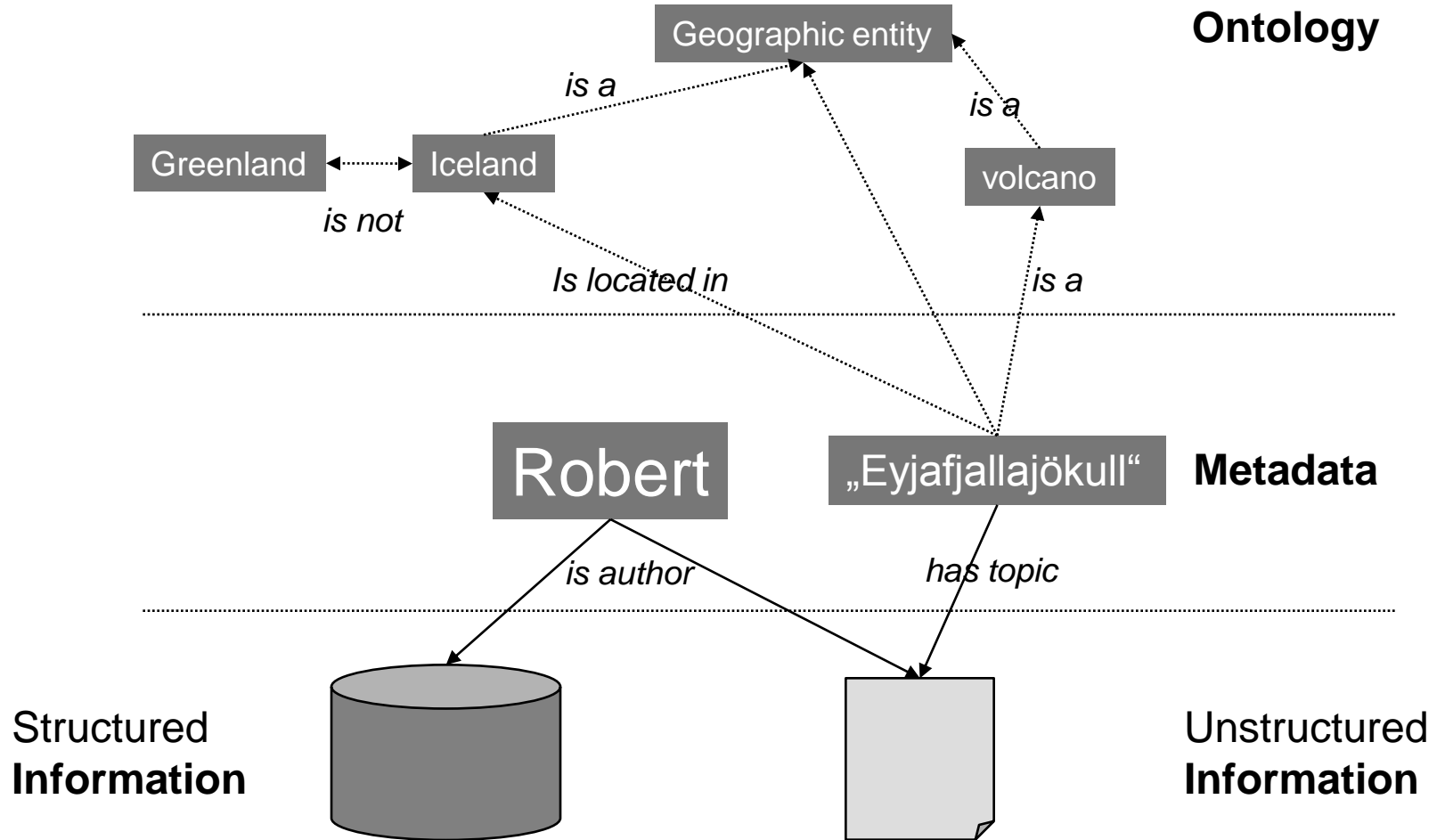
The results include:

