



Combining Learning and Reasoning: A Semantic Web Perspective

Pascal Hitzler

Data Semantics Laboratory (DaSe Lab)
Data Science and Security Cluster (DSSC)
Wright State University
<http://www.pascal-hitzler.de>



Pascal Hitzler, Markus Krötzsch,
Sebastian Rudolph

Foundations of Semantic Web
Technologies

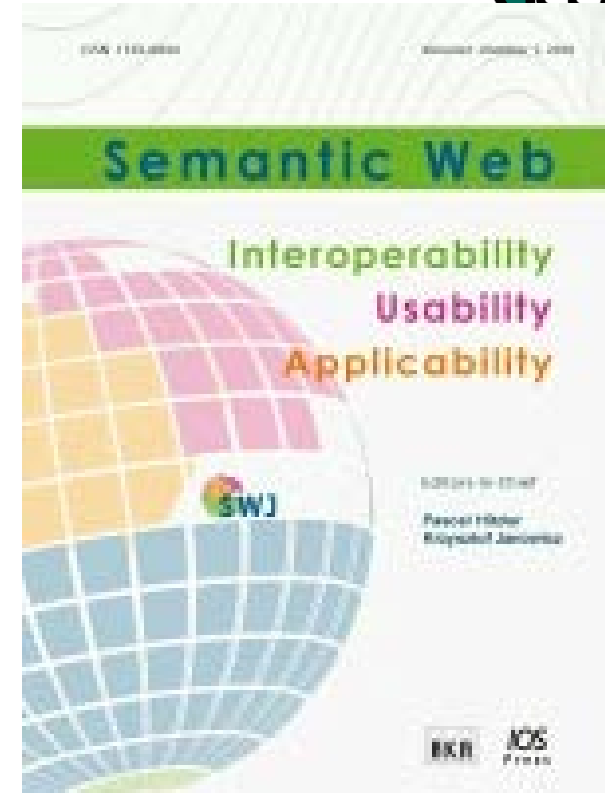
Chapman & Hall/CRC, 2010

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

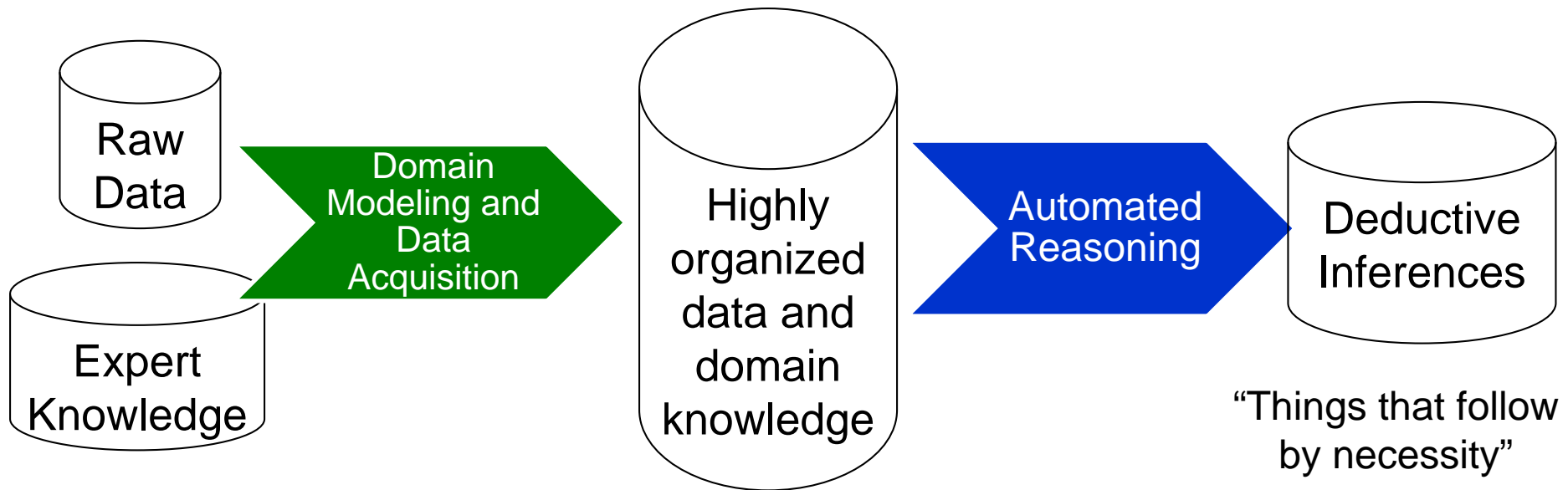
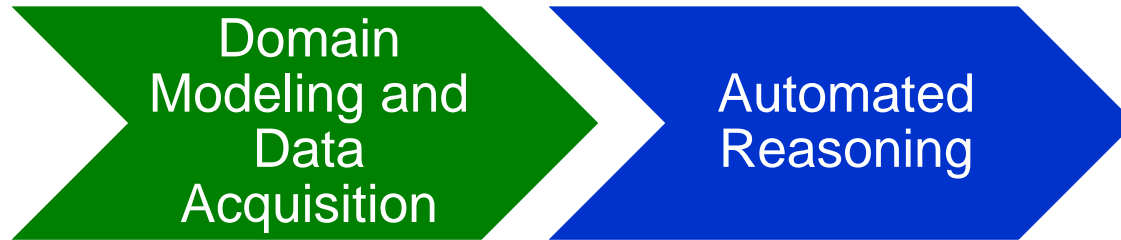
<http://www.semantic-web-book.org>

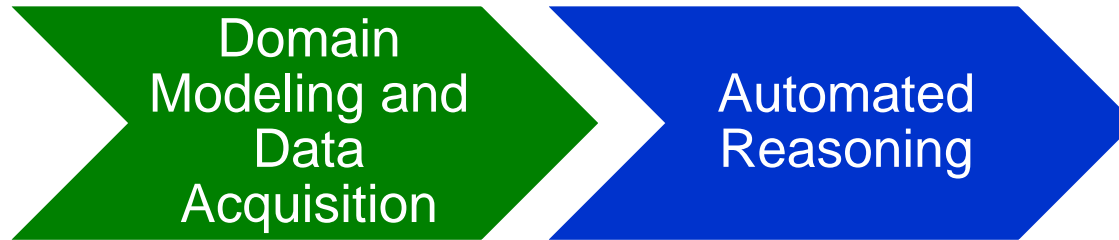


- **EiCs:** Pascal Hitzler
Krzysztof Janowicz
- **Funded 2010**
- **2017 Impact factor of 2.889, top (with 1.3 distance) of all journals with “Web” in the title**
- **We very much welcome contributions at the “rim” of traditional Semantic Web research – e.g., work which is strongly inspired by a different field.**
- **Non-standard (open & transparent) review process.**



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





“Expert Systems” traditionally based on the logic programming paradigm.

In the wake of Semantic Web Technologies, Description Logics have emerged as alternative dominant paradigm.

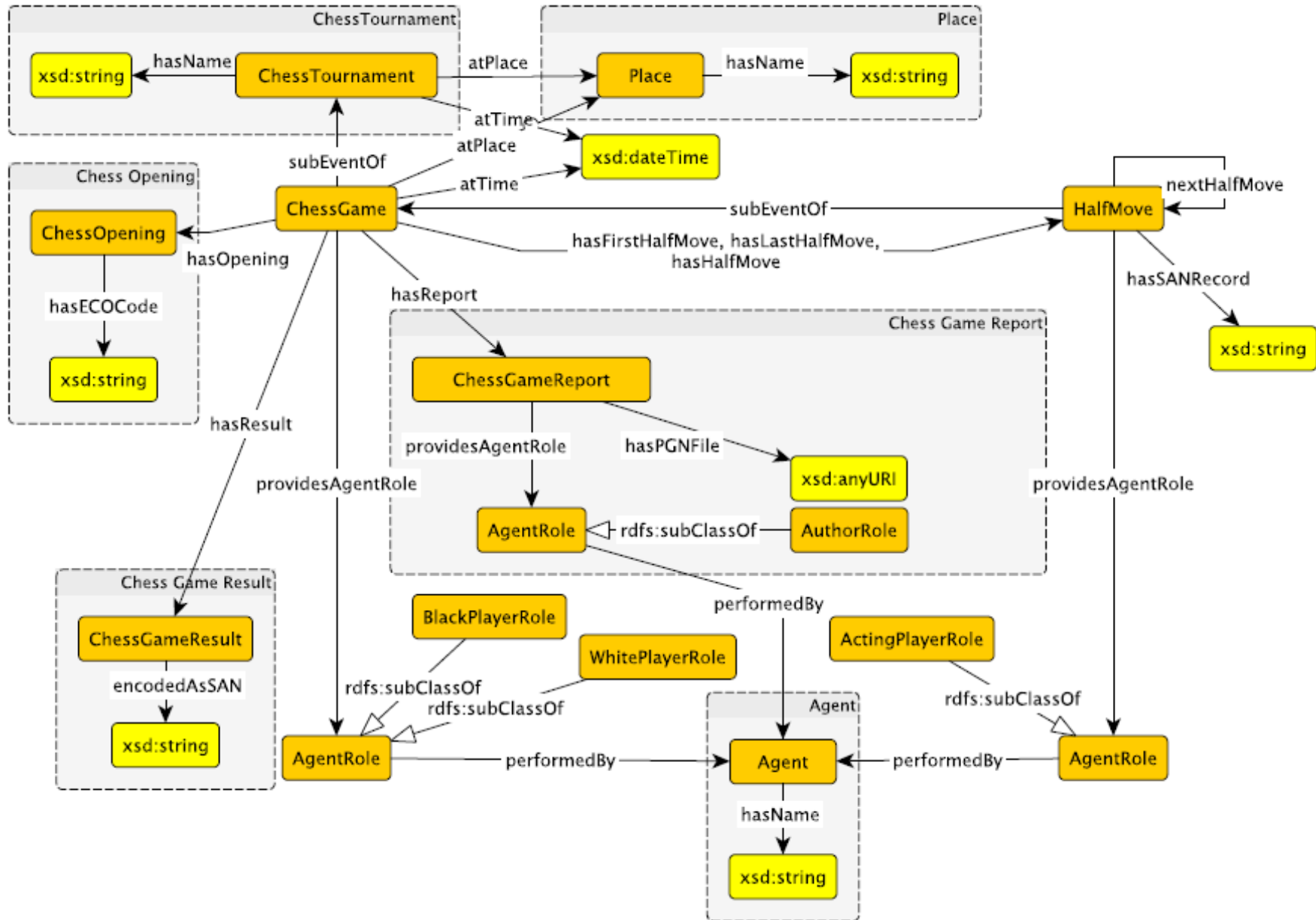
Description logics underlie the W3C standard “Web Ontology Language” (OWL).



