Big Data, Semantic Web, and EarthCube

Pascal Hitzler
Kno.e.sis Center
Wright State University, Dayton, OH
http://www.pascal-hitzler.de/

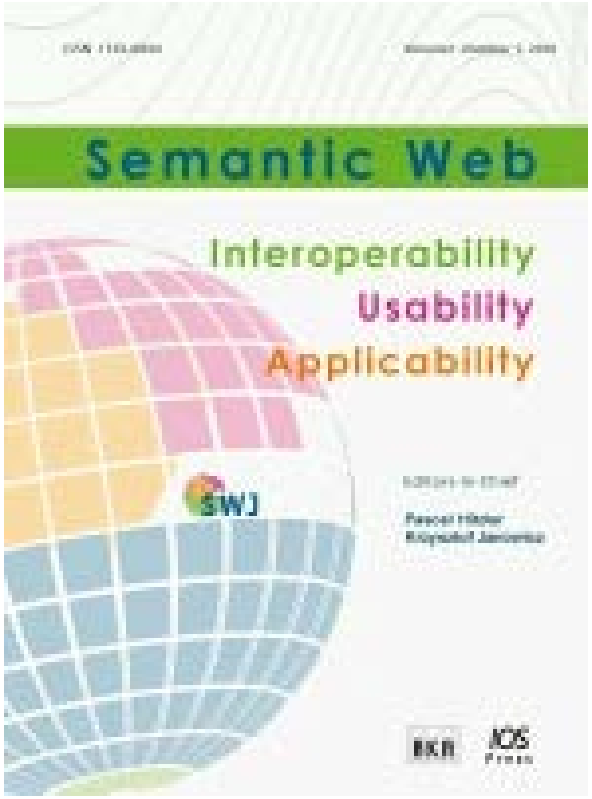
Semantic Web journal

- **EiCs:** **Pascal Hitzler**
  Krzysztof Janowicz

- **New journal with significant initial uptake.**

- **We very much welcome contributions at the "rim" of traditional Semantic Web research – e.g., work which is strongly inspired by a different field.**

- **Non-standard (open & transparent) review process.**



- **http://www.semantic-web-journal.net/**

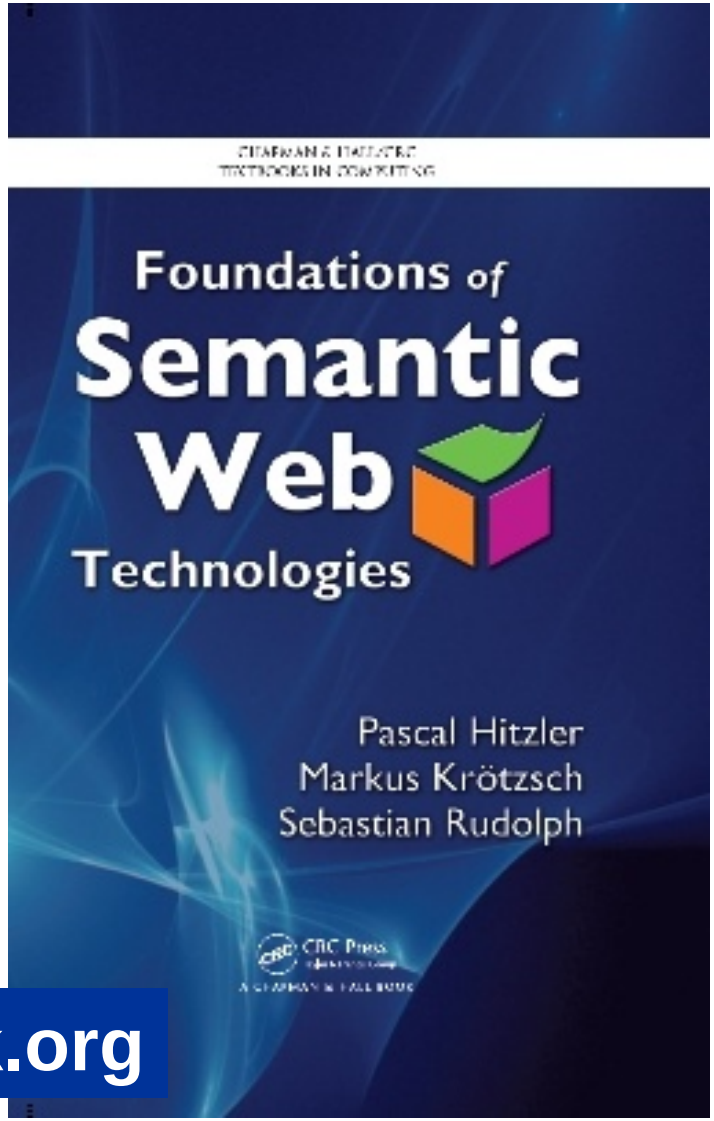
Textbook

Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph

Foundations of Semantic Web Technologies

Chapman & Hall/CRC, 2010

Choice Magazine Outstanding Academic Title 2010 (one out of seven in Information & Computer Science)



http://www.semantic-web-book.org

Contents

- **Big Data**
- **Semantic Web**
- **EarthCube**

Big Data

ITBUSINESSEDGE | An IT Business Edge Site

ITBUSINESSEDGE

Welcome, Guest
Log in | Register

Google Apps for Business
Free 30 day trial
Learn more

Home | IT Projects | Blogs | IT Downloads | White Papers | Newsletters | Google™ Custom Search | Search

Business Alignment & Management | Business Integration | Governance | Infrastructure | Mobile Technology | Security | Sourcing | Vendors & Markets | More

Home → Blogs → Charting Your IT Career → Gartner: Big Data Will Generate 6 Million U.S. Jobs By 2015

Like us:

Gartner: Big Data Will Generate 6 Million U.S. Jobs by 2015

Susan Hall | CHARTING YOUR IT CAREER | 23 OCT, 2012

Share

Gartner predicts that 4.4 million IT **jobs will be created to support Big Data** by 2015, with 1.9 million of them to be in the United States.

In addition, every Big Data-related role in the United States will create employment for three people outside of IT, pushing the total to 6 million U.S. jobs, Peter Sondergaard, senior vice president at Gartner and global head of research, told those attending the Gartner Symposium/ITxpo. He said:

> But there is a challenge. There is **not enough talent** in the industry. Our public and private education systems are failing us. Therefore, only one-third of the IT jobs will be filled. Data experts will be a scarce, valuable commodity," he said. "IT leaders will need immediate focus on how their organization develops and attracts the skills required. These jobs will be needed to grow your business. These jobs are the future of the new information economy.

Though I don't follow Sondergaard's math, we know there's a **shortage of analytics talent for Big Data** and for **engineering talent as well**

Related Content

Topic : Data Management
Data management involves a variety of tasks involved with the full data lifecycle

Blog : Six Reasons IT Can Be Thankful on Thursday

Article : The Fast Evolution - and Transformation - of Mobile Device Management

BuyerZone 2012 BEST of the web

Drive Calendar Gmail

Start Free Trial

Google Apps for Business

Subscribe to our Newsletters

Sign up now and get the best business technology insights

Big Data

Big Data is characterized not only by the enormous volume or the velocity of its generation but also by the heterogeneity, diversity and complexity of the data.

Suzi Iacono, source: http://community.topcoder.com/coeci/nitrd/

- **volume: the sheer size of the data**
- **velocity: new data is added at breathtaking speed**
- **variety: different formats and different perspectives**

Sometimes mentioned:

- **value: how useful is the data?**
- **veracity: how good/reliable is the data?**

however, these can also be subsumed under "variety".

Big Data: My take

Variety can be handled with existing methods if volume and velocity are small.

Big Data research is primarily about methods for handling variety if volume and velocity are so high that existing methods fail.

Linked Data: Volume

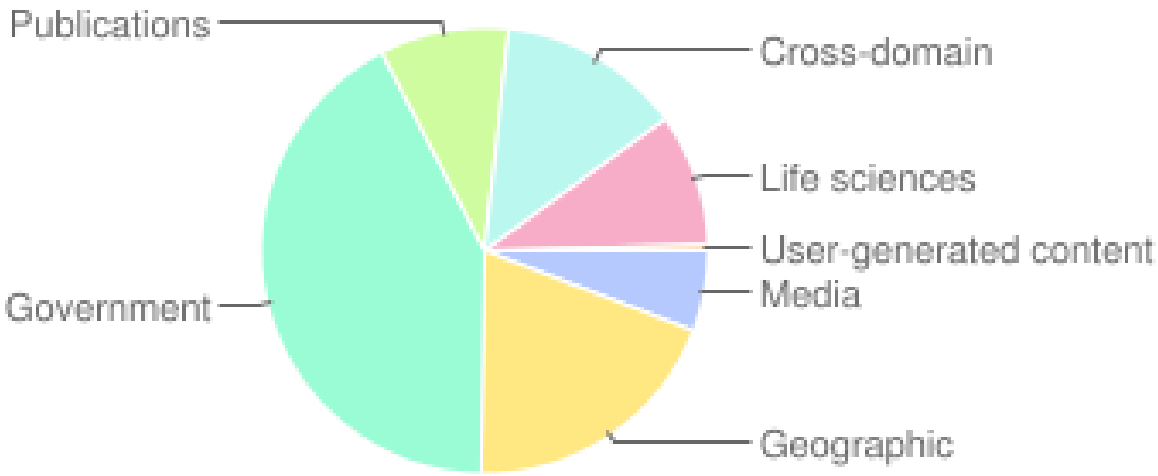
Number of Datasets

2011-09-19	**295**
2010-09-22	**203**
2009-07-14	**95**
2008-09-18	**45**
2007-10-08	**25**
2007-05-01	**12**

Number of triples (Sept 2011)