- Icelandic Canadian**: A snippet mentioning that Canada has the largest ethnic Icelandic population outside Iceland, with about 88,875 people of Icelandic descent as of the Canada 2006 Census. Many Icelandic Canadians are descendants of people who fled an eruption of the Icelandic volcano Askja in 1875.
- Shield volcano**: A snippet defining a shield volcano as a volcano with shallow-sloping sides. The name derives from a translation of "Skjaldbreiður", an Icelandic shield volcano whose name means "broad shield", from its resemblance to a warrior's shield.
- Hekla**: A snippet stating that Hekla is Iceland's most active volcano, with over 20 eruptions having occurred in and around the volcano since 874. It is the only Icelandic volcano to produce calc-alkaline lavas.
- History of Icelandic**: A snippet discussing the history of the Icelandic language, which began in the 9th century with the settlement of Iceland by settlers from Norway. It mentions various locations and events related to the language's development.
- Heklina**: A snippet mentioning Heklina, the stage name of Stefan Grygelko, who performed as a drag queen in San Francisco's alternative drag show Trannyshack from 1996-2008. Heklina is half-Icelandic and named herself after the Icelandic volcano Hekla.
- Eyjafjallajökull**: A snippet describing Eyjafjallajökull (pronounced [ˈɛjja ˈfjatla ˈjœkvtʃ]) as one of the smaller glaciers of Iceland. It mentions satellite evidence of hot lava flows from an Icelandic volcano (CIMSS Satellite Blog) and that the glacier and volcano are related.
- List of famines**: A snippet providing an incomplete list of known major famines, ordered by date. It mentions the Icelandic Volcano Caused Historic Famine In Egypt, Study Shows, and provides a link to "5th century AD • 7th century AD".
- Placeholder names in different languages**: A snippet discussing placeholder names in various languages other than English. It mentions Häckleffjäll, a commonly used name for a generic remote village, which is actually a synonym for the Icelandic volcano Hekla.
- Hvolsvöllur**: A snippet describing Hvolsvöllur as a small town in the south of Iceland, about 106 km to the east of Reykjavík. It mentions that "Hundreds evacuated after Icelandic volcano erupts, first since 1823".


<http://www.bing.com/reference/>



# Information, Metadata & Ontologies



# Semantische Suche = Präzisere Resultate

Austrian Standards Institute   
Setting Standards. Creating Value.

Suche

**Suche nach Kategorie**

- Ausländische & Int. Normen (155132)
- ÖNORMEN (73749)
- Gesetzliche Vorschriften (7500)
- Österreichische Regelwerke (5242)
- Webseite und Pressemeldungen (3947)
- Zertifizierung (3880)
- Fachliteratur & Software (3861)
- Komitees - Kontakte & Projekte (390)
- Seminare, Services & Solutions (57)

**Suche nach Fachgebiet**

- Elektronik und Telekommunikation (38610)
- Chemie Erweiterung (33195)
  - ☐ Erdöl, Erdölprodukte, Erdgas (4199)
  - ☐ Kunststoffe (3558)
    - Kunststoffe im Allgemeinen (1194)
    - Thermoplaste (1073)
    - Duroplaste (556)
    - Schaumstoffe (384)
    - Verstärkte Kunststoffe (351)
- ☐ Textilprodukte (3355)
- ☐ Produkte der chemischen Industrie (3355)

**Suche**  Suchen Hilfe

Suchergebnis weiter verfeinern Treffer/Seite: 10 ▼

Sortieren nach  Dok. Nummer  Relevanz

Aktuelles Suchergebnis erzielt und verfeinert mit: [> alle Verfeinerungen löschen](#)

1. Suchbegriff: **xml** [> löschen](#)

2. Kategorie: **ÖNORMEN** [> löschen](#)

**!** Verfeinern Sie Ihr Ergebnis mittels Navigatoren (z.B. Klick auf eines der Fachgebiete) im linken Fenster oder mit Klickbox "Suchergebnis weiter verfeinern" und weiteren Suchbegriffen.