- **Unary and binary predicates only (classes = concepts
properties = roles)**
- **Conjunction, disjunction, negation and restricted quantifiers to form complex class expressions**
- **SubClass relationships (i.e., implications between complex classes)**
- **Role chains (concatenation of binary predicates)**
- **Variable bindings are severely restricted.**

Result:

- **Decidable logics, with complexities known.**
- **Variable-free notation.**
- **Some are P-complete; some are 2NExpTime**
- **Very efficient reasoners available.**



AgentRole \sqsubseteq ($=1$ performedBy.Agent) \sqcap \forall performedBy.Agent (10.1)

\exists performedBy.Agent \sqsubseteq AgentRole (10.2)

\top \sqsubseteq \forall pAR.AgentRole (10.3)

ChessGame \sqsubseteq \exists atPlace.Place \sqcap \forall atPlace.Place (10.4)

ChessGame \sqsubseteq \exists atTime.xsd:dateTime \sqcap \forall atTime.xsd:dateTime (10.5)

ChessGame \sqsubseteq \exists pAR.BlackPlayerRole \sqcap \exists pAR.WhitePlayerRole (10.6)

\exists subEventOf.ChessTournament \sqcup \exists hasOpening.ChessOpening \sqsubseteq ChessGame (10.7)

\exists hasResult.ChessGameResult \sqcup \exists hasReport.ChessGameReport \sqsubseteq ChessGame (10.8)

ChessGame \sqsubseteq \forall subEventOf.ChessTournament \sqcap \forall hasOpening.ChessOpening (10.9)

ChessGame \sqsubseteq \forall hasResult.ChessGameResult \sqcap \forall hasReport.ChessGameReport (10.10)

BlackPlayerRole \sqcup WhitePlayerRole \sqsubseteq AgentRole \sqcap ($=1$ pAR⁻.ChessGame) (10.11)

ChessGame \sqsubseteq ($=1$ hasFirstHalfMove.HalfMove) \sqcap ($=1$ hasLastHalfMove.HalfMove) (10.12)

ChessGame \sqsubseteq ($=1$ hasLastHalfMove.HalfMove) (10.13)

hasHalfMove \sqsubseteq subEventOf⁻ (10.14)

hasFirstHalfMove \sqsubseteq hasHalfMove (10.15)

hasLastHalfMove \sqsubseteq hasHalfMove (10.16)

HalfMove \sqsubseteq Event \sqcap \exists pAR.ActingPlayerRole \sqcap ($=1$ hasHalfMove⁻.ChessGame) (10.17)

ActingPlayerRole \sqsubseteq AgentRole \sqcap ($=1$ pAR⁻.HalfMove) (10.18)

HalfMove \sqsubseteq (≤ 1 nextHalfMove.HalfMove) \sqcap $\neg \exists$ nextHalfMove.Self (10.19)

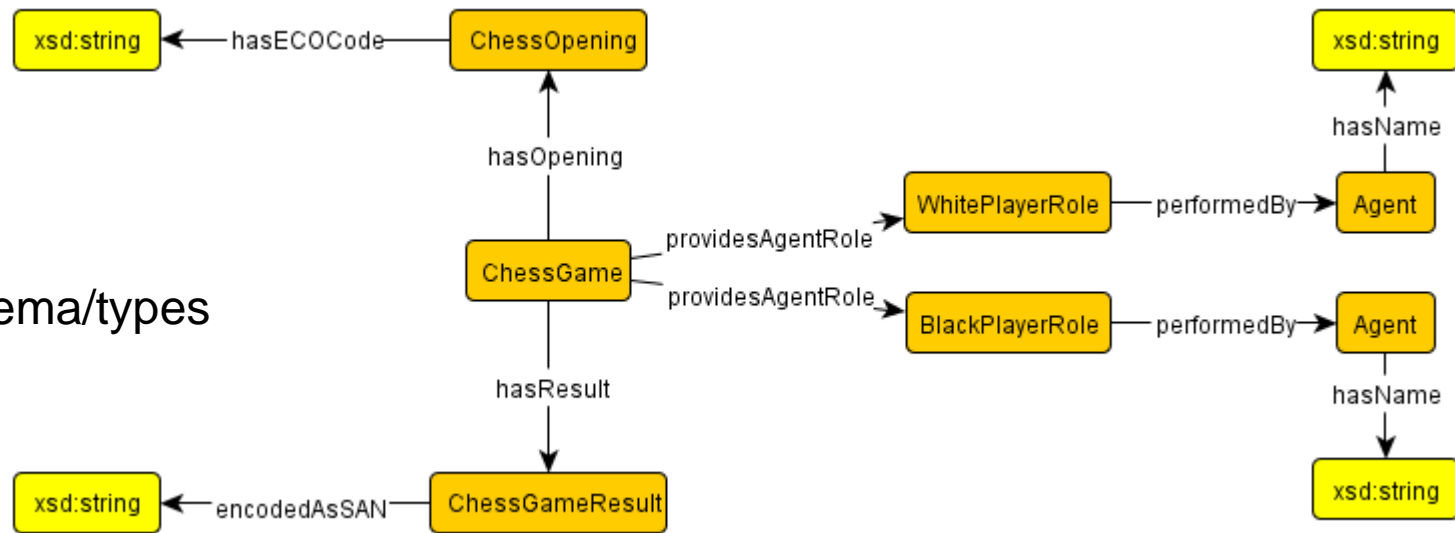
\exists subEventOf.ChessGame \sqcup \exists nextHalfMove.HalfMove \sqsubseteq HalfMove (10.20)

\exists hasSANRecord.xsd:string \sqsubseteq HalfMove (10.21)

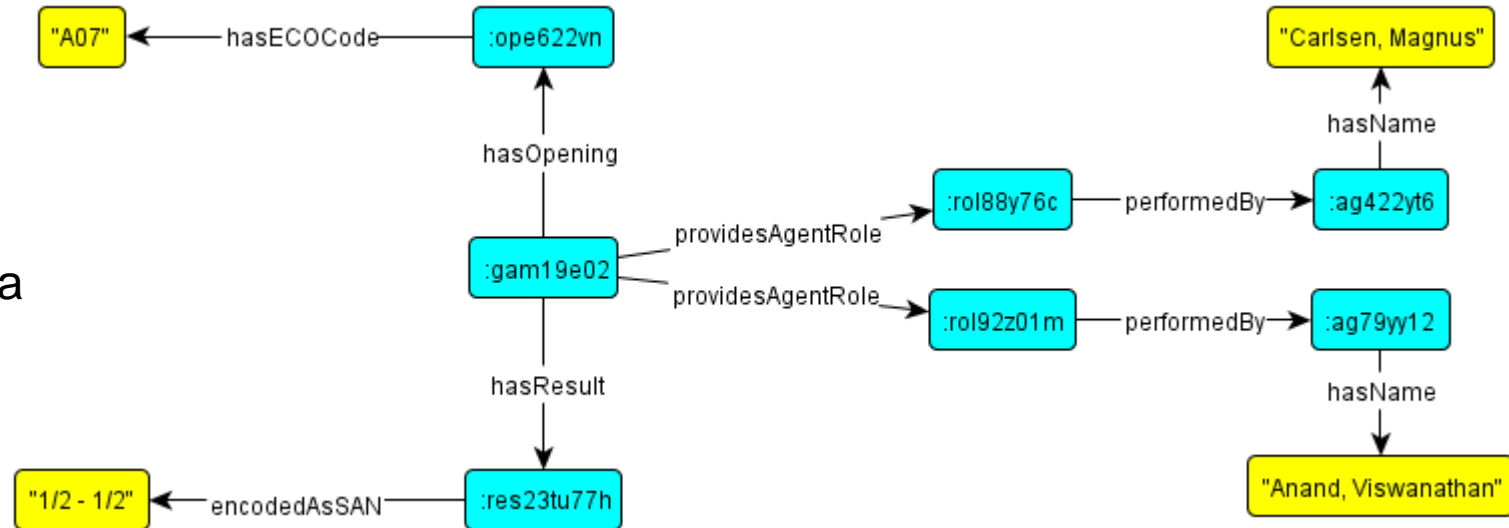
HalfMove \sqcap \forall subEventOf.ChessGame \sqcap \forall nextHalfMove.HalfMove (10.22)