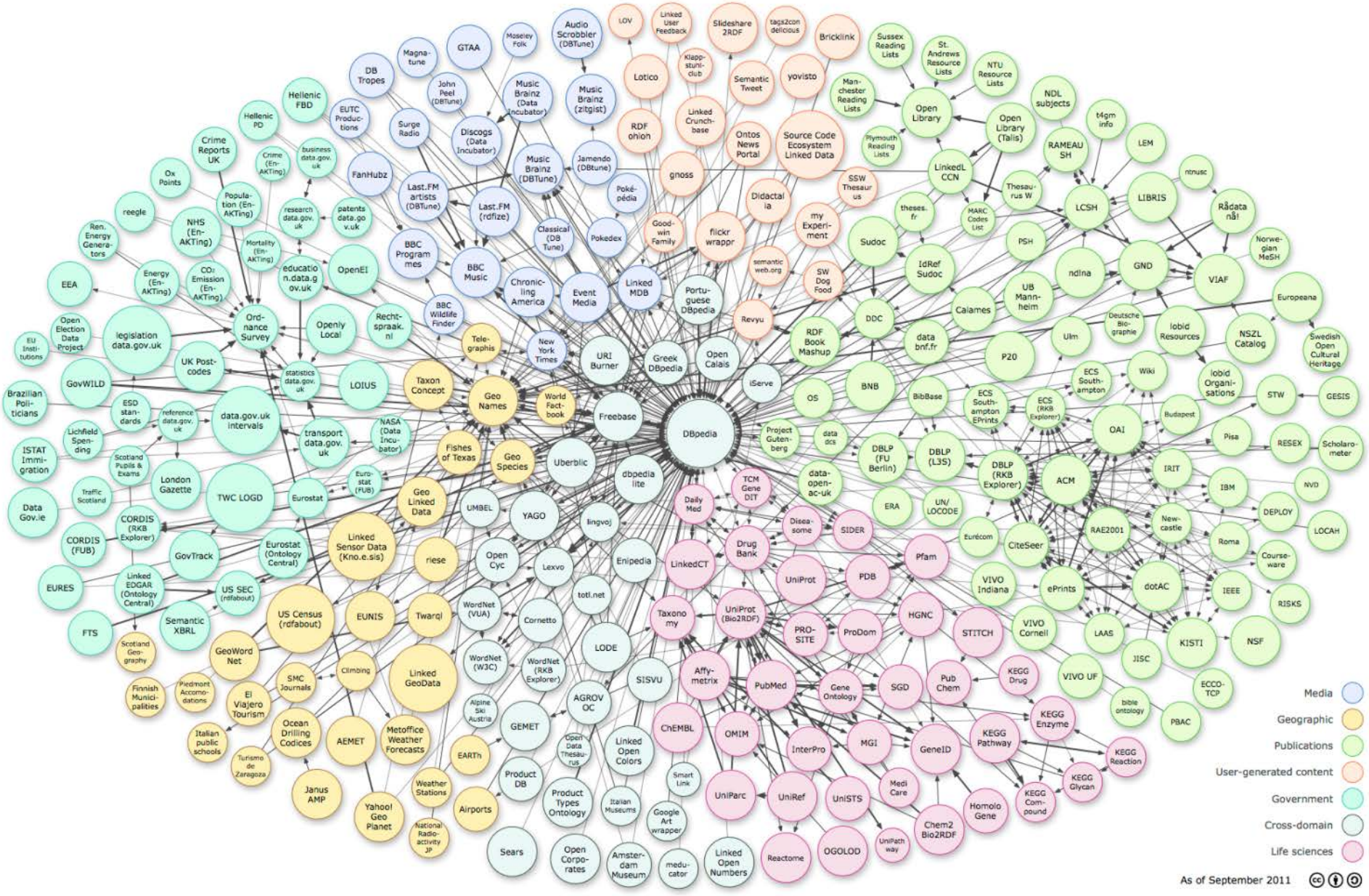
31,634,213,770

with 503,998,829 out-links



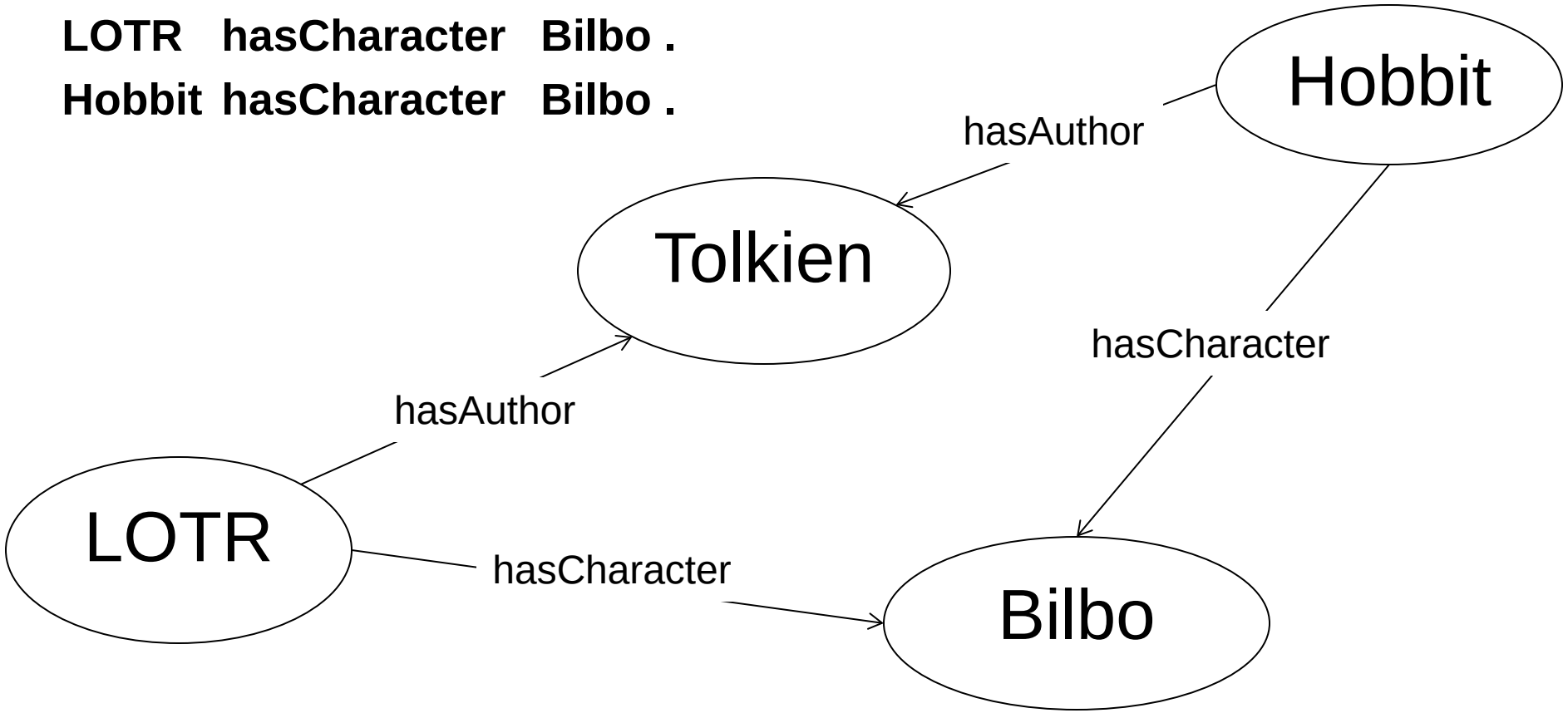
From http://www4.wiwiss.fu-berlin.de/lodcloud/state/

Linked Data 2011



Information as RDF triples / graph

LOTR hasAuthor Tolkien .
Hobbit hasAuthor Tolkien .
LOTR hasCharacter Bilbo .
Hobbit hasCharacter Bilbo .



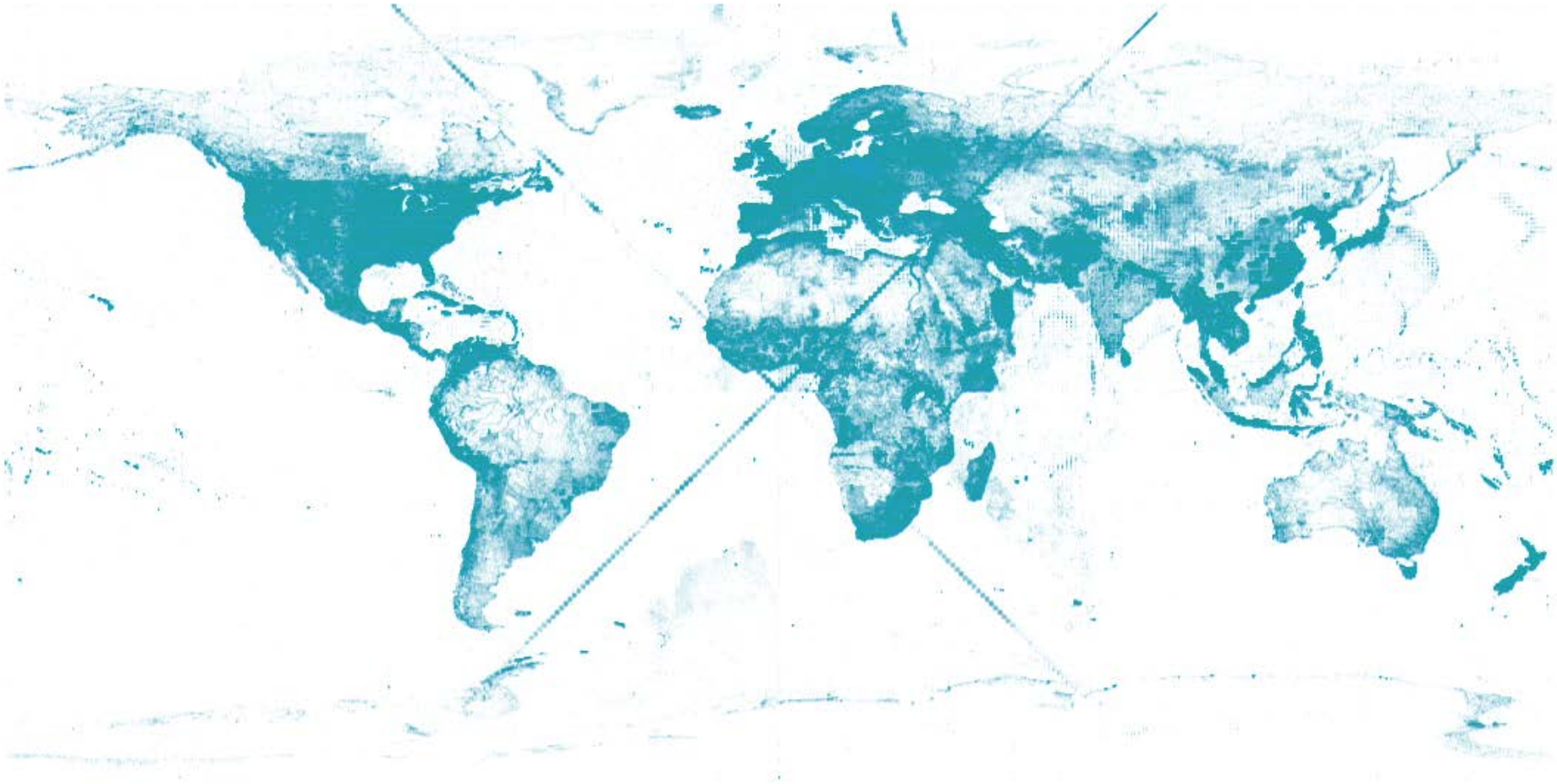
DBpedia: LOTR page

dbpedia-owl:thumbnail	▪ http://upload.wikimedia.org/wikipedia/commons/thumb/6/62/Jrrt_lotr_cover_design.jpg/200px-Jrrt_lotr_cover_design.jpg
dbpedia-owl:wikiPageExternalLink	▪ http://lotr.wikia.com ▪ http://www.glyphweb.com/arda/ ▪ http://www.tolkienlibrary.com/ ▪ http://www.tolkien.co.uk/ ▪ http://www.houghtonmifflinbooks.com/features/lordoftheringstrilogy/
dbpprop:author	▪ dbpedia:J._R._R._Tolkien
dbpprop:books	▪ dbpedia:The_Two_Towers ▪ dbpedia:The_Return_of_the_King ▪ dbpedia:The_Fellowship_of_the_Ring ▪ "Volumes:"
dbpprop:country	▪ England
dbpprop:expiry	▪ 20 (xsd:integer)
dbpprop:genre	▪ dbpedia:Adventure_novel ▪ dbpedia:High_fantasy
dbpprop:hasPhotoCollection	▪ http://www4.wiwiss.fu-berlin.de/flickrwrappr/photos/The_Lord_of_the_Rings
dbpprop:imageCaption	▪ Tolkien's own cover designs for the three volumes
dbpprop:language	▪ English
dbpprop:mediaType	▪ Print
dbpprop:name	▪ The Lord of the Rings
dbpprop:pages	▪ 1216 (xsd:integer)
dbpprop:precededBy	▪ dbpedia:The_Hobbit
dbpprop:pubDate	▪ 21 (xsd:integer)
dbpprop:publisher	▪ dbpedia:Allen_&_Unwin
dbpprop:small	▪ yes
dbpprop:wikiPageUsesTemplate	▪ dbpedia:Template:Infobox_book_series ▪ dbpedia:Template:Pp-vandalism
dcterms:subject	▪ category:Monomyths ▪ category:High_fantasy_novels ▪ category:Middle-earth_books ▪ category:British_fantasy_novels ▪ category:Fantasy_books_by_series ▪ category:1950s_fantasy_novels ▪ category:Sequel_novels ▪ category:The_Lord_of_the_Rings ▪ category:English_novels

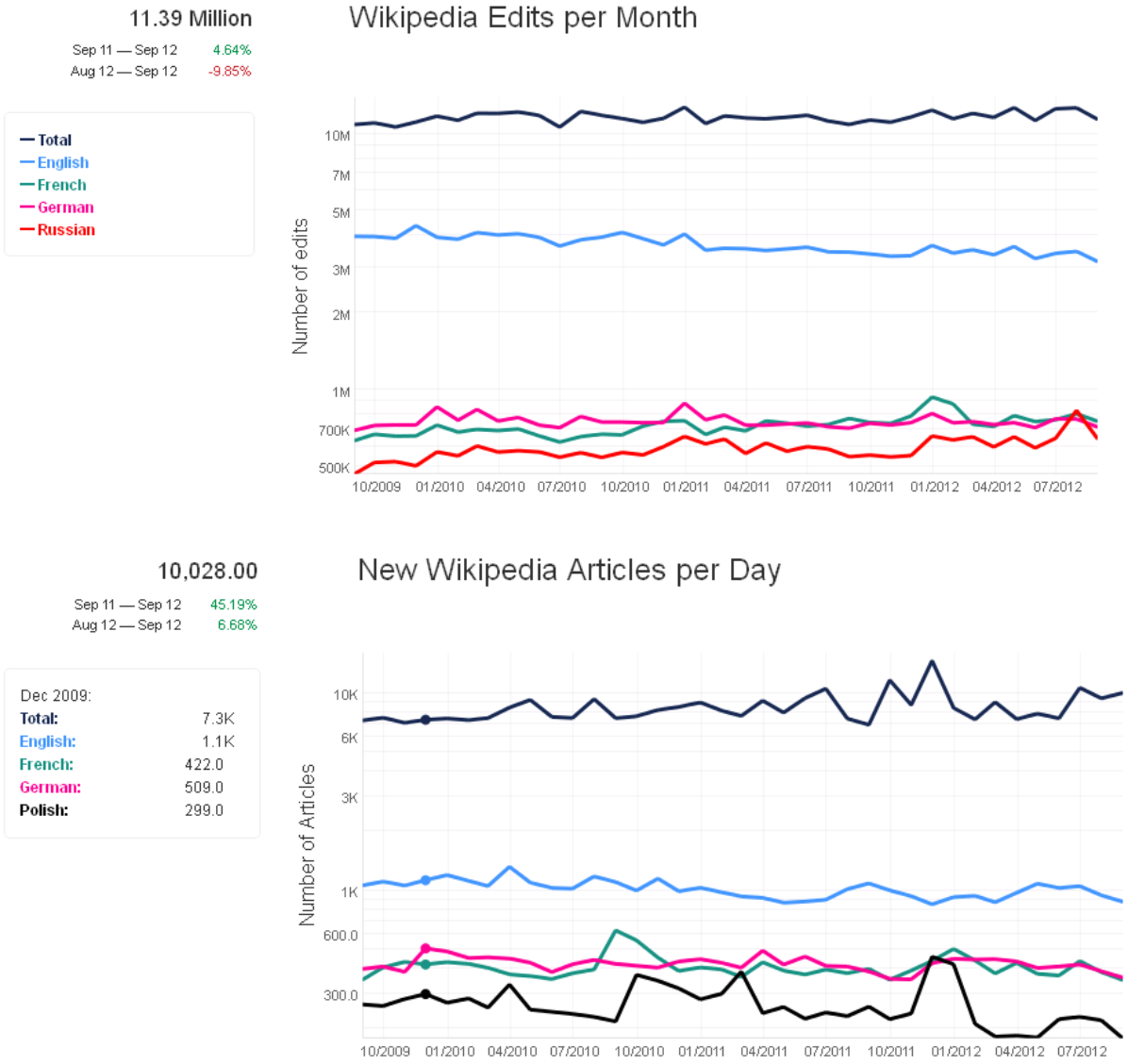
Linked Data: Volume

Geoindexed Linked Data – courtesy of Krzysztof Janowicz

http://stko.geog.ucsb.edu/location_linked_data



Data Velocity



- **Weather sensors**
- **Tweets**
- **Satellite images**
- **…**

Linked Data: Variety

Copernicus lunar crater located on earth – courtesy of Krzysztof Janowicz http://stko.geog.ucsb.edu/location_linked_data (missing reference coordinate system)

Information Workbench

Print Help Login

Copernicus (lunar crater)

You do not have permission to edit this page.