**Suchergebnis**

**253836 Treffer** Seite [1](#) [2](#) [3](#) [4](#) ... [25384](#) ▶

**Aktuelles zum Vergaberecht // Mittwoch, 28. April 2010, 9:00 - 17:00 Uhr // Austrian Standards Plus/1020 Wien, Heinestraße 38**

Aktuell + Partner 1004003N:2010 **Preis: : 470 €**

Status: Training



 [> Details / Buchen](#)

---

**Pflegeprozess - Teil 1: Pflgerisches Assessment**

ÖNORM K 1160-1:2010 04 01 **Preis: ab: 25.6 €**

Status: Norm

[> Vorschau anzeigen](#) |  [> Details / Bestellen](#) |  [> Merken](#)

**Login**

E-Mail Adresse/AS-ID:


Passwort:

[> anmelden](#)

[> Passwort vergessen?](#)

[> Registrierung](#)


[> Rückruf](#)

**Warenkorb** 

Ihr Warenkorb ist leer.

**Suchverlauf**

xml, ÖNORMEN

 [> Suchabfrage per E-Mail versenden](#)

**Topseller**

[ÖVE/ÖNORM E 8001-1/A5](#)

[ÖNORM M 6219-1](#)

[ONR 22251](#)

[ÖNORM EN ISO 10456](#)

**FAQs**

<https://www.astandis.at/shopV5/search/Search.action>



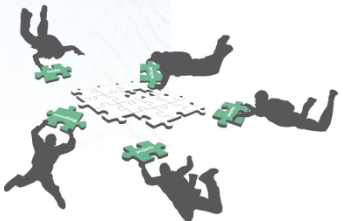
# Agenda

1. Wozu Semantische Suche? Prinzipien Semantischer Technologien & des Semantic Web







2. Anwendungen semantischer Technologien

3. Ganz kurze Einführung in das Semantic Web

4. Trends und Conclusio



# Anwendungsszenarien - Überblick

1	Moderated / Faceted Search		Creation of complex queries
2	Navigation & Browsing		Orientation in Documents
3	Similarity (Recommender Systeme)		Understanding relations
4	Learning (Q&A-Machines)		Discovering facts instead of Documents
5	Dossiers & Trends		Mashing Up Information
6	Sentiment Detection		Capturing moods about certain topics



# ① Moderated / Faceted Search

- **Challenge**

Keywords are not sufficient to define complex queries

- **Solution**

Categories allow to bring order to a collection of documents within predefined contexts

- **Demo**

<http://www.ecoi.net>

<http://www.emteba.at>

<http://www.gopubmed.com>



## ② Navigation & Browsing

- **Challenge**

Working through large amounts of documents is energy and time intensive

- **Solution**

By Entity Extraction orientation and reading speed is enhanced.

- **Demo**

<https://addons.mozilla.org/de/firefox/addon/3999>

<http://www.rsc.org/publishing/journals/projectprospect/index.asp>



## ③ Similarity (Recommender Systeme)

- **Challenge**

Traditional search techniques just make implicit statements about the relation between documents.

- **Solution**

By applying a semantic index the similarity of documents can be calculated with respect to different viewpoints.

- **Demo**

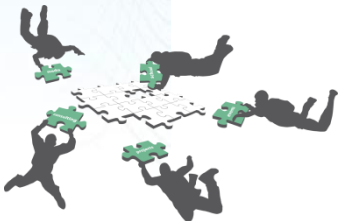
<http://poolparty.punkt.at/>

<http://bookmarks.insuggest.com/>



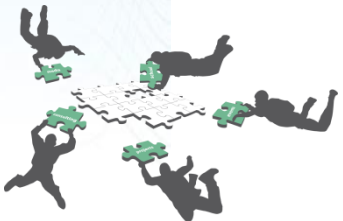
## ④ Learning (Q&A Machines)