Schema/types



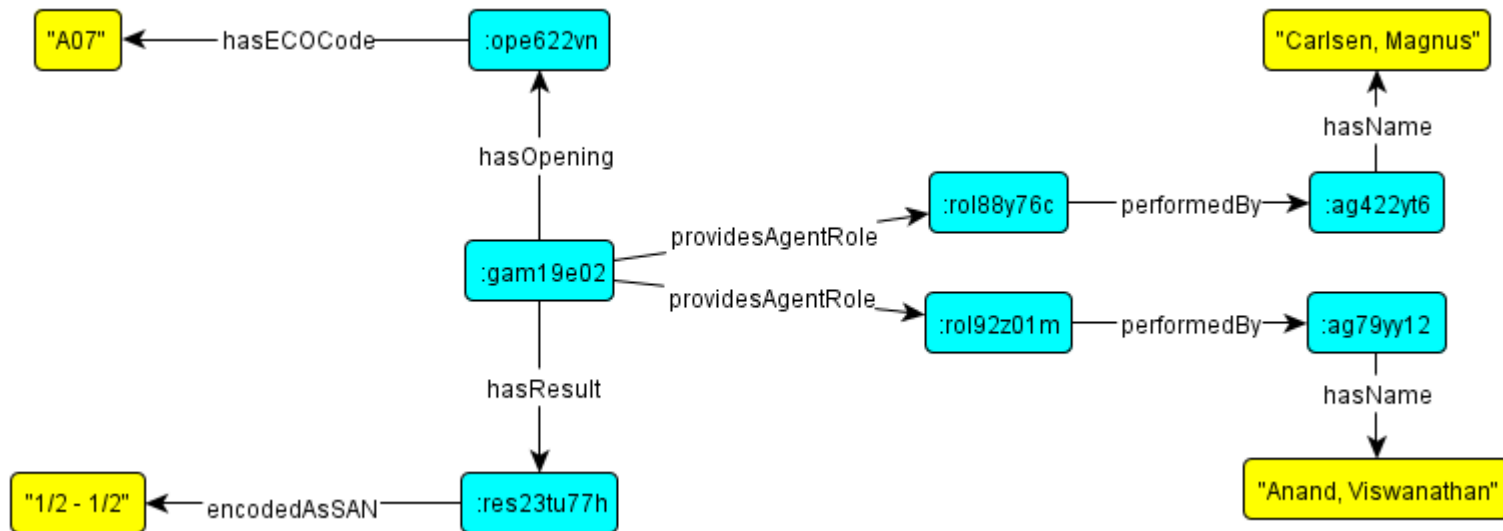
Data





RDF uses Web-referencable identifiers.

E.g., `:ag422yt6` expands to a full URI (for Magnus Carlsen).
Under best practices, it even dereferences.



Knowledge Graph

RDF Graph

Labelled Graph

Abox

Facts

Schema

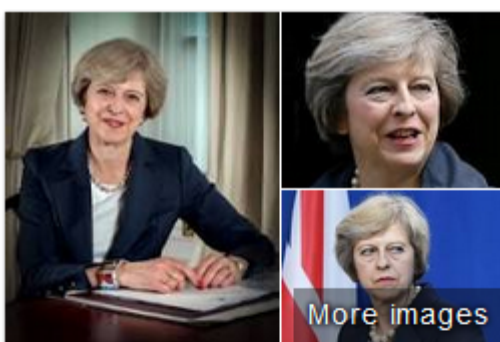
Ontology

Type Logic

Tbox

Logical Theory





More images

Theresa May



British Prime Minister



tmay.co.uk

Theresa Mary May is a British politician who has served as Prime Minister of the United Kingdom and Leader of the Conservative Party since July 2016, the second woman to hold both positions. [Wikipedia](#)

Born: October 1, 1956 (age 60), Eastbourne, United Kingdom

Height: 5' 8"

Party: Conservative Party

Spouse: Philip May (m. 1980)

Education: St Hugh's College, Oxford (1974–1977)

Previous offices: Home Secretary (2010–2016), [MORE](#) ▾

Profiles



Twitter



Facebook

People also search for [View 15+](#)



See photos

St Hugh's College, Oxford

College in Oxford, England

[Website](#)

[Directions](#)

St Hugh's College is one of the constituent colleges of the University of Oxford. It is located on a 14.5-acre site on St Margaret's Road, to the north of the city centre. [Wikipedia](#)

Address: St Margaret's Rd, Oxford OX2 6LE, UK

Principal: Elish Angiolini

Phone: +44 1865 274900

Founder: Elizabeth Wordsworth

Founded: 1886

Named for: Hugh of Lincoln

Undergraduates: 432 (2011–2012)

[Suggest an edit](#) · [Own this business?](#)

Reviews from the web

4.1/5 [University Rooms](#) - 2,310 votes

[Send to your phone](#)

[Send](#)

Notable alumni

[View 40+](#)



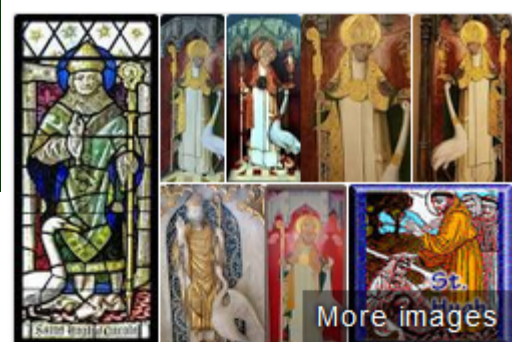
Theresa May



Aung San Suu Kyi



Barbara Castle



More images

Hugh of Lincoln



Saint

Hugh of Lincoln, also known as Hugh of Avalon, was a French noble, Benedictine and Carthusian monk, bishop of Lincoln in the Kingdom of England, and Catholic saint. [Wikipedia](#)

Born: 1140, Avalon, France

Died: November 16, 1200, London, United Kingdom

Feast: 16 November (R.C.C.); 17 November (Anglican)

Major shrine: Lincoln Cathedral

Attributes: a white swan

Patronage: sick children, sick people, shoemakers and swans

People also search for



Little Saint Hugh of Lincoln



Thomas More



William Howard, 1st Visco...

ab



Schema.org



- Collaboratively launched in 2011 by Google, Microsoft, Yahoo, Yandex.
2011: 297 classes, 187 relations
2015: 638 classes, 965 relations
- Simple schema, request to web site providers to annotate their content with schema.org markup. Promise: They will make better searches based on this.
- 2015: 31.3% of Web pages have schema.org markup, on average 26 assertions per page.

Ramanathan V. Guha, Dan Brickley, Steve Macbeth:
Schema.org: Evolution of Structured Data on the
Web. ACM Queue 13(9): 10 (2015)

- [TrainTrip](#)
- [Organization](#)
 - [Airline](#)
 - [Corporation](#)
 - [EducationalOrganization](#)
 - [CollegeOrUniversity](#)
 - [ElementarySchool](#)
 - [HighSchool](#)
 - [MiddleSchool](#)
 - [Preschool](#)
 - [School](#)
 - [GovernmentOrganization](#)
 - [LocalBusiness](#)
 - [AnimalShelter](#)
 - [AutomotiveBusiness](#)
 - [AutoBodyShop](#)
 - [AutoDealer](#)
 - [AutoPartsStore](#)
 - [AutoRental](#)
 - [AutoRepair](#)
 - [AutoWash](#)
 - [GasStation](#)
 - [MotorcycleDealer](#)
 - [MotorcycleRepair](#)
 - [ChildCare](#)
 - [Dentist](#)
 - [DryCleaningOrLaundry](#)
 - [EmergencyService](#)
 - [FireStation](#)
 - [Hospital](#)
 - [PoliceStation](#)
 - [EmploymentAgency](#)
 - [EntertainmentBusiness](#)
 - [AdultEntertainment](#)
 - [AmusementPark](#)
 - [ArtGallery](#)
 - [Casino](#)
 - [ComedyClub](#)
 - [MovieTheater](#)
 - [NightClub](#)
 - [FinancialService](#)
 - [AccountingService](#)
 - [AutomatedTeller](#)
 - [BankOrCreditUnion](#)
 - [InsuranceAgency](#)
 - [FoodEstablishment](#)
 - [Bakery](#)
 - [BarOrPub](#)
 - [Brewery](#)
 - [CafeOrCoffeeShop](#)
 - [FastFoodRestaurant](#)



- Main page
- Community portal
- Project chat
- Create a new item
- Recent changes
- Random item
- Query Service
- Nearby
- Help
- Donate

Print/export

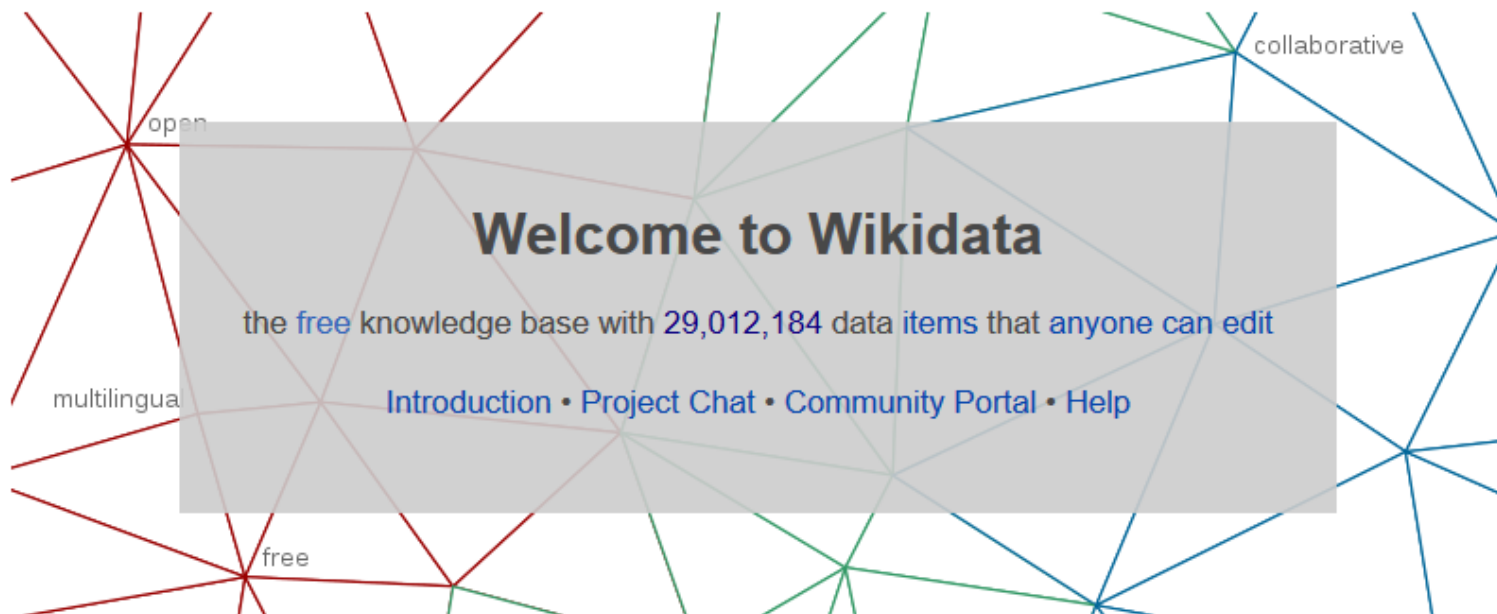
- Create a book
- Download as PDF
- Printable version