View Revisions

Copernicus is a lunar impact crater named after the astronomer Nicolaus Copernicus, located in eastern Oceanus Procellarum. It is estimated to be about 800 million years old, and typifies craters that formed during the Copernican period in that it has a prominent ray system.

Contents

Characteristics
Names
Satellite craters
See also
References
External links

Characteristics

Copernicus is visible using binoculars, and is located slightly northwest of the center of the Moon's Earth-facing hemisphere. South of the crater is the Mare Insularum, and to the south-south west is the crater Reinhold. North of Copernicus are the Montes Carpatus, which lie at the south edge of Mare Imbrium. West of Copernicus is a group of dispersed lunar hills. Due to its relative youth, the crater has remained in a relatively pristine shape since it formed.

The circular rim has a discernible hexagonal form, with a terraced inner wall and a 30 km wide, sloping rampart that descends nearly a kilometer to the surrounding mare. There are three distinct terraces visible, and arc-shaped landslides due to slumping of the inner wall as the crater debris subsided.

Most likely due to its recent formation, the crater floor has not been flooded

Location of Copernicus.

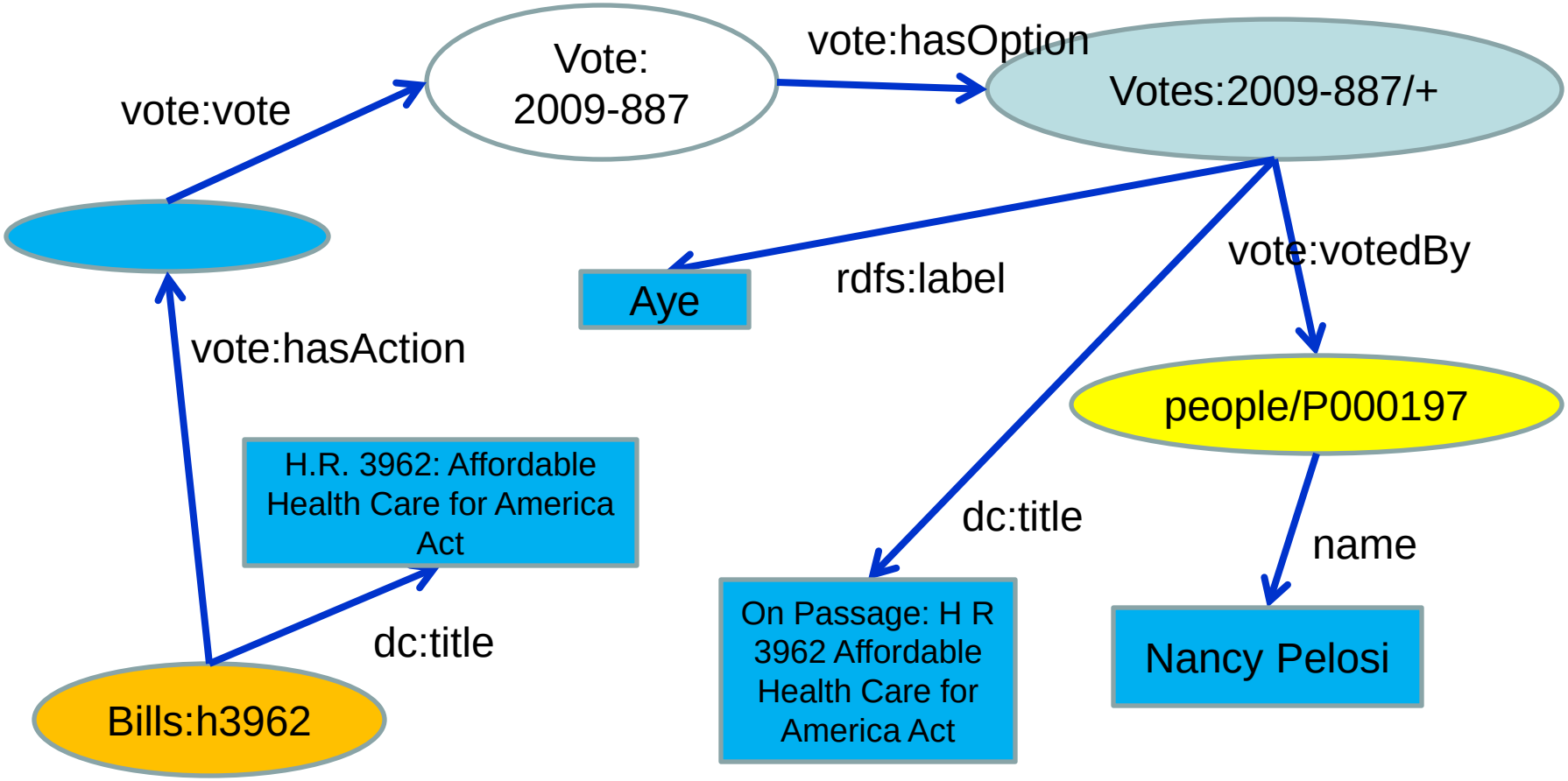
Location of Copernicus.

Image

Google Map

Linked Data: Variety and Value (GovTrack)

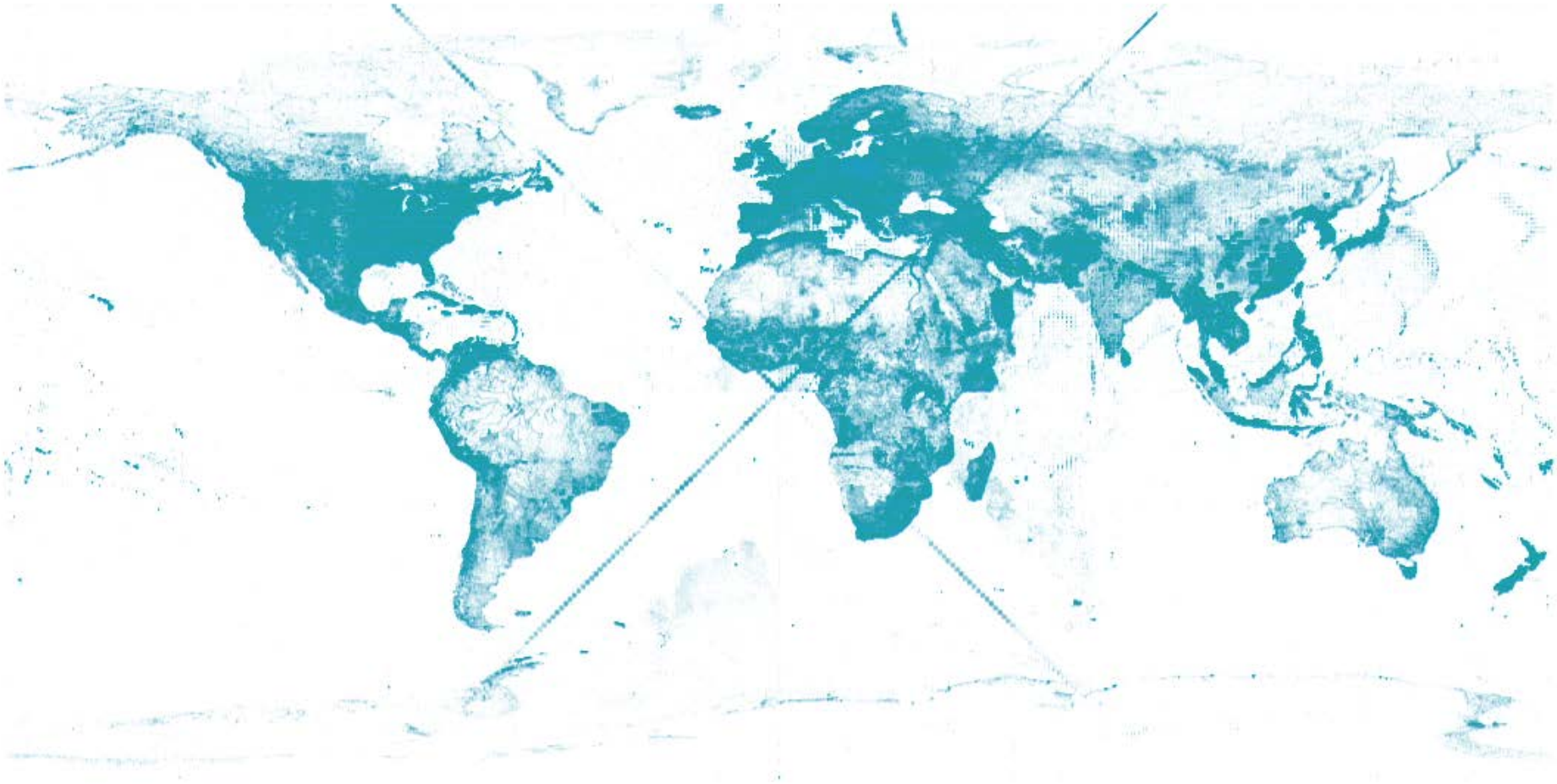
“Nancy Pelosi voted in favor of the Health Care Bill.”



Linked Data: Variety and Veracity

Geoindexed Linked Data – courtesy of Krzysztof Janowicz

http://stko.geog.ucsb.edu/location_linked_data



Linked Data: Variety and Veracity

Courtesy of Krzysztof Janowicz

http://stko.geog.ucsb.edu/location_linked_data

FactForge

RDF Search and Explore | SPARQL | RelFinder | About | Contact |

SPARQL Query

Results for your query (6) - Edit query

View as Exhibit Download SPARQL Results in: JSON | XML

place	populationCount
dbpedia:Keta	18077
http://sws.geonames.org/2304548/	29748
w-flick:Aneho	47579
http://sws.geonames.org/6295630/	6814400000
dbpedia:Lomé	749700
http://sws.geonames.org/2393947/	9847

Linked Data: Veriety and Veracity

Courtesy of Krzysztof Janowicz

http://stko.geog.ucsb.edu/location_linked_data

FactForge

RDF Search and Explore | SPARQL | RelFinder | About | Contact

Welt RDF Rank

Source: http://sws.geonames.org/6295630/

RDF Search and Explore

Subject (100 of 8935127) | Predicate | Object | All

Download in: JSON | RDF | N3/Turtle | N-Triples

Statements in which the resource exists as a subject. Named Gr

Predicate	Object
rdf:type	http://schema.org/Place, geo-ont:Feature,
rdfs:seeAlso	dbpedia:Earth, http://sws.geonames.org/6
rdfs:isDefinedBy	http://sws.geonames.org/6295630/about.r
rdfs:label	Earth@en, Globe@en, World@en
skos:altLabel	Earth@en, Globe@en, World@en
dc:type	geo-ont:L, geo-ont:L.AREA
dc-term:type	geo-ont:L, geo-ont:L.AREA
geo-pos:long	0
geo-pos:lat	0

Coordinates: 1°0′N 4°0′E

Map of the Gulf of Guinea, showing the chain of islands formed by the Cameroon line of volcanoes.

Big Data: My take

Variety can be handled with existing methods if volume and velocity are small.

Big Data research is primarily about methods for handling variety if volume and velocity are so high that existing methods fail.

Linked Data and Big Data

- **Linked Data is a kind of structured Big Data**
- **Linked Data is Big Data in a nutshell**