- **Challenge**  
Traditional search techniques just reference documents but not facts.
- **Solution**  
A Semantic Index supports the extraction of facts
- **Demo**  
<http://www.wolframalpha.com>  
<http://www.ask.com>  
<http://start.csail.mit.edu/>



## 5 Dossiers & Trends

- **Challenge**  
How can I collect information that is distributed over several databases?
- **Solution**  
Mash it up (and use RDF)!
- **Demo**  
<http://trends.ellerdale.com/>  
<http://sig.ma>  
<http://www.silobreaker.com/>  
<http://topics.nytimes.com/>



## ⑥ Sentiment Detection

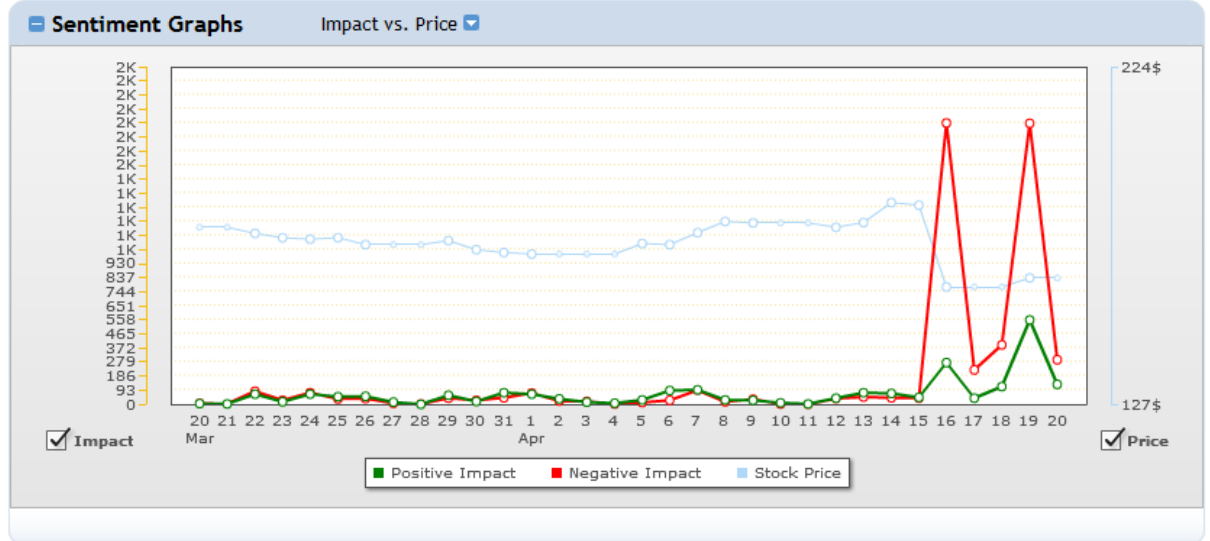
- **Challenge**  
Traditional search techniques tell us little about the moods & motives of people
- **Solution**  
Mine for information that give statements about the sentiment towards anything.
- **Demo**  
<http://www.wingify.com/labs.php>



# Sentiment analysis - further example

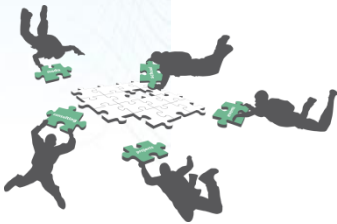
Name:  Analyze

★ Goldman Sachs Group Inc (The) (NYSE:GS) W M 3M 6M Y From: 3/20/2010 To: 4/20/2010