In other projects

- Wikimedia Commons
- MediaWiki
- Meta-Wiki
- Wikispecies
- Wikibooks
- Wikinews
- Wikipedia
- Wikiquote
- Wikisource
- Wikiversity
- Wikivoyage
- Wiktionary

Tools

What links here



Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

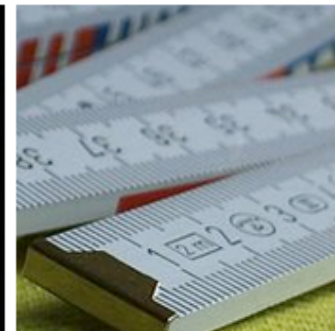
Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource, and others.

Wikidata also provides support to many other sites and services beyond just Wikimedia projects! The content of Wikidata is available under a [free license](#), [exported using standard formats](#), and can be [interlinked to other open data sets](#) on the linked data web.



Learn about data

New to the wonderful world of data? [Develop and improve your data literacy through content](#) designed to get you up to speed and feeling comfortable with the fundamentals in no time.



A bit older but somewhat more expressive: Linked Data on the Web



Number of Datasets	2017-01-26	1,146
	2014-08-30	570
	2011-09-19	295
	2010-09-22	203
	2009-07-14	95
	2008-09-18	45
	2007-10-08	25
	2007-05-01	12

38.606.408.854 triples and counting!

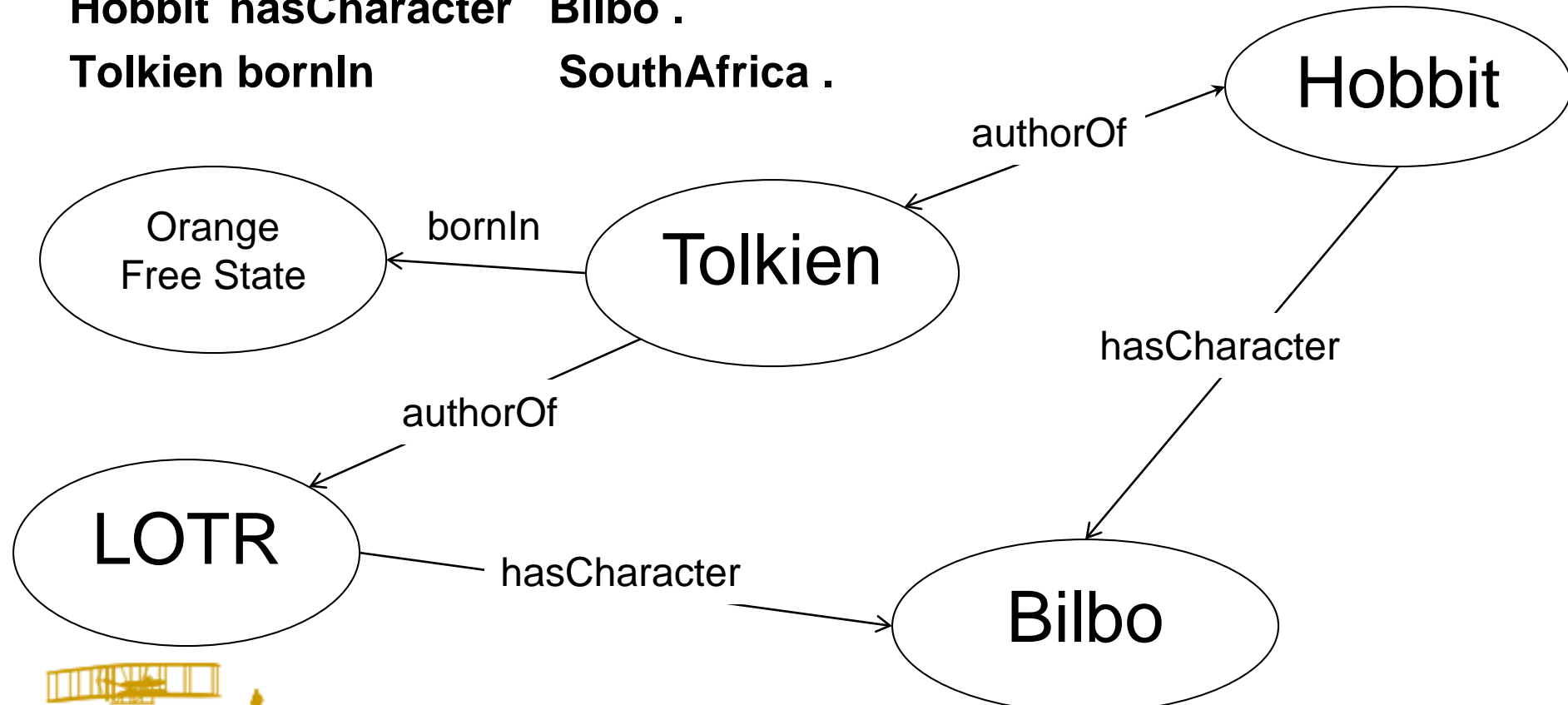


LOD Laundromat

Information as RDF graph



LOTR hasAuthor Tolkien .
Hobbit hasAuthor Tolkien .
LOTR hasCharacter Bilbo .
Hobbit hasCharacter Bilbo .
Tolkien bornIn SouthAfrica .