Many of the same problems

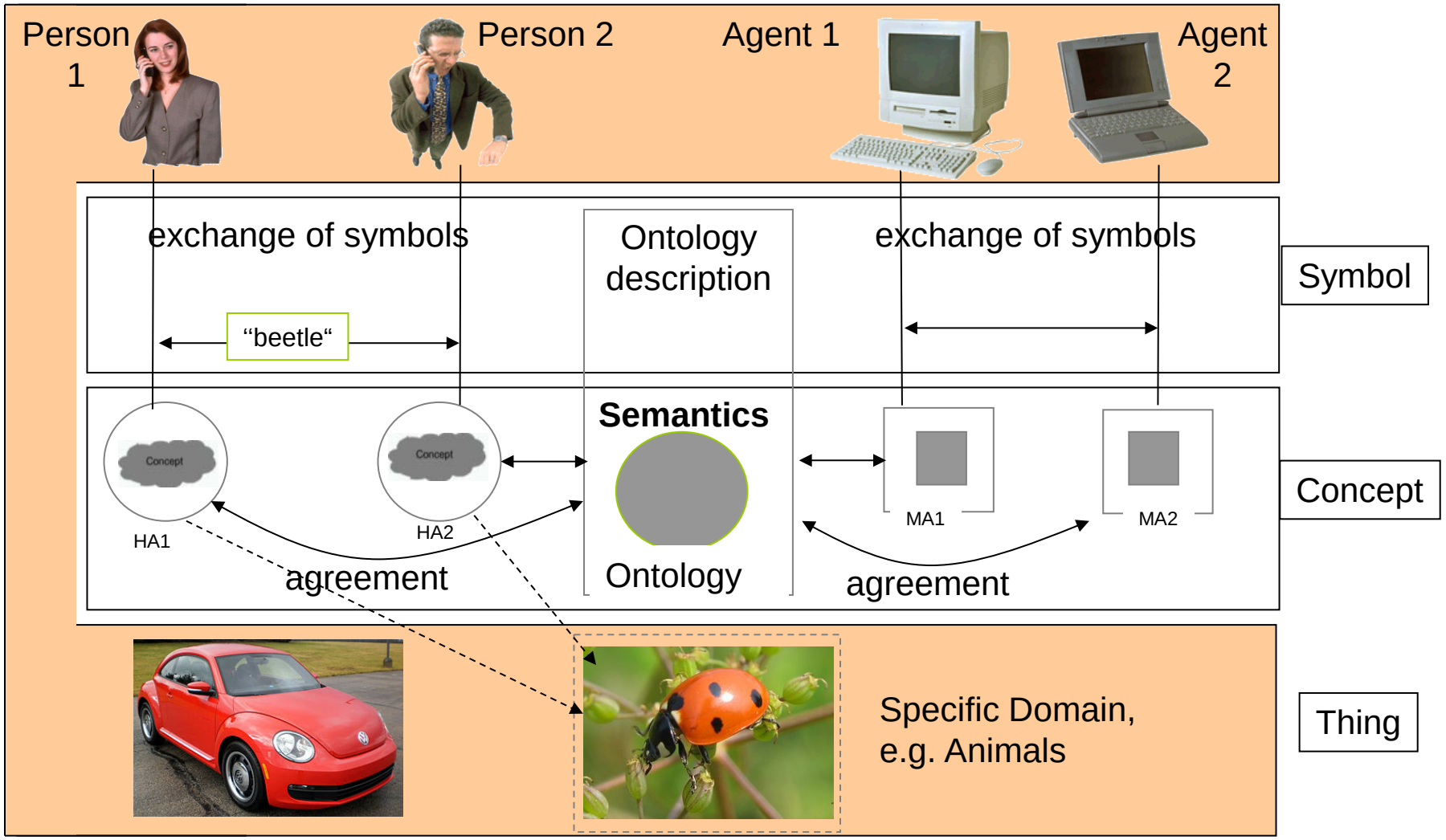
Testbed for Big Data solutions

Intermediate stage for getting *semantics* into Big Data

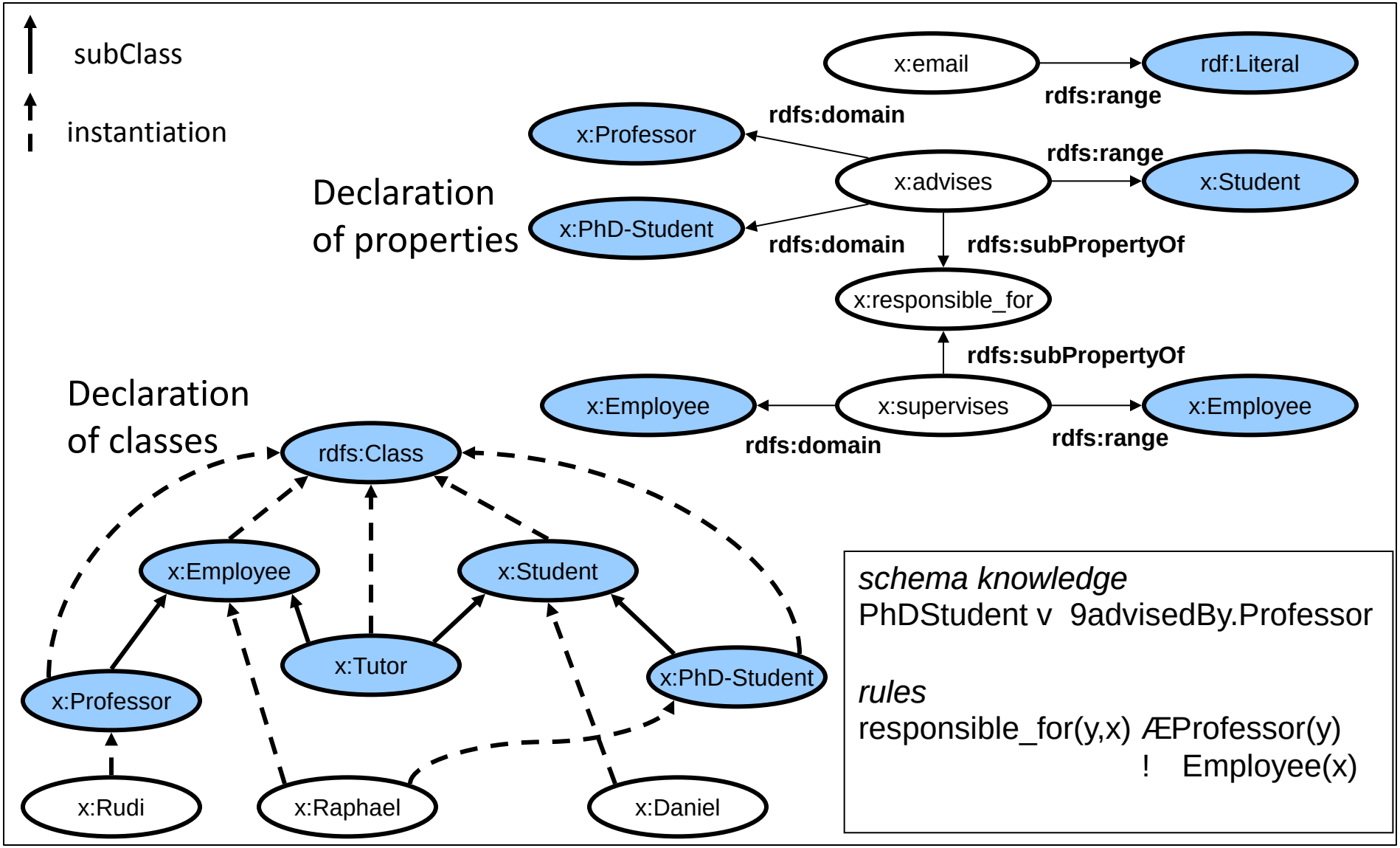
Contents

- **Big Data**
- **Semantic Web**
- **EarthCube**

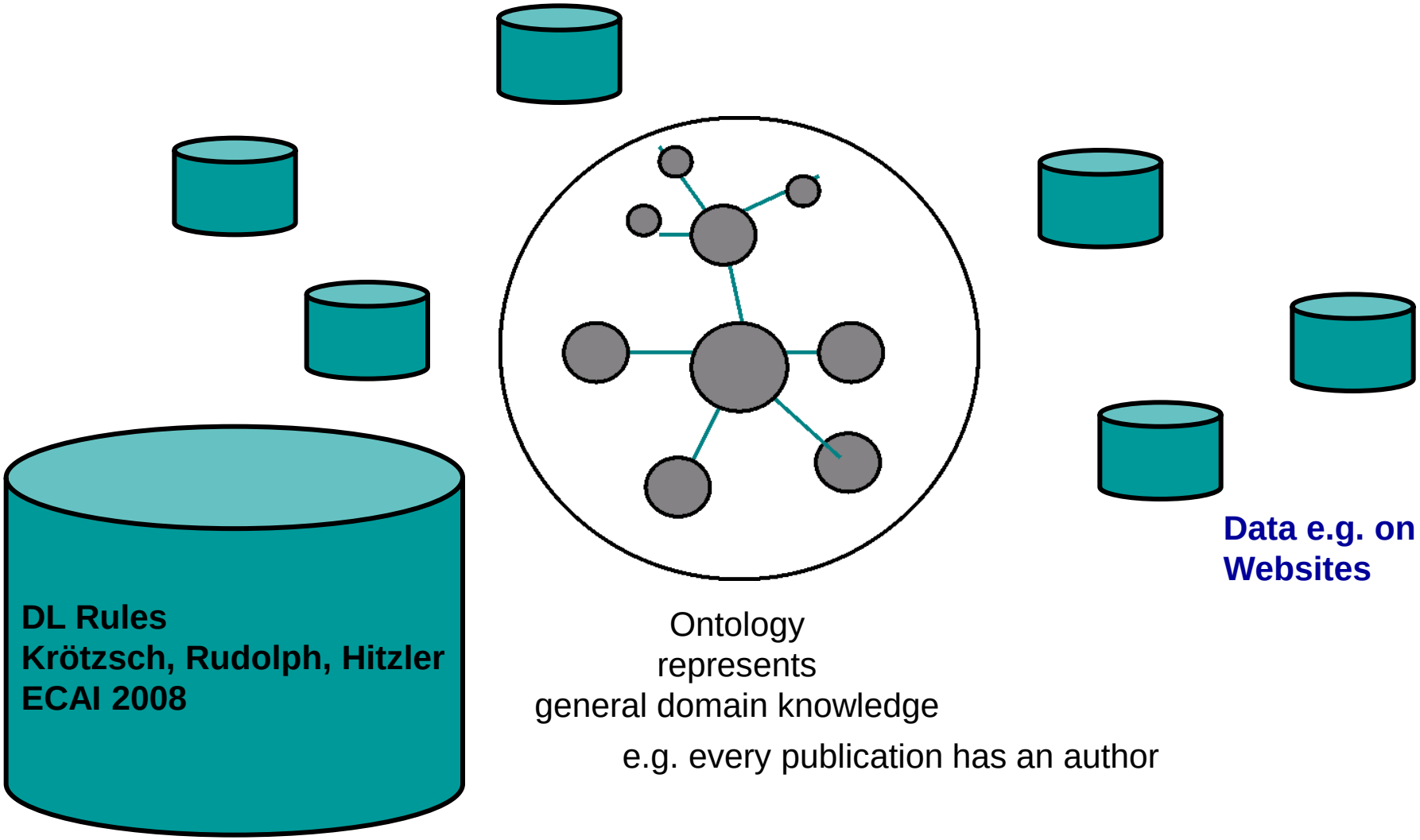
Basic Idea of the Semantic Web



Ontology Example

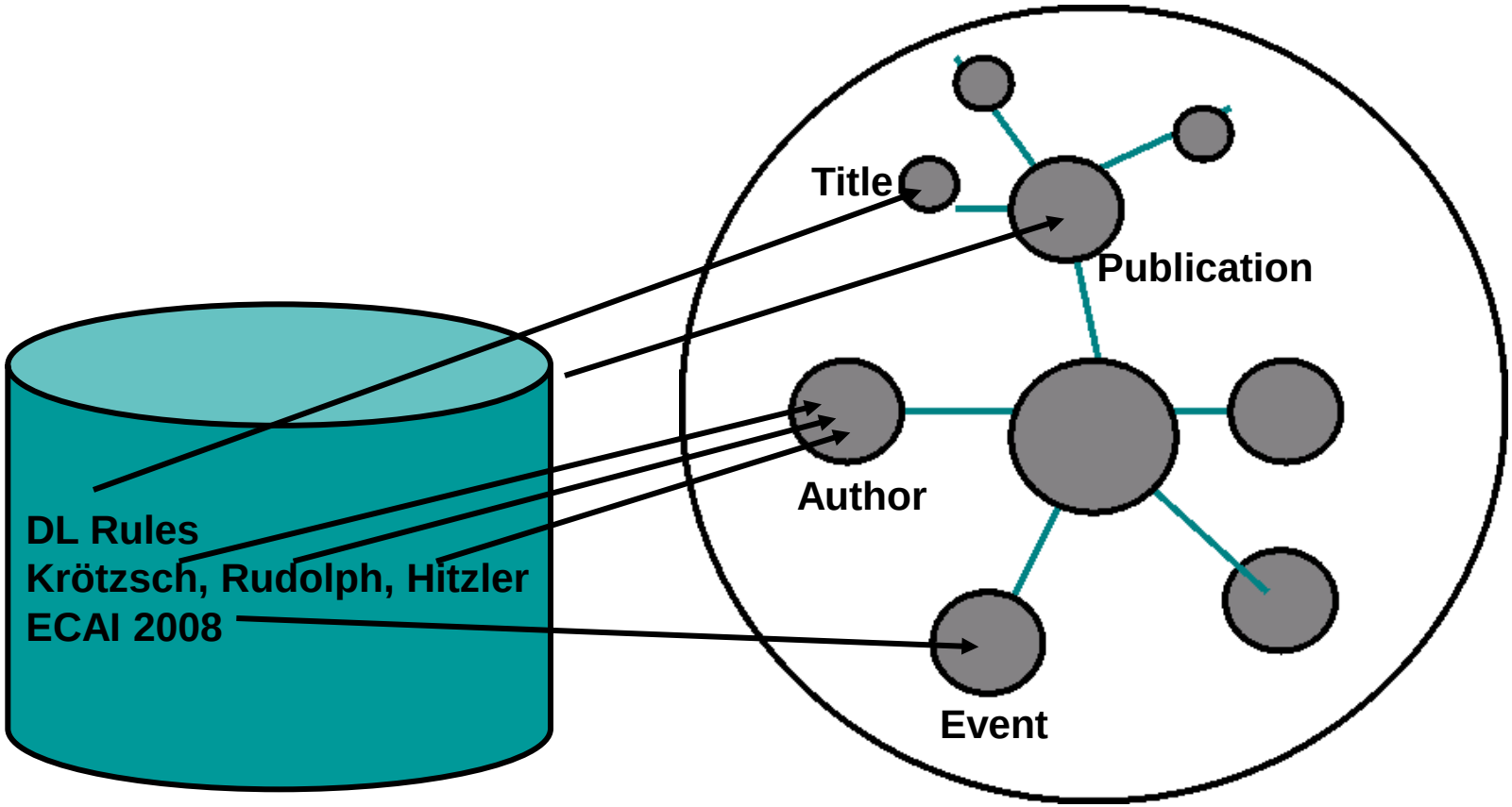


Basic Idea of the Semantic Web



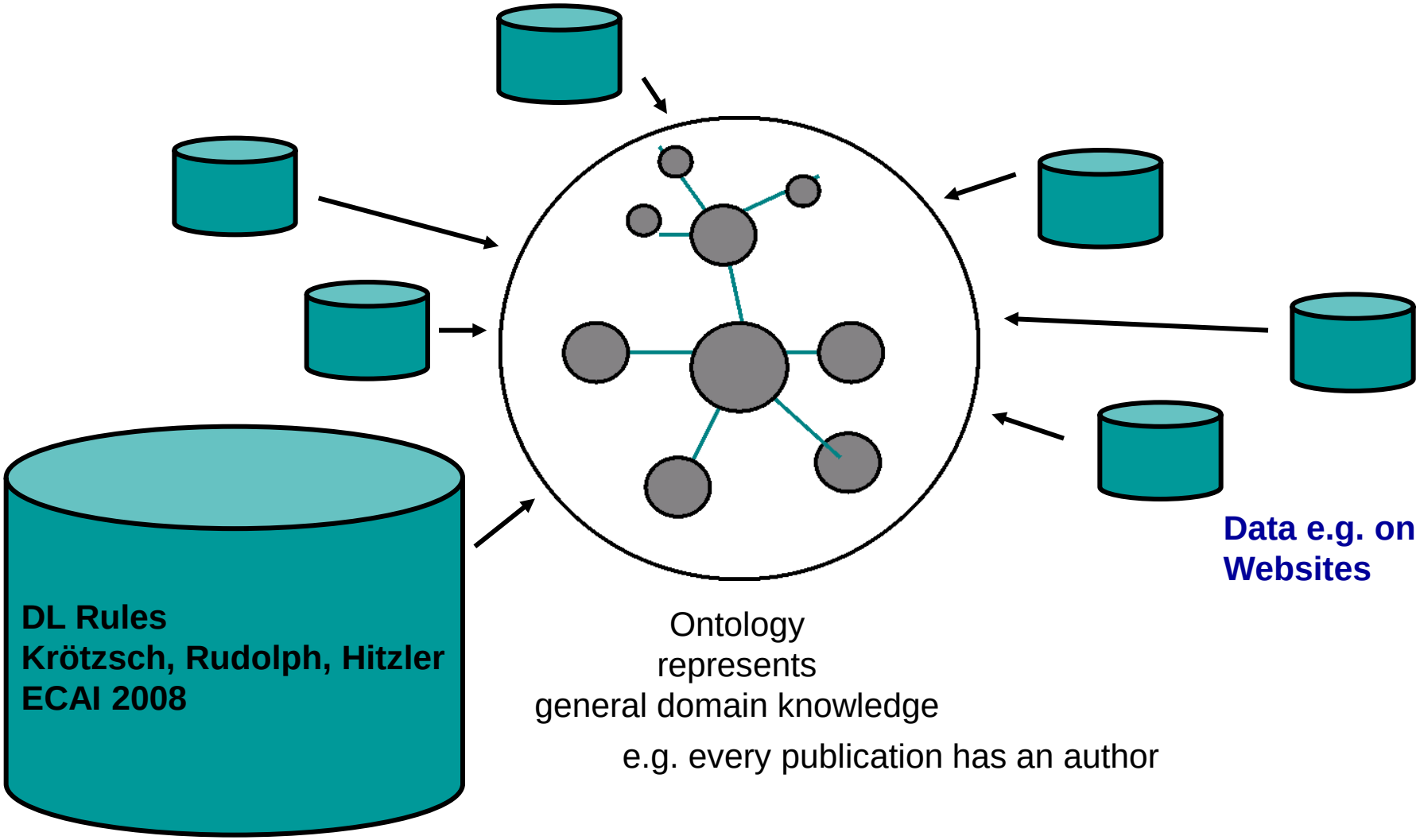
e.g. every publication has an author

Basic Idea of the Semantic Web



e.g. every publication has an author

Basic Idea of the Semantic Web



IBM

IBM WATSON

The Science Behind an Answer

Watson performs so fast that it can rival the greatest human contestants in understanding a Jeopardy! clue and arriving at a single, precise answer. The significance of this accomplishment can be difficult to comprehend.

Watch the video to see how the computing system designed to play Jeopardy! works.

Possible Answers

bake
balance
ban
bang
bat
bathe
battle
be
beam
bear

The first person mentioned by name
in 'The Man in th
is this hero of a
by the same aut

IBM Watson Solutions and WellPoint, America's largest health benefits company with

The DeepQA hypothesis is that by complementing classic knowledge-based approaches with recent advances in NLP, Information Retrieval, and Machine Learning to interpret and reason over huge volumes of widely accessible naturally encoded knowledge (or "unstructured knowledge") we can build effective and adaptable open-domain QA systems. While they may not be able to formally prove an answer is correct in purely logical terms, they can build confidence based on a combination of reasoning methods that operate directly on a combination of the raw natural language, automatically extracted entities, relations and available structured and semi-structured knowledge available from for example the Semantic Web.

What is Watson?

Implications for analytics, system design and industry transformation ›

Watson for a Smarter Planet™

Join the conversation on IBM Watson Connect

Watson - A System Designed for Answers

Optimize Your Growing

Store | Mac | iPod | iPhone | iPad | iTunes | Support

iPhone

Features | Built-in Apps | From the App Store | iOS | iCloud | Tech Specs | Buy iPhone

Siri. Beta

Your wish is its command.

Siri on iPhone 4S lets you use yo
voice to send messages, schedu
meetings, place phone calls, an
more. Ask Siri to do things just
talking the way you talk. Siri

mediabistro | semanticweb.com | SemTechBiz SF | more >>

semanticweb.com™ The Voice of Semantic Web Technology and Linked Data Business

Home | Events | Community | Learning | Industry Verticals | Answers | Search semanticweb.com | SEARCH

Apple Buys Siri: Once Again The Back Story Is About Semantic Web

By Bernard Lunn on April 28, 2010 2:33 PM

According to Robert Scoble who got it from tracking FTC, Apple is buying Siri. (This has yet to be confirmed by Siri or Apple). The front story is mobile, specifically a bruising battle between Apple and Google. But once again the back story is semantic technology. Siri is not some cute iPhone app banged together in a garage over a Red Bull fueled long weekend. Siri has hard core semantic tech that originated from Darpa (just like that little system called the Internet).

Like the Facebook OpenGraph story, this is another example of semantic web going mainstream. The Open Graph front story was all about social media, but the back story was their adoption of RDFa. That has been a big boost to the semantic web community.

Siri looks like a good exit for investors and will give them confidence to invest more in companies

Semanticweb.com Newsletter
Enter your email for updates | and your ZIP | SIGN UP

Semanticweb.com Event Updates
Enter your email for updates | and your ZIP | SIGN UP

Send an anonymous tip | SEND

SemanticWeb.com on Facebook
Like 1,467

Twitter
Mobile
RSS

SEMANTIC TECH & BUSINESS CONFERENCE
June 3-7, 2012
San Francisco
presented by SemanticWeb.com and WebMediaBrands

WRIGHT STATE UNIVERSITY

- **Thing**: additionalType, description, image, name, url
  - **CreativeWork**: about, accountablePerson, aggregateRating, alternativeHeadline, associatedMedia, audience, audio, author, award, awards, comment, contentLocation, contentRating, contributor, copyrightHolder, copyrightYear, creator, dateCreated, dateModified, datePublished, discussionUrl, editor, encoding, encodings, genre, headline, inLanguage, interactionCount, isFamilyFriendly, keywords, mentions, offers, provider, publisher, publishingPrinciples, review, reviews, sourceOrganization, text, thumbnailUrl, version, video
    - **Article**: articleBody, articleSection, wordCount
      - **BlogPosting**
      - **NewsArticle**: dateline, printColumn, printEdition, printPage, printSection
      - **ScholarlyArticle**
        - **MedicalScholarlyArticle**: citation, publicationType
    - **Blog**: blogPost, blogPosts
    - **Book**: bookEdition, bookFormat, illustrator, isbn, numberOfPages
    - **Comment**
    - **Diet**: dietFeatures, endorsers, expertConsiderations, overview, physiologicalBenefits, proprietaryName, risks
    - **ExercisePlan**: activityDuration, activityFrequency, additionalVariable, exerciseType, intensity, repetitions, restPeriods, workload