**Articles** Sources Score Range: -1 to 1

Date	Score	Title	Original article
04/20/2010	0 (Pos: 5, Neg: 5)	UK regulator begins Goldman Sachs probe (AP)	UK regulator begins Goldman Sachs probe (AP)
04/20/2010	-0.55 (Pos: 1, Neg: 7)	UK Markets Regulator Launches Goldman Probe (at CNBC)	Search for share prices Get Quotes Finance Search Tue, Apr 20, 2010, 6:38AM ET - U.S. Markets open in 2 hours and 52 minutes. People enter Goldman Sachs headquarters, Monday, April 19, 2010, in New York. Stocks are falling on
04/20/2010	0.55 (Pos: 7, Neg: 1)	European shares extend gains as banks recover (at Reuters)	concern about the fallout over Goldman Sachs



# Agenda

1. Wozu Semantische Suche? Prinzipien Semantischer Technologien & des Semantic Web

2. Anwendungen semantischer Technologien

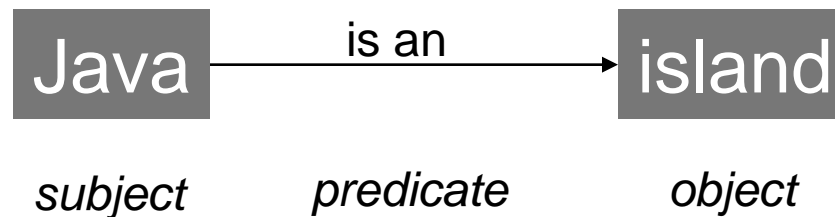
3. Ganz kurze Einführung in das Semantic Web

4. Trends und Conclusio



# Das „Triple“ - die „Grundwährung“ im Semantic Web

- Das Semantic Web ist ein globales Web von Fakten
- Fakten werden hier als „Triples“ gespeichert
- Jedes Triple gehorcht diesem Muster: *subject, predicate, object*



# Uniform Resource Identifier (URI)

- Strings können mehrdeutig sein, z.B. „*Java = island = programming language*“
- Um „Java“ eindeutig bestimmen zu können, benötigen wir so genannte URIs:



 <http://dbpedia.org/resource/Java>

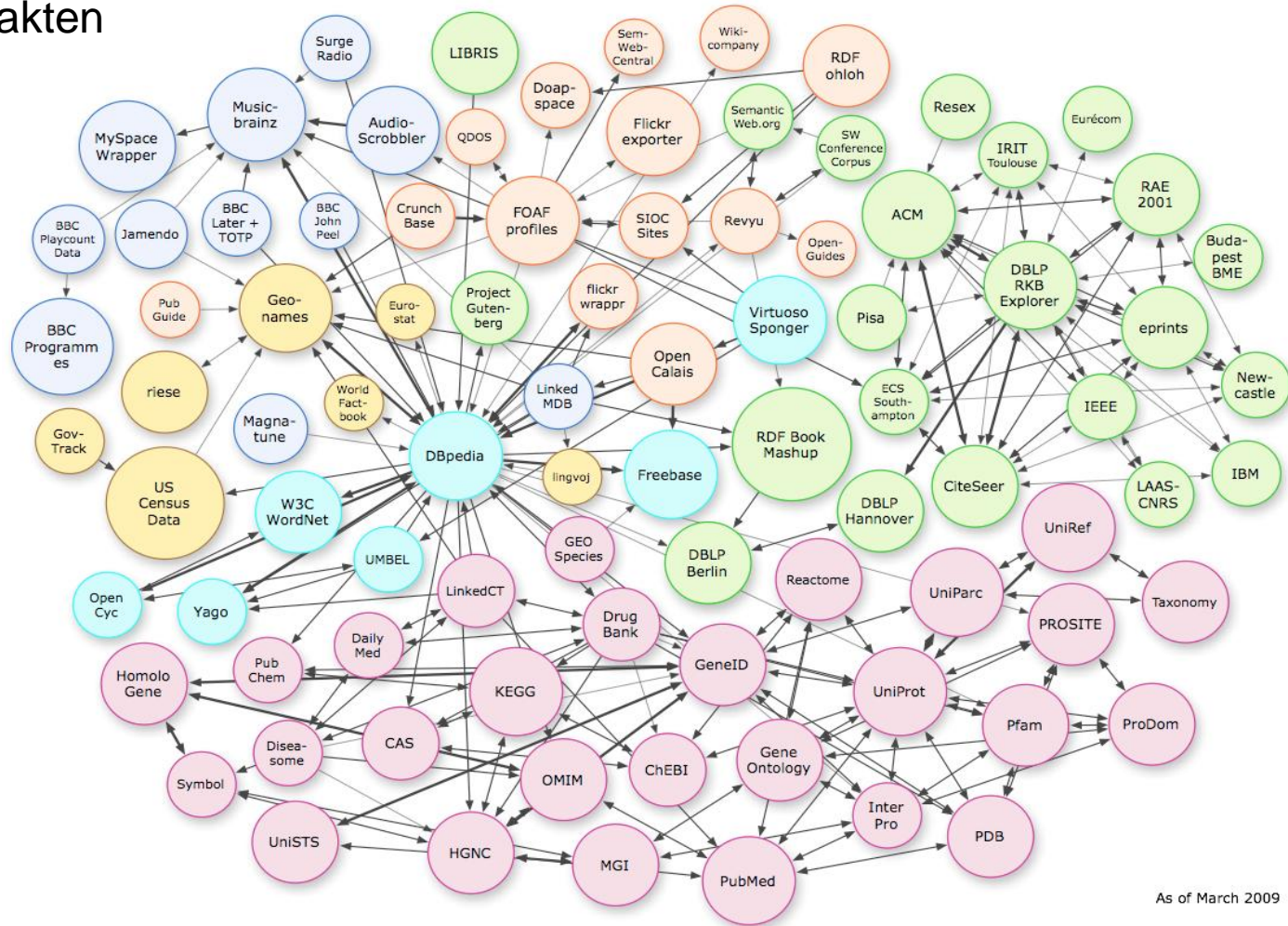


 [http://dbpedia.org/resource/Java\\_\(programming\\_language\)](http://dbpedia.org/resource/Java_(programming_language))



# Ja, das Semantic Web existiert ... und es ist riesig!

LOD Cloud  
~ 4.7 Mrd. Fakten



As of March 2009



# Agenda

1. Wozu Semantische Suche? Prinzipien Semantischer Technologien & des Semantic Web

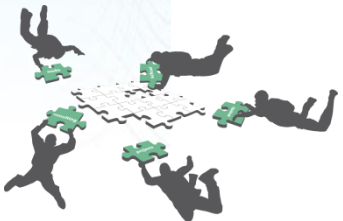
2. Anwendungen semantischer Technologien