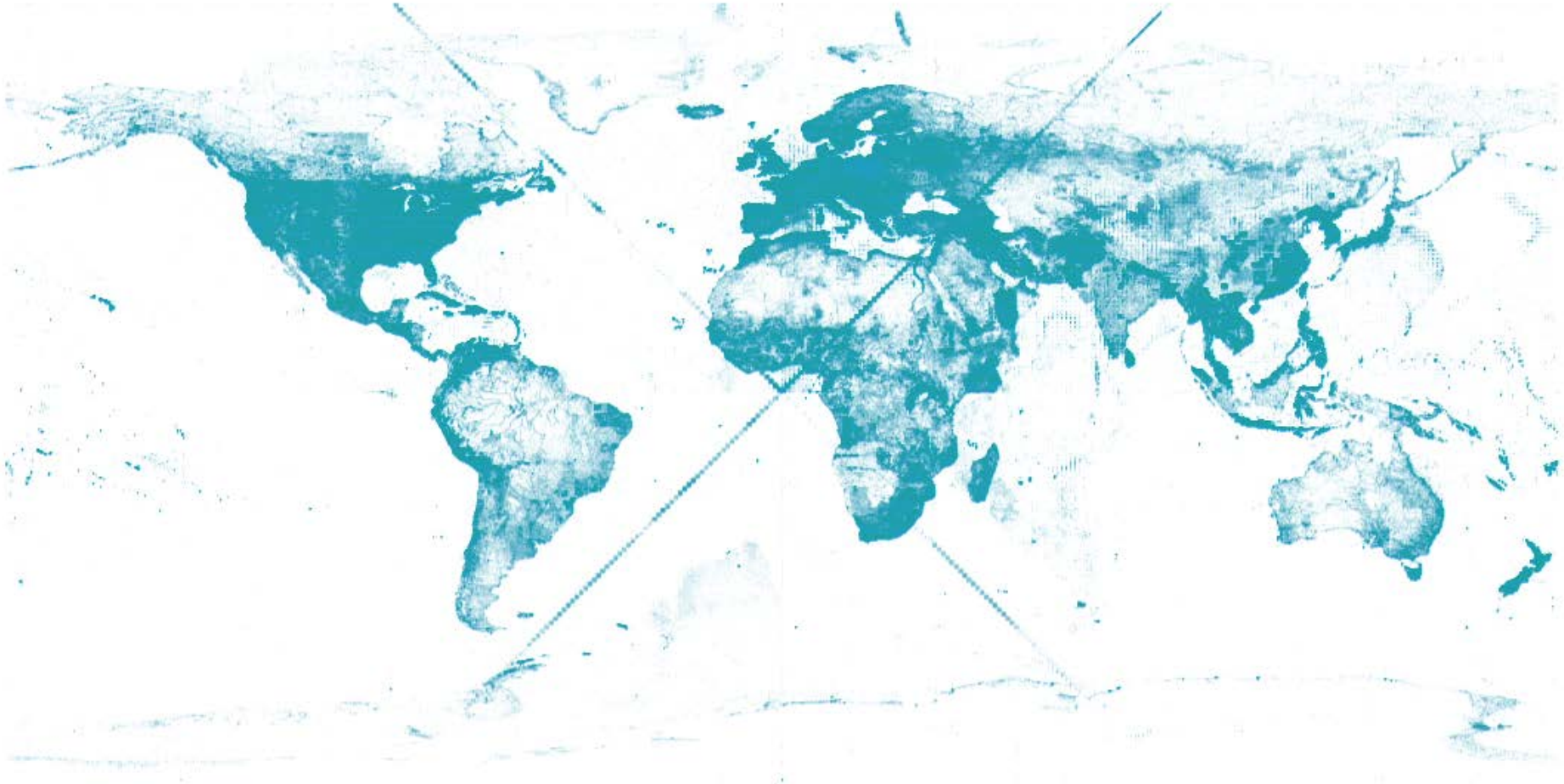
DBpedia: LOTR page

dbpedia-owl:thumbnail	<ul style="list-style-type: none">▪ http://upload.wikimedia.org/wikipedia/commons/thumb/6/62/Jrrt_lotr_cover_design.jpg/200px-Jrrt_lotr_cover_design.jpg
dbpedia-owl:wikiPageExternalLink	<ul style="list-style-type: none">▪ 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	<ul style="list-style-type: none">▪ dbpedia:J._R._R._Tolkien
dbpprop:books	<ul style="list-style-type: none">▪ dbpedia:The_Two_Towers▪ dbpedia:The_Return_of_the_King▪ dbpedia:The_Fellowship_of_the_Ring▪ "Volumes:"
dbpprop:country	<ul style="list-style-type: none">▪ England
dbpprop:expiry	<ul style="list-style-type: none">▪ 20 (xsd:integer)
dbpprop:genre	<ul style="list-style-type: none">▪ dbpedia:Adventure_novel▪ dbpedia:High_fantasy
dbpprop:hasPhotoCollection	<ul style="list-style-type: none">▪ http://www4.wiwiss.fu-berlin.de/flickwrappr/photos/The_Lord_of_the_Rings
dbpprop:imageCaption	<ul style="list-style-type: none">▪ Tolkien's own cover designs for the three volumes
dbpprop:language	<ul style="list-style-type: none">▪ English
dbpprop:mediaType	<ul style="list-style-type: none">▪ Print
dbpprop:name	<ul style="list-style-type: none">▪ The Lord of the Rings
dbpprop:pages	<ul style="list-style-type: none">▪ 1216 (xsd:integer)
dbpprop:precededBy	<ul style="list-style-type: none">▪ dbpedia:The_Hobbit
dbpprop:pubDate	<ul style="list-style-type: none">▪ 21 (xsd:integer)
dbpprop:publisher	<ul style="list-style-type: none">▪ dbpedia:Allen_&_Unwin
dbpprop:small	<ul style="list-style-type: none">▪ yes
dbpprop:wikiPageUsesTemplate	<ul style="list-style-type: none">▪ dbpedia:Template:Infobox_book_series▪ dbpedia:Template:Pp-vandalism
dcterms:subject	<ul style="list-style-type: none">▪ 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

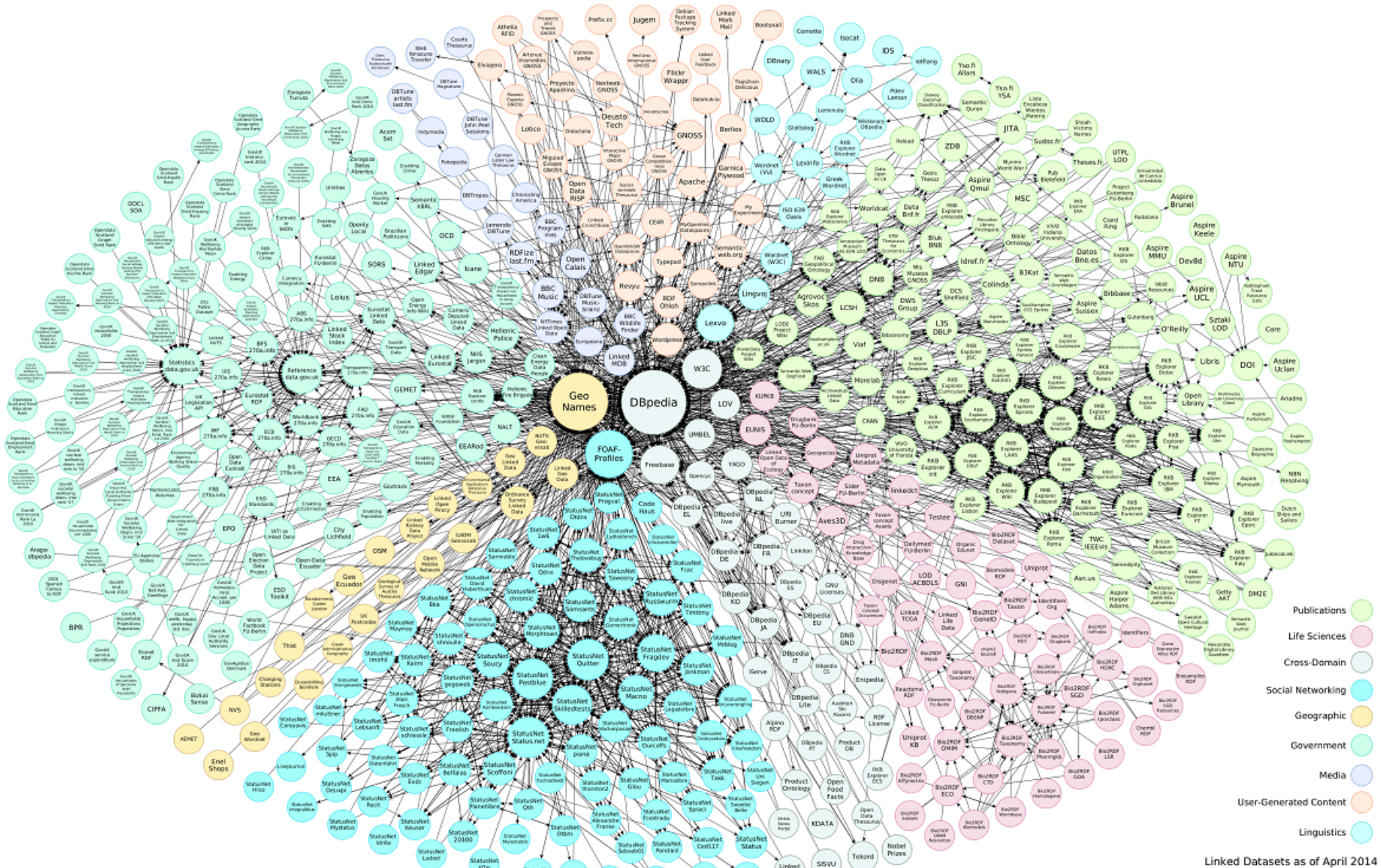
Linked Data: Volume

Geoindexed Linked Data – courtesy of Krzysztof Janowicz, 2012

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

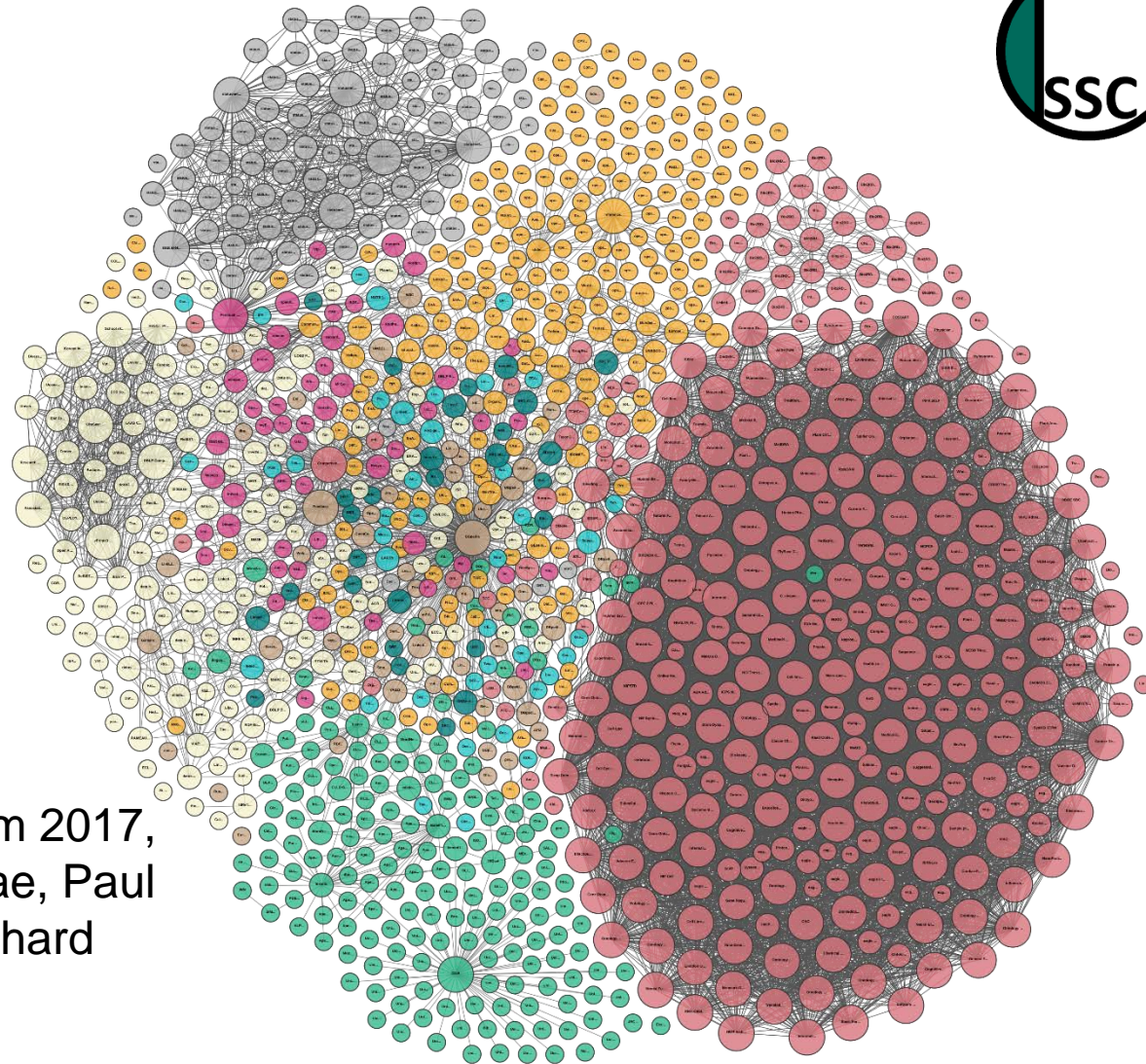
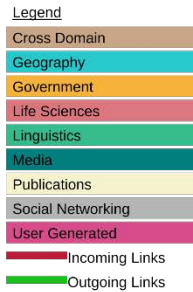