    - **ItemList**: itemListElement, itemListOrder
    - **Map**
    - **MediaObject**: associatedArticle, bitrate, contentSize, contentUrl, duration, embedUrl, encodesCreativeWork, encodingFormat, expires, height, interactionCount, offers, playerType, regionsAllowed, requiresSubscription, uploadDate, width
      - **AudioObject**: transcript
      - **ImageObject**: caption, exifData, representativeOfPage, thumbnail
      - **MusicVideoObject**
      - **VideoObject**: caption, productionCompany, thumbnail, transcript, videoFrameSize, videoQuality
    - **Movie**: actor, actors, director, duration, musicBy, producer, productionCompany, trailer
    - **MusicPlaylist**: numTracks, track, tracks
      - **MusicAlbum**: byArtist
    - **MusicRecording**: byArtist, duration, inAlbum, inPlaylist
    - **Painting**
    - **Photograph**
    - **Recipe**: cookingMethod, cookTime, ingredients, nutrition,
    - **Review**: itemReviewed, reviewBody, reviewRating
    - **Sculpture**
    - **SoftwareApplication**: applicationCategory, applicationSub
      downloadUrl, featureList, fileFormat, fileSize, installUr
      releaseNotes, requirements, screenshot, softwareVersion, storageRequirements
      - **MobileApplication**: carrierRequirements
      - **WebApplication**: browserRequirements
    - **TVEpisode**: actor, actors, director, episodeNumber, musicBy, partOfSeason, partOfTVSeries, producer, productionCompany, trailer
    - **TVSeason**: endDate, episode, episodes, numberOfEpisodes, partOfTVSeries, seasonNumber, startDate, trailer
    - **TVSeries**: actor, actors, director, endDate, episode, episodes, musicBy, numberOfEpisodes, producer, productionCompany, season, seasons, startDate, trailer

schema.org for enhancing web search
joint effort including Bing, Google, Yahoo, Yandex

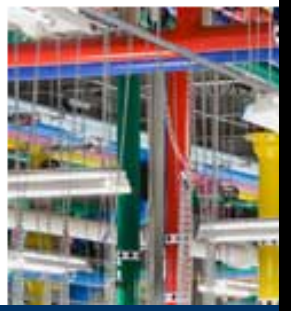
Google and the future of search: Amit Singhal and the Knowledge Graph

Google has revolutionised the way we holiday, shop, work and play. Now, with Knowledge Graph, it plans to radically transform the way we search the internet... again. But some voice qualms about the company's ambitions

Share 1253
Tweet 697
+1 371
Share 215
Email

Tim Adams
The Obs

Jump



senses, that attribute is in the process of changing. This year, Google will roll out what it calls its Knowledge Graph, the closest any system has yet come to creating what Tim Berners-Lee, originator of the web itself, called "the semantic web", the version that had understanding as well as data, that could itself provide answers, not links to answers.

The Knowledge Graph is a database of the 500 million most searched for people, places and

Online M.Ed. Mathematics

Become a College Instructor, Consultant, Curriculum Director or Dept Chair.

Wright State University

Sign i

Web Images Maps Shopping More Search tools

About 32,400,000 results (0.23 seconds)

Wright State University

www.wright.edu/

Expanded scholarship program offered. **Wright State University** has unveiled a new scholarship program designed to make a college education more affordable.

Score: **21** / 30 - 11 Google reviews - Write a review

3640 Colonel Glenn Hwy Dayton, OH 45435
(937) 775-3333

Wings
Get a WINGS username and password. Alternative Login for ...

Graduate School
Programs - Admissions - Apply - Check Application Status - ...

Undergraduate Admissions
... your new home away from home. Visit Undergraduate ...

The Official Wright State ...
Playing in the new and updated facilities, the Raiders have won ...

Academics
We think Wright State is pretty amazing—and national ...

Libraries
Research Databases - Hours - Contact Us - Staff - ...

Search wright.edu

Wright State University - Wikipedia, the free encyclopedia

en.wikipedia.org/wiki/Wright_State_University

Wright State University is a public research university in Fairborn, Ohio just outside of Dayton. The school offers degrees at the associate, bachelor's, master's, ...

Wright State University

Directions

Wright State University is a public research university in Fairborn, Ohio just outside of Dayton. The school offers degrees at the associate, bachelor's, master's, and doctoral level. Wikipedia

Nickname: Raiders

Address: 3640 Colonel Glenn Hwy, Dayton, OH 45435

Phone: (937) 775-3333

Mascot: Rowdy Raider

Founded: 1967

Colors: Gold, Green

Getting Started
Core Concepts
Advanced Topics
Technical Guides
API Reference
SDK Reference

Getting Started

Key Concepts
Tutorial
Tutorial Videos
Login Dialog
Social Plugins
Open Graph Samples
Open Graph Distribution

Open Graph Tool

Define Objects
Define Actions
Define Aggregations
Using Text Templates

Approval

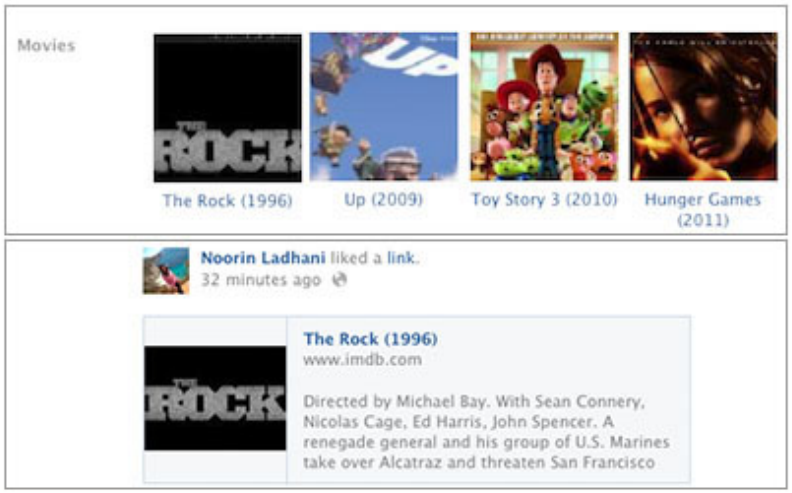
Open Graph Checklist
Submit for Approval
Open Graph Guidelines

Open Graph Protocol

We announced the next version of the Open Graph at f8 2011. It is currently released to the public. You can view the Open Graph documentation here. The below documentation refers to the first version of the Open Graph that shipped with the Like Button at f8 2010.

We have also updated how the Like Button will function with respect to the next version of Open Graph. Please read the following developer doc about the Like Button Migration.