3. Ganz kurze Einführung in das Semantic Web

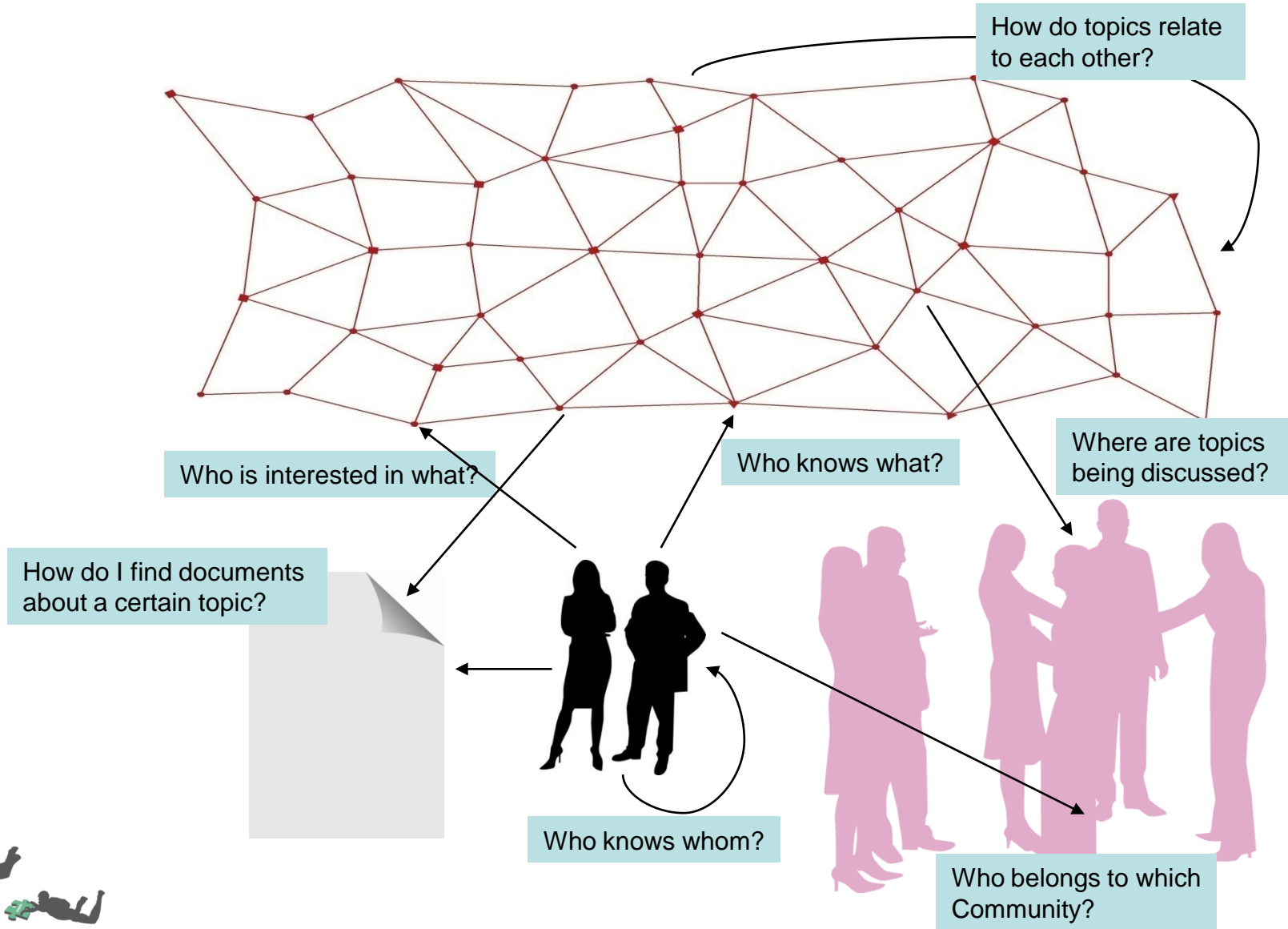
4. Trends und Conclusio



- **Search Engines** start combining text mining with ontologies
- Text Mining available as **Web-Service**
- **Enterprise Search** starts to augment enterprise data with data from outside sources (LOD Cloud)



# It's not only about „searching“ but also about discovering „relations“



- Kostenlose Webinare der SWC



- Nächste Termine
  - Donnerstag, 6.5.2010  
09:00 - 10:30
  - Donnerstag, 20.5.2010  
09:00 - 10:30
- Bitte melden Sie sich auf <http://www.semantic-web.at/> an!



# Zum sechsten Mal:

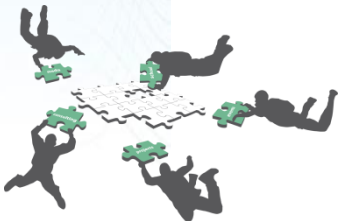
## I-semantics

1. - 3. September 2010, Graz



<http://i-semantics.at/>

co-located with: 



## Andreas Blumauer Semantic Web Company

[a.blumauer@semantic-web.at](mailto:a.blumauer@semantic-web.at)



Semantic Web Company GmbH  
Lerchenfelder Gürtel 43  
A-1160 Wien

Tel. +43 1 402 12 35  
[office@semantic-web.at](mailto:office@semantic-web.at)