Some Linked Datasets 2014

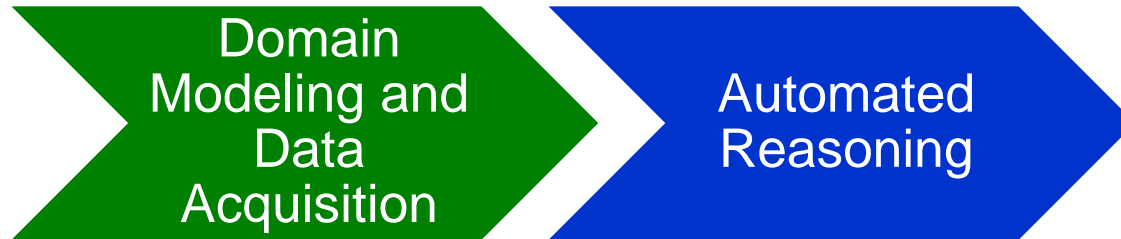


Linked Datasets as of April 2014

Some Linked Datasets 2017

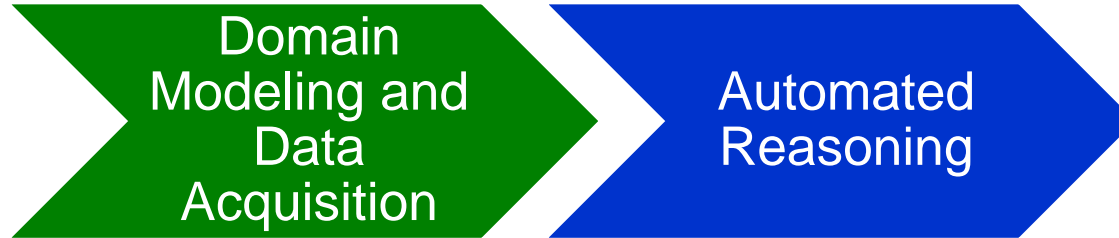


Linking Open Data cloud diagram 2017,
by Andrejs Abele, John P. McCrae, Paul
Buitelaar, Anja Jentzsch and Richard
Cygniak. <http://lod-cloud.net/>



Phylogenetic tree: Reconstructed evolutionary history/lines of descent using organismal traits etc.

“Phyloreferencing” uses ontological representations and automated reasoning for this purpose.



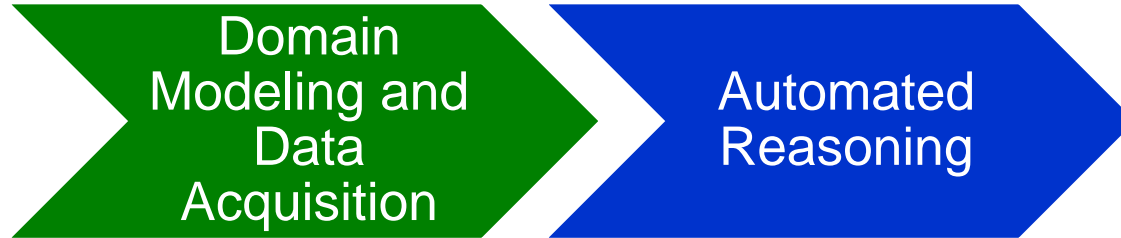
Data *management*: SWT promises:

If information is structured **well**, then this significantly reduces data management cost (discovery, reuse, repurposing, integration, revision).

Use of standards.

Best practices.

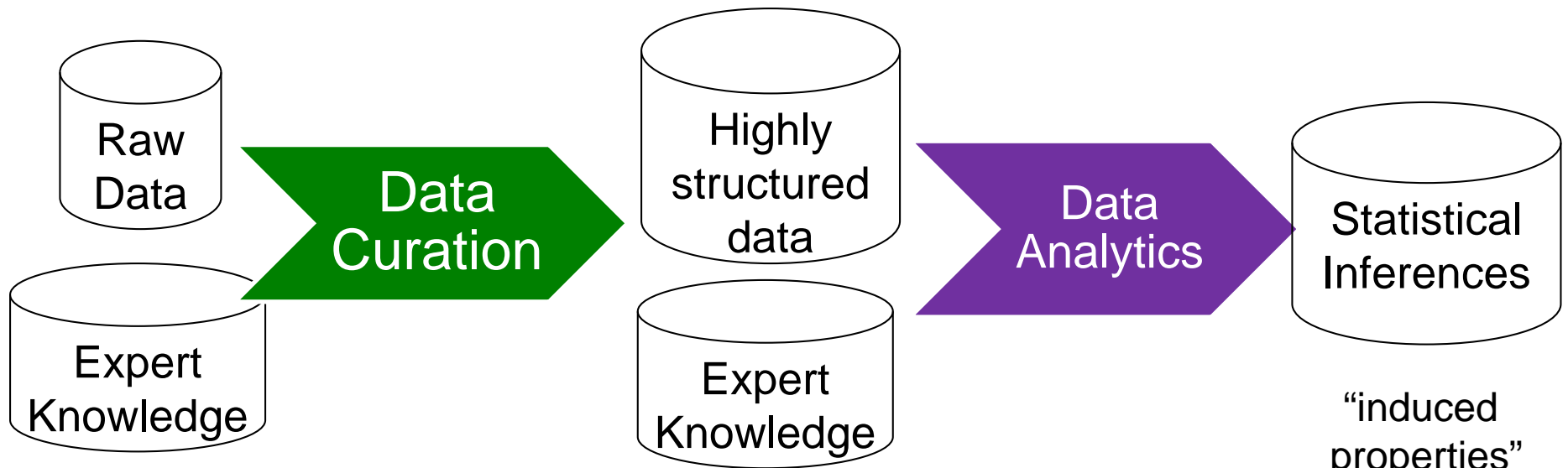
Data and ontology quality principles.



Reasoning as a tool for Data Management:

Data Curation by identifying e.g.

- inconsistencies
- violations of schema/model

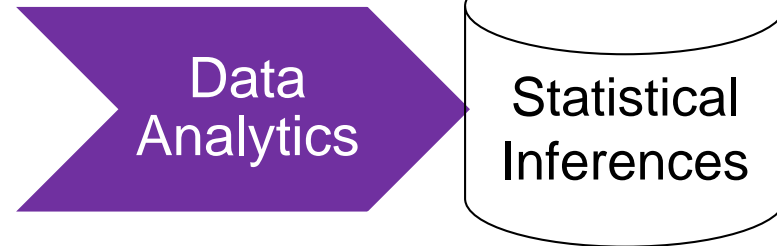




Semantic Web Technologies simplify the data curation part.

- **Easier integration of different sources.**
- **Clearer relationship between data organization and expert knowledge (i.e., data easier to understand).**
- **Easier discovery of relevant data.**

Examples:

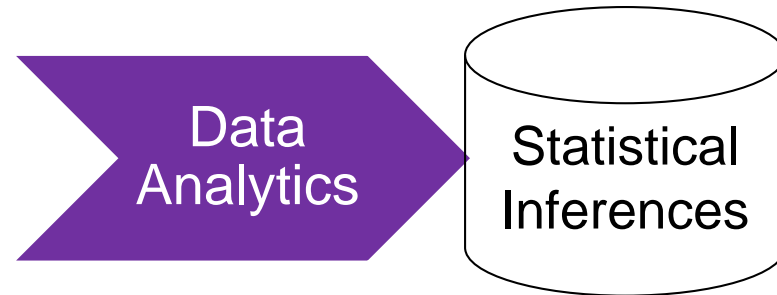


“induced
properties”

- **Scene depicted by an image.**
- **Event detection (from multi- and social media).**
- **Recommender systems (e.g., in a complex disaster scenario).**
- **Risk assessment (e.g., indicators for an epidemic).**

In most cases, you would like to know much more about what you induced.

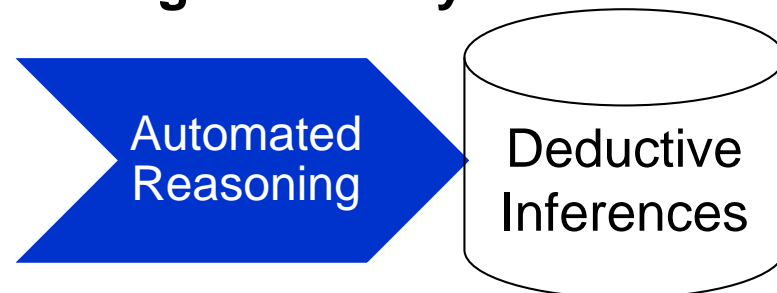
In most cases, you would like to know much more about what you induced.