The Open Graph Protocol enables you to integrate your Web pages into the social graph. It is currently designed for Web pages representing profiles of real-world things — things like movies, sports teams, celebrities, and restaurants. Including Open Graph tags on your Web page, makes your page equivalent to a Facebook Page. This means when a user clicks a Like button on your page, a connection is made between your page and the user. Your page will appear in the "Likes and Interests" section of the user's profile, and you have the ability to publish updates to the user. Your page will show up in the same places that Facebook pages show up around the site (e.g. search), and you can target ads to people who like your content. The structured data you provide via the Open Graph Protocol defines how your page will be represented on Facebook.



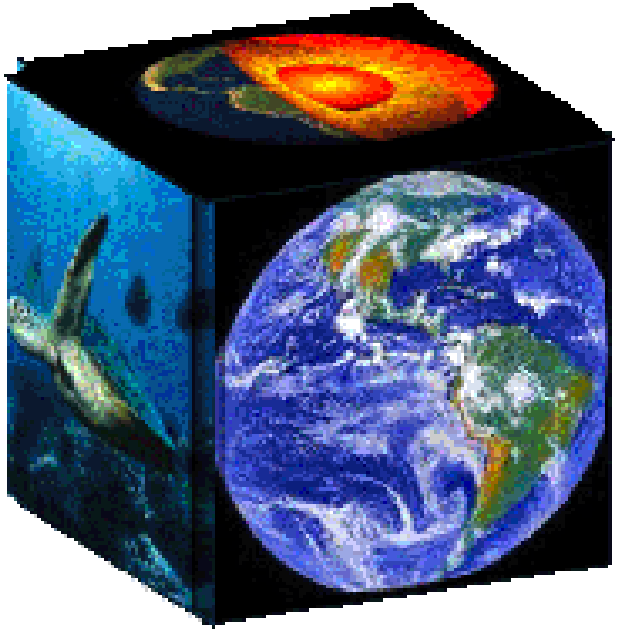
Contents

- **Big Data**
- **Semantic Web**
- **EarthCube**

EarthCube

NSF effort for the earth sciences

Goal:
To transform the conduct of research in the geosciences by developing IT solutions for the integration of information and data in the geosciences.



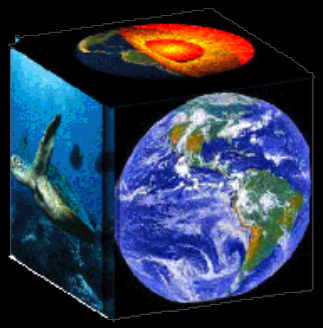
How this is going to be done is still in the making.

Semantic Technologies have been part of the mix from the start.

[Berg-Cross, …, Hitzler et al., GIBDa 2012]

EarthCube requires

- information integration
- interoperability
- conceptual modeling
- intelligent search
- data-model intercomparison
- data publishing support

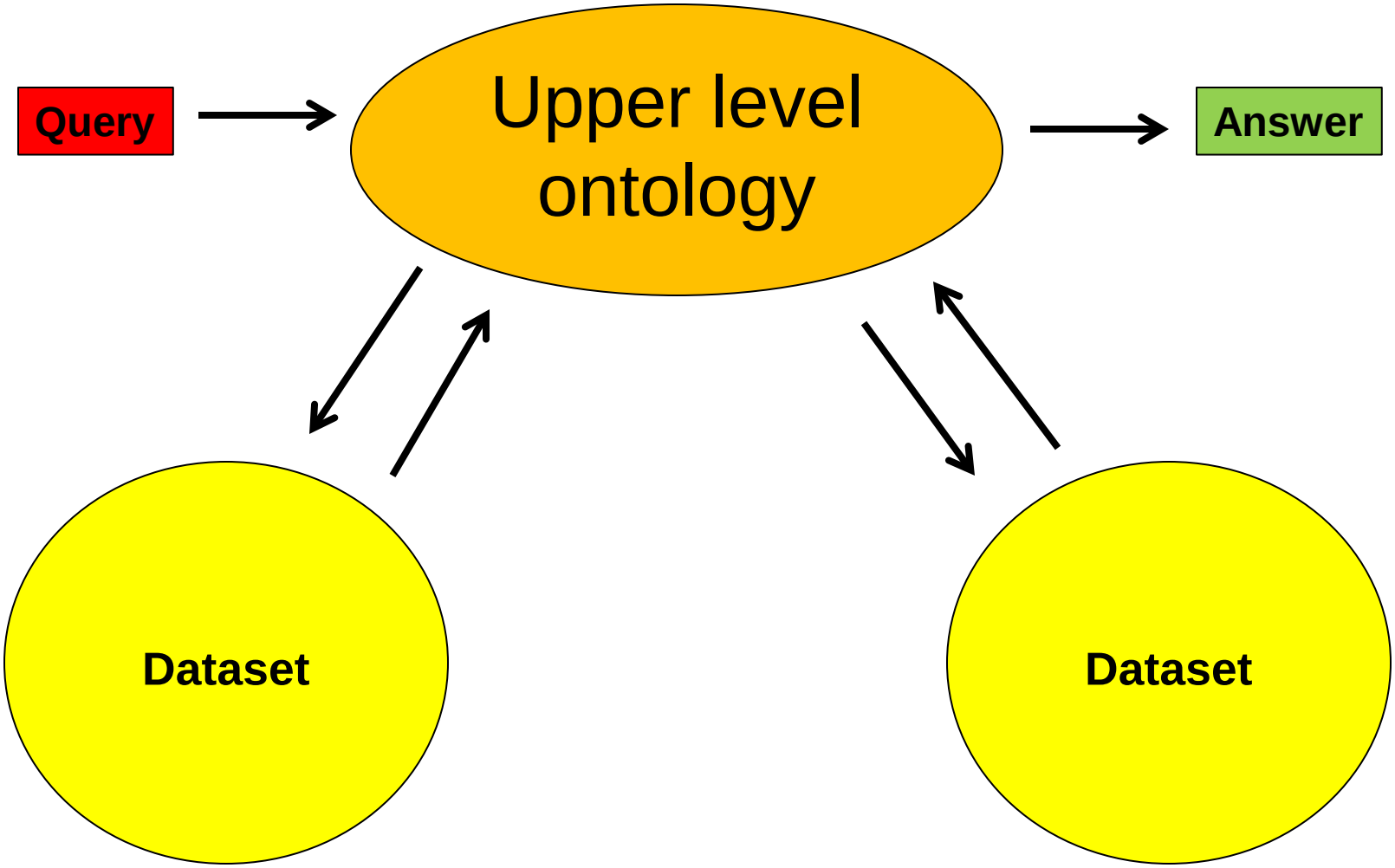


Semantic Web studies

- information integration
- interoperability
- conceptual modeling
- intelligent search
- data-model intercomparison
- data publishing support

Pascal Hitzler, WSU; Krzysztof Janowicz, UCSB

Vertical data integration



[Joshi, Jain, Hitzler et al. ODBASE 2012]

Ontological commitments

$$\text{a:flowsInto} \sqsubseteq \text{a:IsConnected} \quad (1)$$
$$\text{a:IrrigationCanal} \sqsubseteq \text{a:Canal} \quad (2)$$
$$\exists \text{a:flowsInto.a:AgriculturalField} \sqsubseteq \text{a:IrrigationCanal} \quad (3)$$
$$\text{a:Waterbody} \sqcap \text{a:Land} \sqsubseteq \bot \quad (4)$$
$$\text{a:AgriculturalField} \sqsubseteq \text{a:Land} \quad (5)$$

$$\text{b:flowsInto} \sqsubseteq \text{b:IsConnected} \quad (6)$$
$$\text{b:Canal} \sqsubseteq (\geq 2\ \text{b:IsConnected.b:Waterbody}) \quad (7)$$
$$\text{b:IrrigationCanal} \equiv (=1\ \text{b:isConnected.b:Waterbody}) \sqcap (=1\ \text{b:flowsInto.b:AgriculturalField}) \quad (8)$$

Two ontologies.

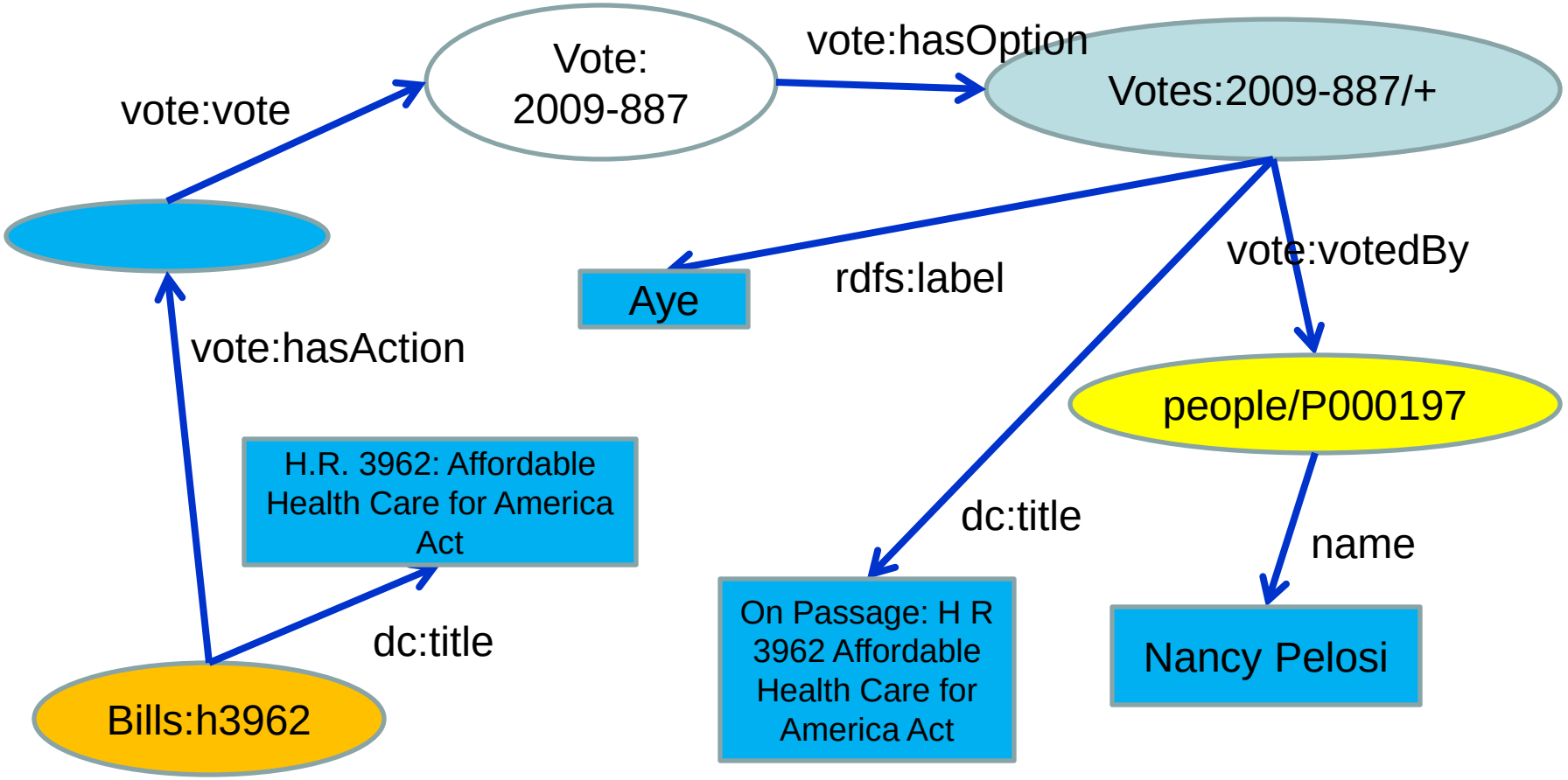
Left: transportation domain

Right: agriculture domain

We cannot simply equate a:Canal and b:Canal !

Linked Data: Variety and Value (GovTrack)

“Nancy Pelosi voted in favor of the Health Care Bill.”