“induced properties”

Two interesting cases:

- 1) Explanations for the inference.
- 2) Post-processing by knowledge-based systems.



e.g.,
disaster recommendations

“Things that follow by necessity”

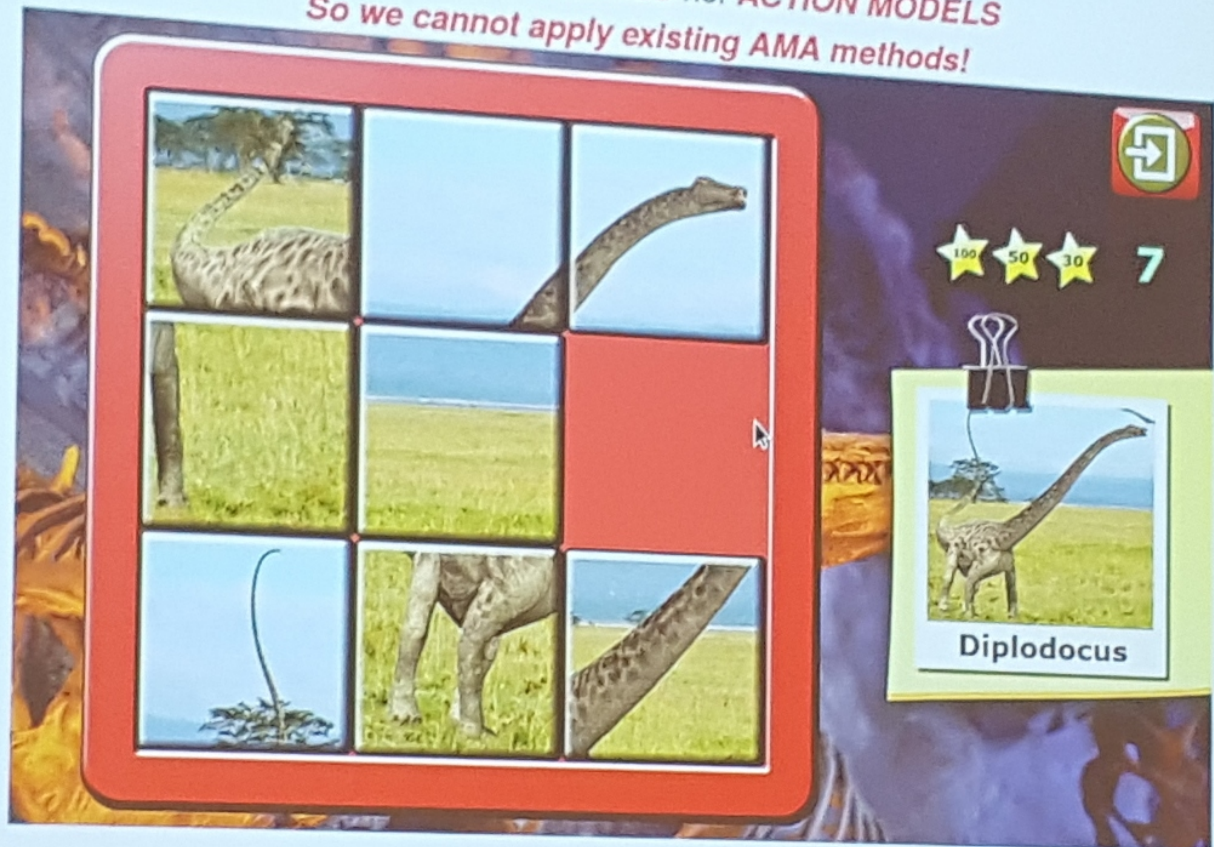


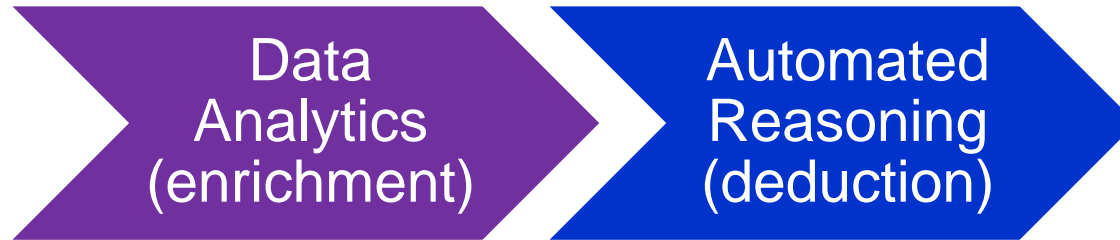
At NeSy'17

img_static_puzzle.jpg

8. Problem Description

We do not have **SYMBOLS** nor **ACTION MODELS**
So we cannot apply existing **AMA** methods!

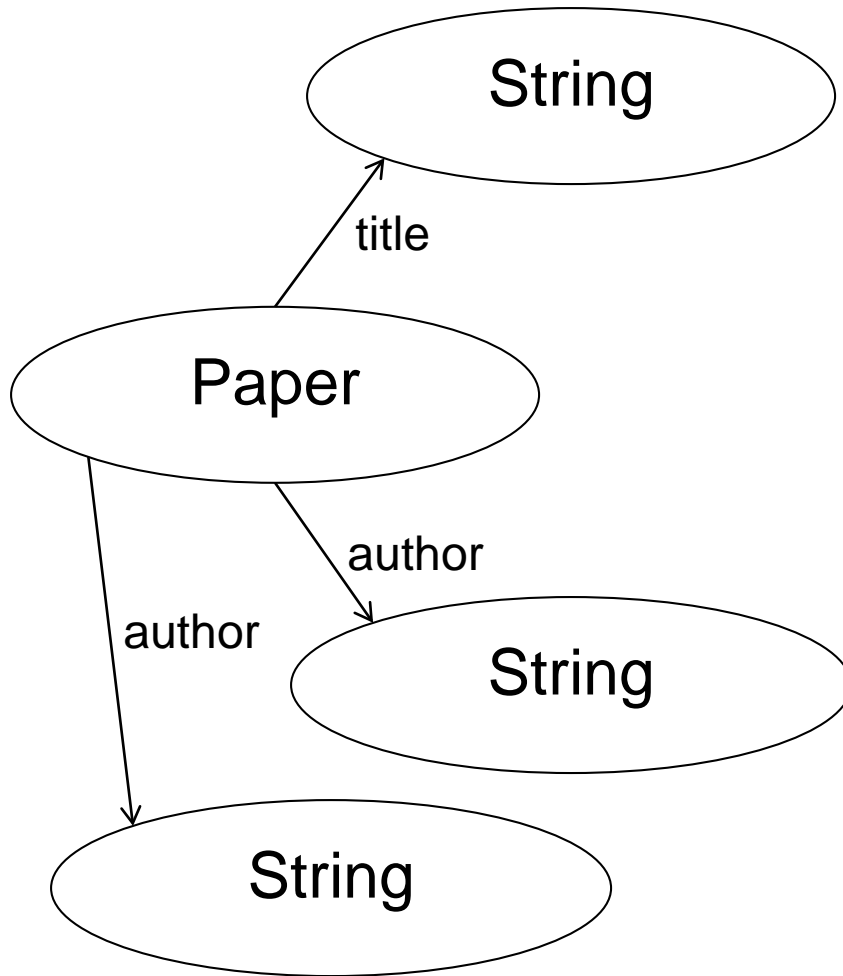




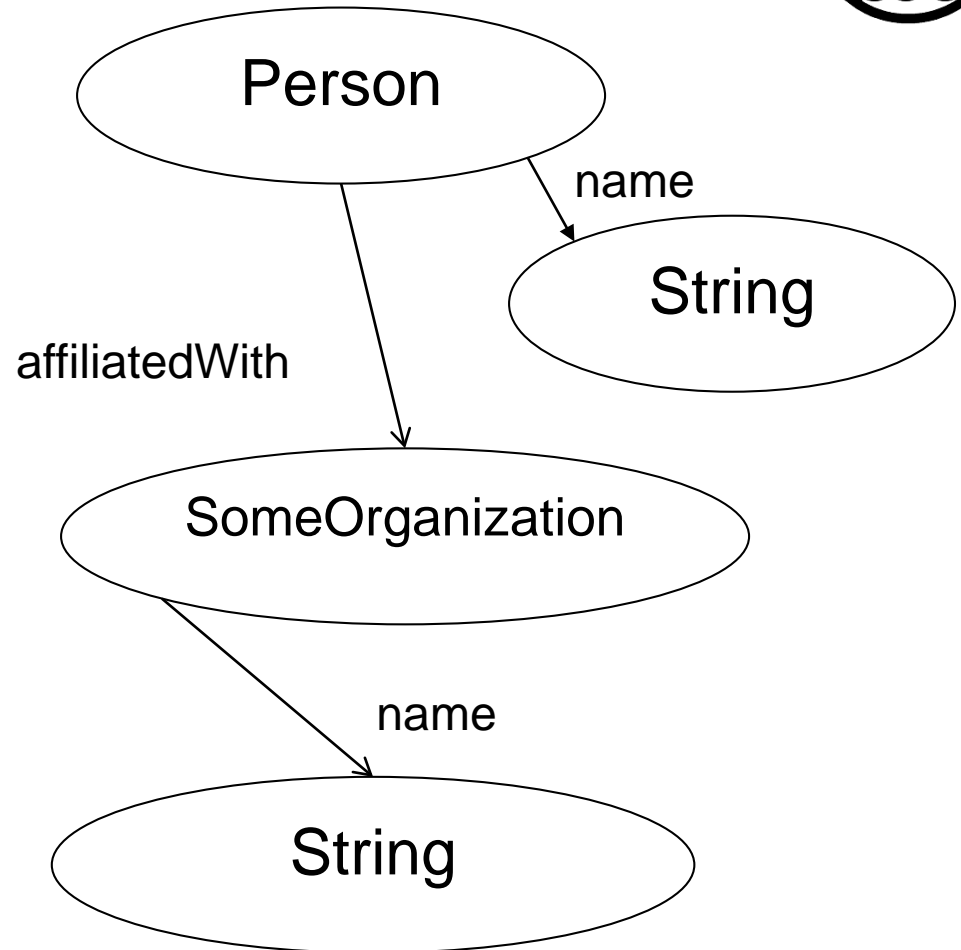
**We can hardly ever get this to work seamlessly,
even in seemingly simple scenarios.**

Find all publications from MIT

Dataset with publication data



Dataset with affiliation data





Most systems can do only equivalences.

$$\forall x(\text{SomeOrganization}(x) \longleftrightarrow \text{Organization}(x))$$

$$\forall x \forall y(\text{affiliatedWith}(x, y) \longleftrightarrow \text{hasAffiliation}(x, y))$$

Moreover, almost all of the performance of current systems is based on string similarity metrics.

[Cheatham and Hitzler, ISWC 2013]

Property subsumption

Detecting property subsumptions is still in infancy.

$$\forall x \forall y (\text{title}(x, y) \rightarrow \text{hasName}(x, y))$$

$$\forall x \forall y (\text{name}(x, y) \rightarrow \text{hasName}(x, y))$$



[Cheatham, Hitzler OM 2014]

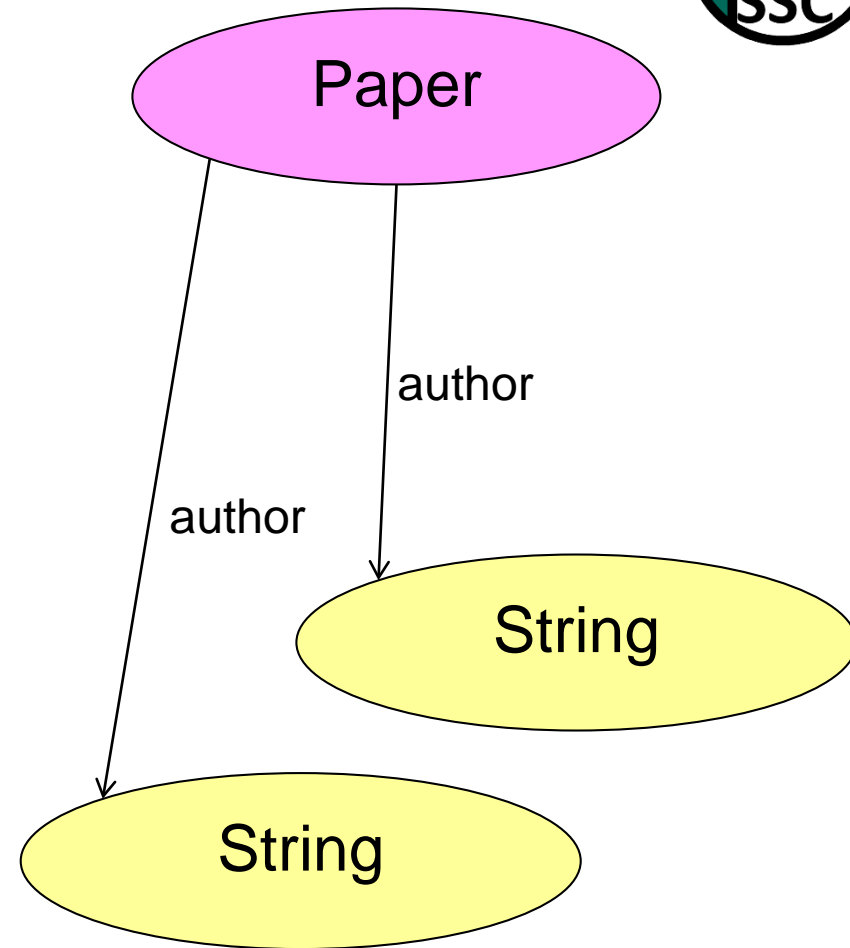
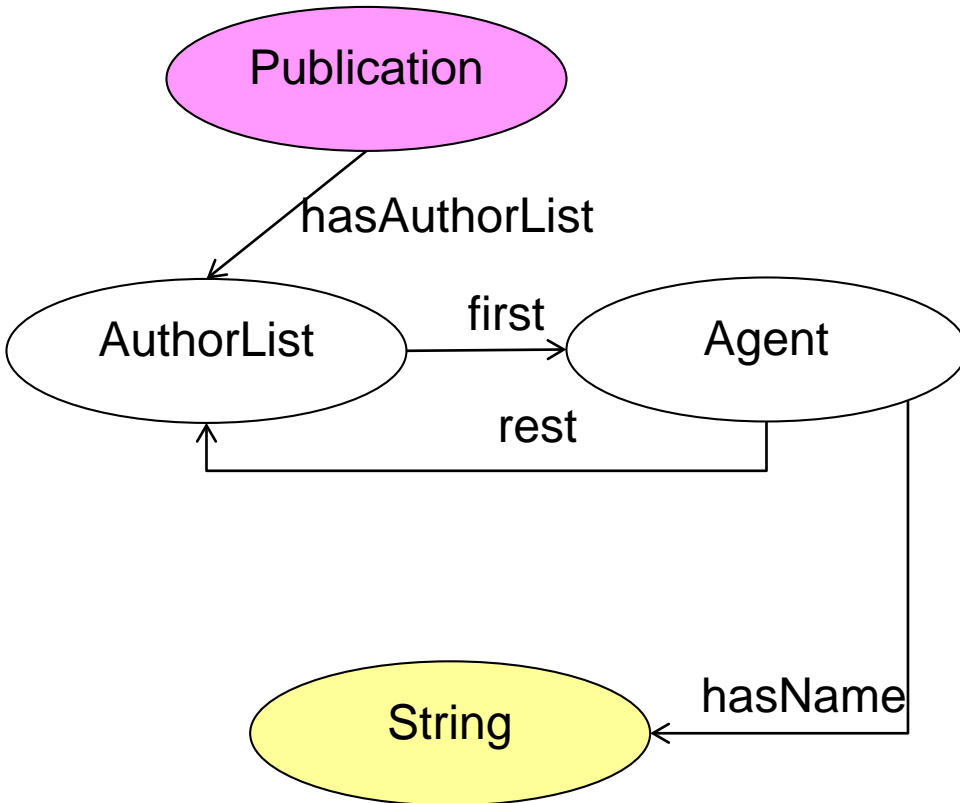
System	Class Prec	Class Rec	Class Fms	Prop Prec	Prop Rec	Prop Fms
AML	0.86	0.62	0.72	1.00	0.20	0.33
AMLback	0.86	0.64	0.73	1.00	0.24	0.39
CIDER_CL	0.46	0.59	0.52	0.07	0.22	0.11
HerTUDA	0.84	0.56	0.67	0.26	0.20	0.23
HotMatch	0.81	0.57	0.67	0.24	0.20	0.22
IAMA	0.87	0.55	0.67	0.14	0.07	0.09
LogMap	0.82	0.65	0.73	0.62	0.28	0.39

System	Precision	Recall	F-measure
PropString (prec)	1.0	0.26	0.41
PropString (rec)	0.34	0.5	0.4
Soft TF-IDF	0.2	0.24	0.22

Table 4. Results on the OAEI Conference track

YAM++	0.82	0.71	0.76	0.68	0.57	0.62
Average	0.79	0.60	0.68	0.36	0.18	0.21

We need mapping *rules*



- How to build a **good** ontology for an application domain?
 - If done by humans, it's a very evolved process.
 - Can we support this by inductive inferencing?
- E.g., by extracting structure from texts or data?





Given ontology and (separate) data or text corpora,

create a knowledge graph from the data/text corpus which adheres to the ontology as schema.

Add-on: Do it with sufficient quality such that you can meaningfully reason over the ontology+knowledge graph afterwards.



- **Voice control of your phone for all tasks as a “personal assistant”.**
- **Health care recommender system.**
- **Generally, Watson-like systems.**



Thanks!

- 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.



- **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.**
- **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.**
- **Krzysztof Janowicz, Frank van Harmelen, James A. Hendler, Pascal Hitzler, Why the Data Train Needs Semantic Rails. AI Magazine 26 (1), 2015, 5-14.**



- 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.

- 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: Paronymy 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.





- **Michelle Cheatham, Pascal Hitzler, String Similarity Metrics for Ontology Alignment.** In: H. Alani, L. Kagal, A. Fokoue, P. Groth, C. Biemann, J.X. Parreira, L. Aroyo, N. Noy, C. Welty, K. Janowicz (eds.), **The Semantic Web - ISWC 2013.** 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013, Proceedings, Part II. Lecture Notes in Computer Science Vol. 8219, Springer, Heidelberg, 2013, pp. 294-309.
- **Michelle Cheatham, Pascal Hitzler, The Properties of Property Alignment.** In: Proceedings OM-2014, The Ninth International Workshop on Ontology Matching, at the 13th International Semantic Web Conference, ISWC 2014, Riva del Garda, Trentino, Italy, October 2014. To appear.

- Cellinese, N., Lapp, H.: An Ontology-Based system for querying life in a Post-Taxonomic age (2015),
https://figshare.com/articles/An_Ontology_Based_System_for_Querying_Life_in_a_Post_Taxonomic_Age/1401984