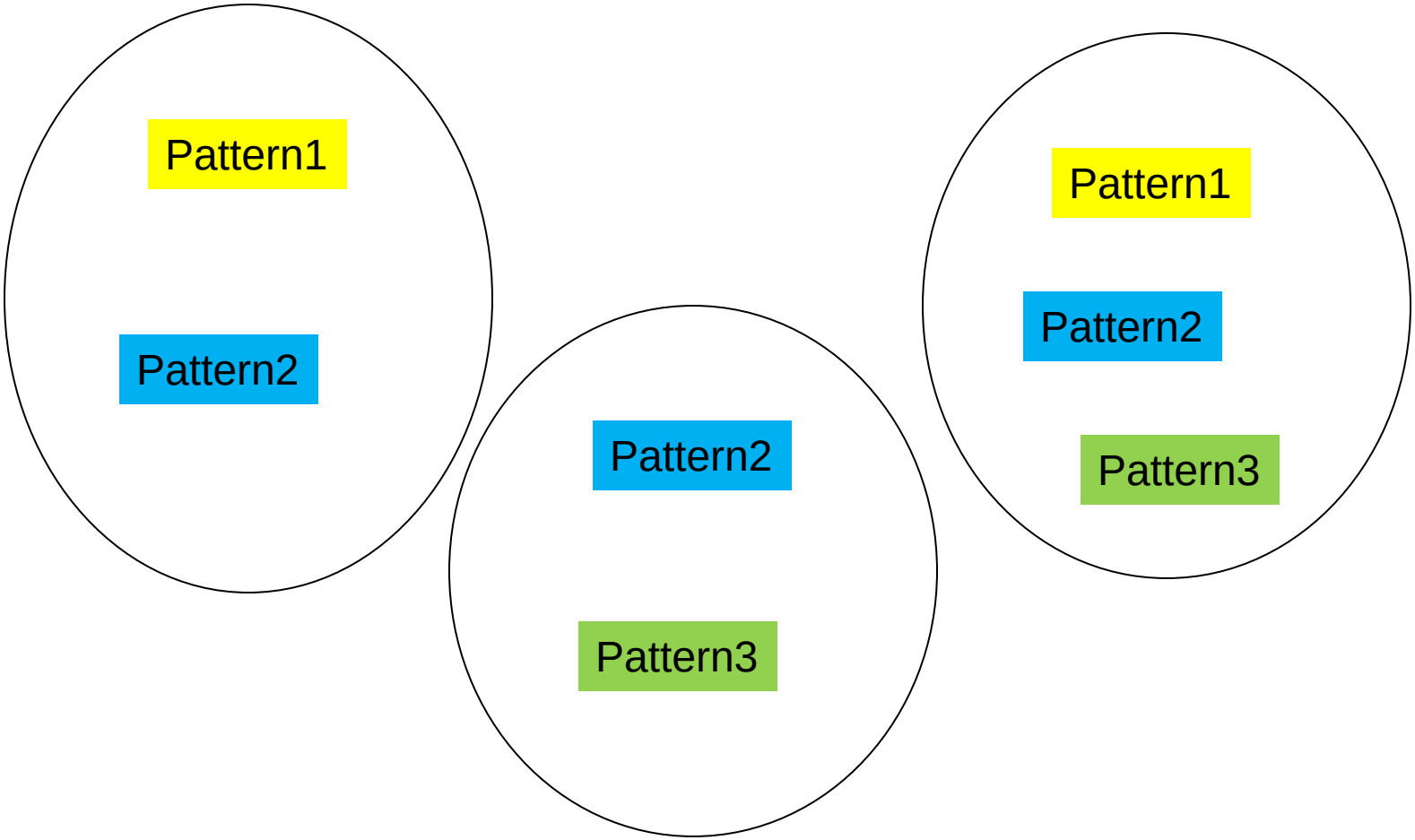


Ontology Design Patterns

- **Bottom-up homogeneization of data representation.**
- **Avoidance of strong ontological commitments.**
- **Avoidance of standardization.**
- **Well thought-out patterns can be very strong and versatile, thus serve many needs.**

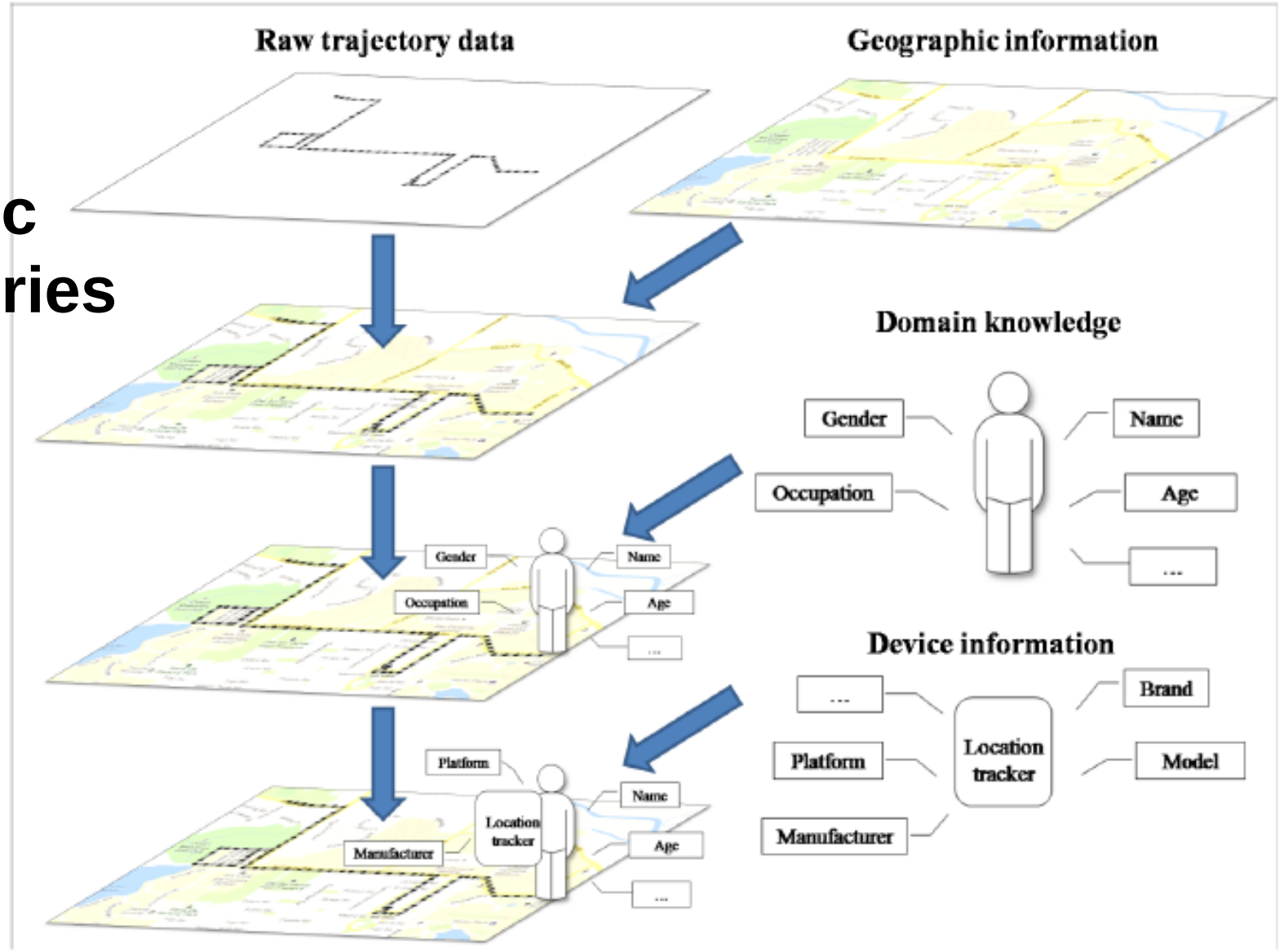
We are currently establishing many geo-patterns in a series of hands-on workshops, the GeoVoCamps, see http://vocamp.org/

Ontology Design Patterns



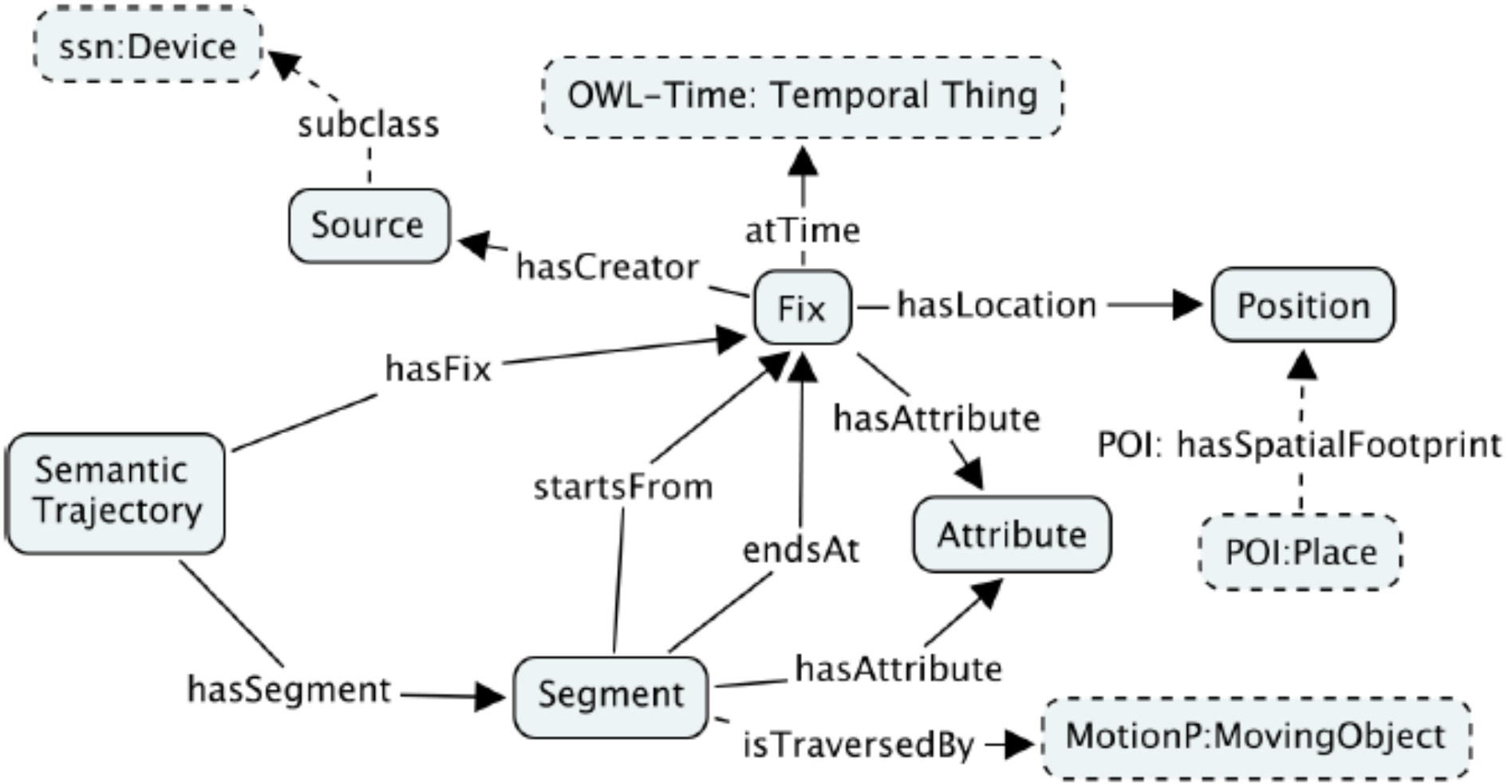
“Horizontal” alignment via patterns

Semantic Trajectories



[Hu, Janowicz, Carral, Scheider, Kuhn, Berg-Cross, Hitzler, Dean, COSIT2013, to appear]

Semantic Trajectories



Semantics in OWL

$$Fix \sqsubseteq \exists atTime.OWL\text{-}Time{:}Temporal\ Thing \sqcap \exists hasLocation.Position \sqcap \exists hasFix^{-}.SemanticTrajectory \quad (1)$$

$$Segment \sqsubseteq \exists startsFrom.Fix \sqcap \exists endsAt.Fix \quad (2)$$

$$\top \sqsubseteq \leq 1 startsFrom.\top \quad (3)$$

$$\top \sqsubseteq \leq 1 endsAt.\top \quad (4)$$

$$Segment \sqsubseteq \exists hasSegment^{-}.SemanticTrajectory \quad (5)$$

$$startsFrom^{-} \circ endsAt \sqsubseteq hasNext \quad (6)$$

$$hasNext \sqsubseteq hasSuccessor \quad (7)$$

$$hasSuccessor \circ hasSuccessor \sqsubseteq hasSuccessor \quad (8)$$

$$hasNext^{-} \sqsubseteq hasPrevious \quad (9)$$

$$hasSuccessor^{-} \sqsubseteq hasPredecesor \quad (10)$$

Semantics in OWL

$$Fix \sqcap \neg \exists endsAt.Segment \sqsubseteq StartingFix \quad (11)$$

$$Fix \sqcap \neg \exists startsFrom.Segment \sqsubseteq EndingFix \quad (12)$$

$$Segment \sqcap \exists startsFrom.StartingFix \sqsubseteq StartingSegment \quad (13)$$

$$Segment \sqcap \exists endsAt.EndingFix \sqsubseteq EndingSegment \quad (14)$$

$$SemanticTrajectory \sqsubseteq \exists hasSegment.Segment \quad (15)$$

$$hasSegment \circ startsFrom \sqsubseteq hasFix \quad (16)$$

$$hasSegment \circ endsAt \sqsubseteq hasFix \quad (17)$$

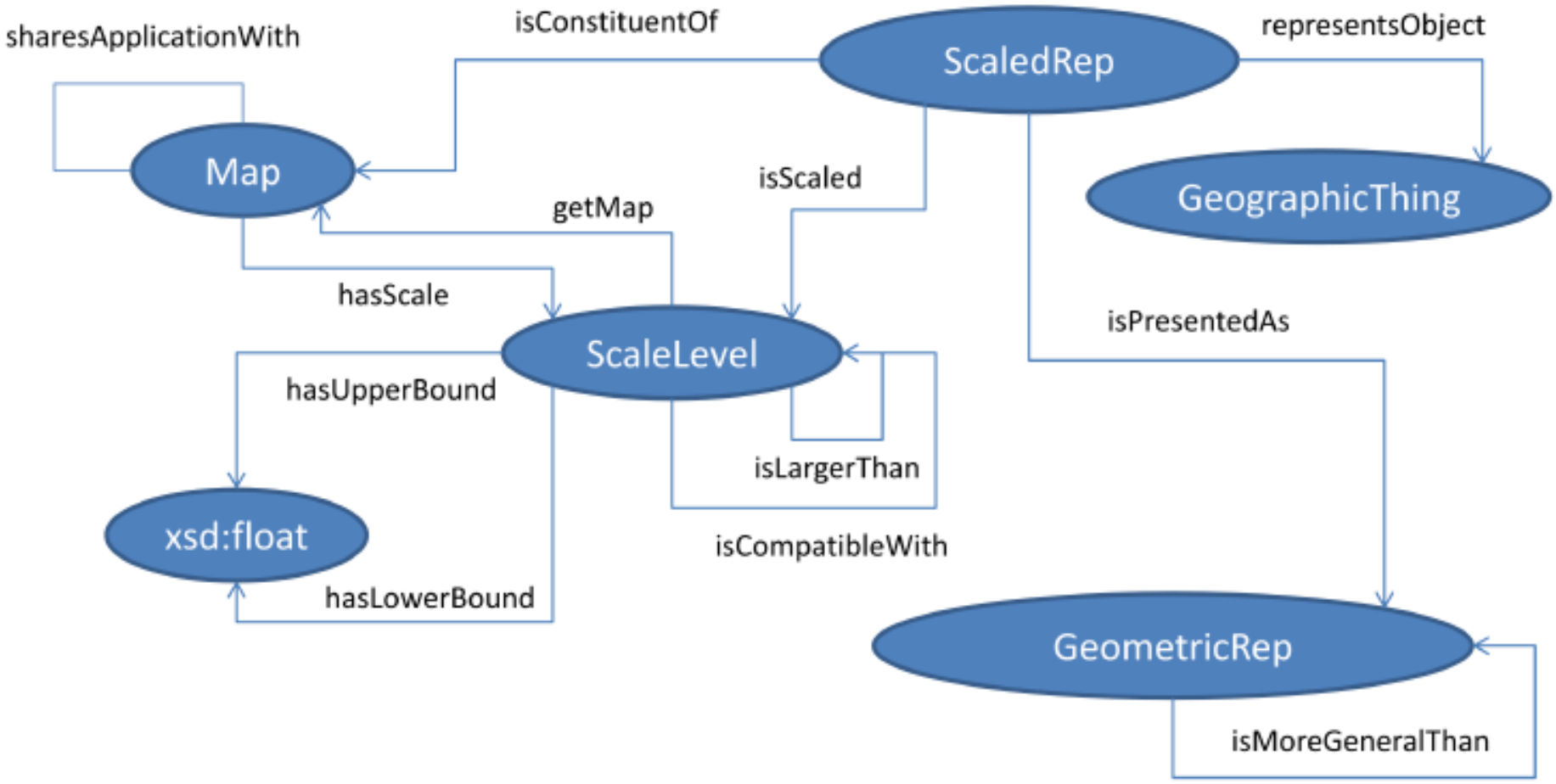
$$\exists hasSegment.Segment \sqsubseteq SemanticTrajectory \quad (18)$$

$$\exists hasSegment^{-}.SemanticTrajectory \sqsubseteq Segment \quad (19)$$

$$\exists hasFix.Segment \sqsubseteq SemanticTrajectory \quad (20)$$

$$\exists hasFix^{-}.SemanticTrajectory \sqsubseteq Fix \quad (21)$$

Cartographic Map Scaling



[Carral, Scheider, Janowicz, Vardeman, Krisnadhi, Hitzler, ESWC2013, to appear]

Semantics in OWL

$$sharesApplicationWith \circ sharesApplicationWith \sqsubseteq sharesApplicationWith \quad (1)$$

$$sharesApplicationWith^{-} \sqsubseteq sharesApplicationWith \quad (2)$$

$$Map \sqsubseteq \exists sharesApplicationWith.Self \quad (3)$$

$$\top \sqsubseteq \leq 1 hasScale.\top \quad (4)$$

$$hasScale^{-} \sqsubseteq getMap \quad (5)$$

$$\top \sqsubseteq \leq 1 (getMap \circ sharesApplicationWith).\top \quad (6)$$

$$isConstituentOf \circ hasScale \sqsubseteq isScaled \quad (7)$$

$$isLargerThan \circ isLargerThan \sqsubseteq isLargerThan \quad (8)$$

$$\exists (isLargerThan \sqcap isLargerThan^{-}).\top \sqsubseteq \bot \quad (9)$$

$$isMoreGeneralThan \circ isMoreGeneralThan \sqsubseteq isMoreGeneralThan \quad (10)$$

$$\exists (isMoreGeneralThan \sqcap isMoreGeneralThan^{-}).\top \sqsubseteq \bot \quad (11)$$

Semantics in OWL

$$isCompatibleWith^{-} \sqsubseteq isCompatibleWith \quad (12)$$

$$ScaleLevel \sqsubseteq \exists isCompatibleWith.Self \quad (13)$$

$$ScaleLevel \sqsubseteq \exists hasLowerBound.xsd{:}float \quad (14)$$

$$ScaleLevel \sqsubseteq \exists hasUpperBound.xsd{:}float \quad (15)$$

$$\top \sqsubseteq \leq 1 hasLowerBound.\top \quad (16)$$

$$\top \sqsubseteq \leq 1 hasUpperBound.\top \quad (17)$$

$$\top \sqsubseteq \leq 1 isPresentedAs.\top \quad (18)$$

$$\top \sqsubseteq \leq 1 isScaled.\top \quad (19)$$

$$\top \sqsubseteq \leq 1 representsObject.\top \quad (20)$$

$$ScaledRep \sqsubseteq \exists isPresentedAs.GeometricRep \quad (21)$$

$$ScaledRep \sqsubseteq \exists isScaled.ScaleLevel \quad (22)$$

$$ScaledRep \sqsubseteq \exists representsObject.GeographicThing \quad (23)$$

$$isConstituentOf^{-} \circ representsObject \circ representsObject^{-} \sqsubseteq R_{aux} \quad (24)$$

$$\top \sqsubseteq \leq 1 (R_{aux} \sqcap isConstituentOf^{-}).\top \quad (25)$$

Semantics in OWL

$$sharesApplicationWith(m_x, m_y) \wedge hasScale(s_y, m_y) \wedge hasScale(s_x, m_x)$$
$$\wedge isLargerThan(s_x, s_y) \wedge$$
$$isConstituentOf(m_x, sr_x) \wedge isConstituentOf(m_y, sr_y) \wedge$$
$$representsObject(sr_x, g) \wedge representsObject(sr_y, g) \wedge$$
$$isPresentedAs(sr_x, grr_x) \wedge isPresentedAs(sr_y, grr_y) \wedge$$
$$isMoreGeneralThan(grr_x, grr_y) \rightarrow \bot(m_x)$$

This rule enforces that the ontology becomes inconsistent if

- there exist maps m_1 and m_2 belonging to the same application with scales s_1 and s_2,
- scale s_1 is larger than scale s_2,
- maps m_1 and m_2 contain scaled representations sr_1 and sr_2 that represent the same geographic thing g, and
- the geographic representation record grr_1 for sr_1 is more general than the one for sr_2, namely grr_2.

Semantics in OWL

$$sharesApplicationWith(m_x, m_y) \wedge hasScale(s_y, m_y) \wedge hasScale(s_x, m_x)$$
$$\wedge isLargerThan(s_x, s_y) \wedge$$
$$isConstituentOf(m_x, sr_x) \wedge isConstituentOf(m_y, sr_y) \wedge$$
$$representsObject(sr_x, g) \wedge representsObject(sr_y, g) \wedge$$
$$isPresentedAs(sr_x, grr_x) \wedge isPresentedAs(sr_y, grr_y) \wedge$$
$$isMoreGeneralThan(grr_x, grr_y) \rightarrow \bot(m_x)$$

$$hasScale^- \circ sharesApplicationWith \circ hasScale \sqsubseteq R_1$$
$$R_1 \sqcap isLargerThan \sqsubseteq R_2$$
$$isScaled \circ R_2 \circ isScaled \sqsubseteq R_3$$
$$isPresentedAs \circ isMoreGeneralThan^- \circ isPresentedAs^- \sqsubseteq R_4$$
$$representsObject \circ representsObject^- \sqsubseteq R_5$$
$$R_3 \sqcap R_4 \sqcap R_5 \sqsubseteq R_\bot$$
$$\exists R_\bot . \top \sqsubseteq \bot$$

Take-home messages

- **Linked Data is Big Data in a nutshell.**
- **Semantic Web Technologies address the meaning of data.**
- **Ontology design patterns is a powerful concept for Web data sharing.**
- **GIScience needs semantics for data and information integration.**

Thanks!

References

- Pascal Hitzler, Frank van Harmelen, A reasonable Semantic Web. Semantic Web 1 (1-2), 39-44, 2010.
- Prateek Jain, Pascal Hitzler, Peter Z. Yeh, Kunal Verma, Amit P. Sheth, Linked Data is Merely More Data. In: Dan Brickley, Vinay K. Chaudhri, Harry Halpin, Deborah McGuinness: Linked Data Meets Artificial Intelligence. Technical Report SS-10-07, AAAI Press, Menlo Park, California, 2010, pp. 82-86. ISBN 978-1-57735-461-1. Proceedings of LinkedAI at the AAAI Spring Symposium, March 2010.
- Pascal Hitzler, Krzysztof Janowicz, *What's Wrong with Linked Data?* http://blog.semantic-web.at/2012/08/09/whats-wrong-with-linked-data/ , August 2012.
- Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph, *Foundations of Semantic Web Technologies*. Chapman and Hall/CRC Press, 2009.

References

- Pascal Hitzler, Krzysztof Janowicz, Linked Data, Big Data, and the 4th Paradigm. Semantic Web 4 (3), 2013, 233-235.
- Krzysztof Janowicz, Pascal Hitzler, The Digital Earth as Knowledge Engine. Semantic Web 3 (3), 213-221, 2012.
- Gary Berg-Cross, Isabel Cruz, Mike Dean, Tim Finin, Mark Gahegan, Pascal Hitzler, Hook Hua, Krzysztof Janowicz, Naicong Li, Philip Murphy, Bryce Nordgren, Leo Obrst, Mark Schildhauer, Amit Sheth, Krishna Sinha, Anne Thessen, Nancy Wiegand, Ilya Zaslavsky, Semantics and Ontologies for EarthCube. In: K. Janowicz, C. Kessler, T. Kauppinen, D. Kolas, S. Scheider (eds.), Workshop on GIScience in the Big Data Age, In conjunction with the seventh International Conference on Geographic Information Science 2012 (GIScience 2012), Columbus, Ohio, USA. September 18th, 2012. Proceedings.
- Krzysztof Janowicz, Pascal Hitzler, Thoughts on the Complex Relation Between Linked Data, Semantic Annotations, and Ontologies. In: Paul N. Bennett, Evgeniy Gabrilovich, Jaap Kamps, Jussi Karlgren (eds.), Proceedings of the 6th International Workshop on Exploiting Semantic Annotation in Information Retrieval, ESAIR 2013, ACM, San Francisco, 2013, pp. 41-44.

References

- Prateek Jain, Pascal Hitzler, Amit P. Sheth, Kunal Verma, Peter Z. Yeh, Ontology Alignment for Linked Open Data. In P. Patel-Schneider, Y. Pan, P. Hitzler, P. Mika, L. Zhang, J. Pan, I. Horrocks, B. Glimm (eds.), The Semantic Web - ISWC 2010. 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-11, 2010, Revised Selected Papers, Part I. Lecture Notes in Computer Science Vol. 6496. Springer, Berlin, 2010, pp. 402-417.
- Amit Krishna Joshi, Prateek Jain, Pascal Hitzler, Peter Z. Yeh, Kunal Verma, Amit P. Sheth, Mariana Damova, Alignment-based Querying of Linked Open Data. In: Meersman, R.; Panetto, H.; Dillon, T.; Rinderle-Ma, S.; Dadam, P.; Zhou, X.; Pearson, S.; Ferscha, A.; Bergamaschi, S.; Cruz, I.F. (eds.), On the Move to Meaningful Internet Systems: OTM 2012, Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2012, Rome, Italy, September 10-14, 2012, Proceedings, Part II. Lecture Notes in Computer Science Vol. 7566, Springer, Heidelberg, 2012, pp. 807-824.
- Yingjie Hu, Krzysztof Janowicz, David Carral, Simon Scheider, Werner Kuhn, Gary Berg-Cross, Pascal Hitzler, Mike Dean, Dave Kolas, A Geo-Ontology Design Pattern for Semantic Trajectories. In: Thora Tenbrink, John G. Stell, Antony Galton, Zena Wood (Eds.): Spatial Information Theory - 11th International Conference, COSIT 2013, Scarborough, UK, September 2-6, 2013. Proceedings. Lecture Notes in Computer Science Vol. 8116, Springer, 2013, pp. 438-456.

References

- Prateek Jain, Peter Z. Yeh, Kunal Verma, Reymonrod G. Vasquez, Mariana Damova, Pascal Hitzler, Amit P. Sheth, Contextual Ontology Alignment of LOD with an Upper Ontology: A Case Study with Proton. In: Grigoris Antoniou, Marko Grobelnik, Elena Paslaru Bontas Simperl, Bijan Parsia, Dimitris Plexousakis, Pieter De Leenheer, Jeff Pan (Eds.): The Semantic Web: Research and Applications - 8th Extended Semantic Web Conference, ESWC 2011, Heraklion, Crete, Greece, May 29-June 2, 2011, Proceedings, Part I. Lecture Notes in Computer Science 6643, Springer, 2011, pp. 80-92.
- Prateek Jain, Pascal Hitzler, Kunal Verma, Peter Yeh, Amit Sheth, Moving beyond sameAs with PLATO: Partonomy detection for Linked Data. In: Ethan V. Munson, Markus Strohmaier (Eds.): 23rd ACM Conference on Hypertext and Social Media, HT '12, Milwaukee, WI, USA, June 25-28, 2012. ACM, 2012, pp. 33-42.

References

- David Carral Martinez, Krzysztof Janowicz, Pascal Hitzler, A Logical Geo-Ontology Design Pattern for Quantifying over Types. In: Isabel F. Cruz, Craig Knoblock, Peer Kröger, Egemen Tanin, Peter Widmayer (Eds.): SIGSPATIAL 2012 International Conference on Advances in Geographic Information Systems (formerly known as GIS), SIGSPATIAL'12, Redondo Beach, CA, USA, November 7-9, 2012. ACM 2012, pp. 239-248.
- David Carral, Simon Scheider, Krzysztof Janowicz, Charles Vardeman, Adila A. Krisnadhi, Pascal Hitzler, An Ontology Design Pattern for Cartographic Map Scaling. In: Philipp Cimiano, Oscar Corcho, Valentina Presutti, Laura Hollink, Sebastian Rudolph (Eds.), The Semantic Web: Semantics and Big Data. 10th International Conference, ESWC 2013, Montpellier, France, May 26-30, 2013. Proceedings. Lecture Notes in Computer Science Vol. 7882, Springer, Heidelberg, 2013, pp. 76-93.