

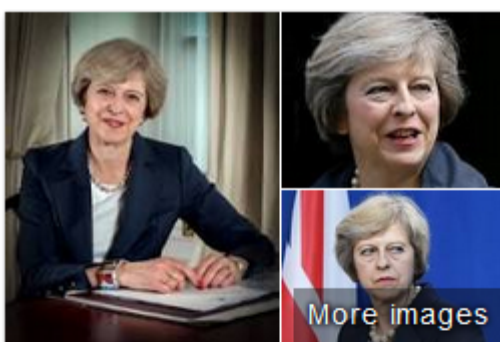


Explaining Trained Networks Through Semantic Web Background Knowledge

Pascal Hitzler

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





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+](#)



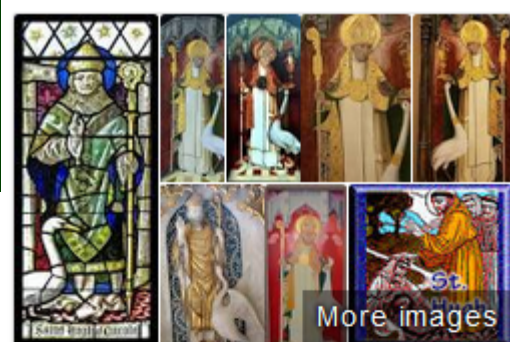
Theresa May



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

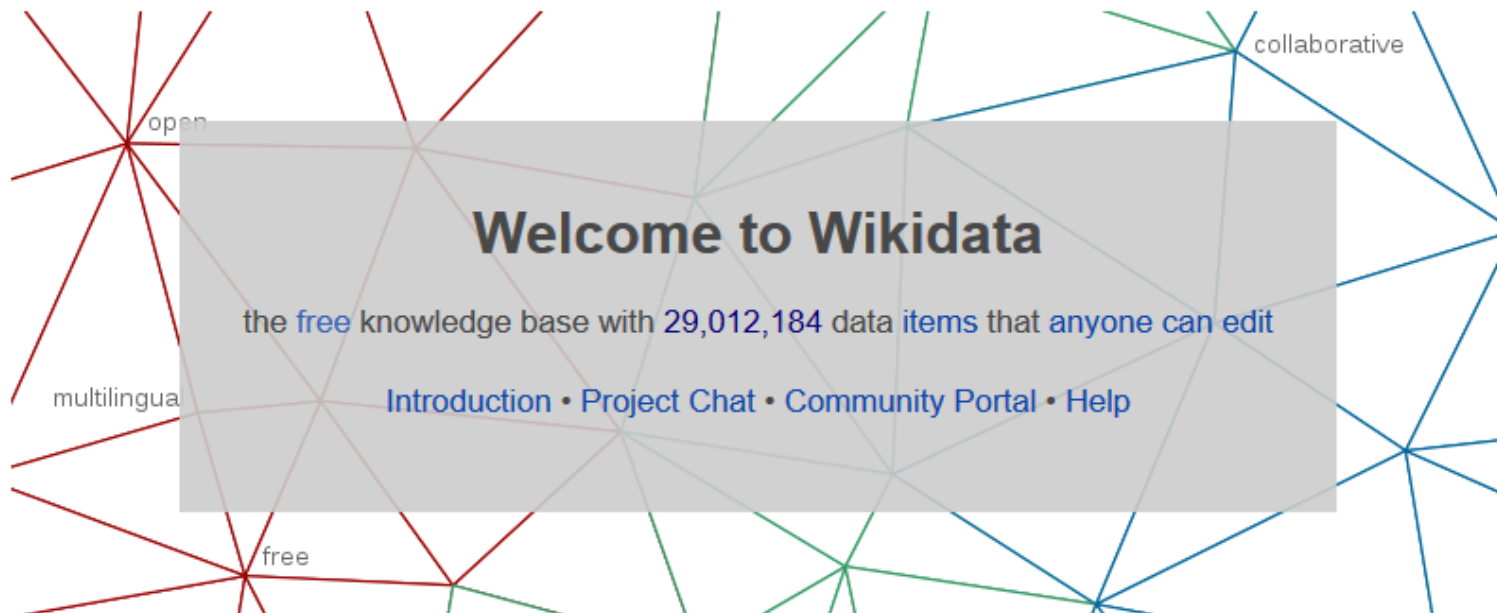
- Create a book
- Download as PDF
- Printable version

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.

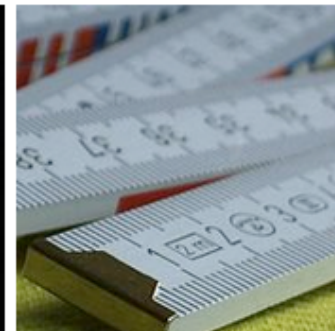
Wikidata acts as central storage for the **structured data** of its Wikimedia sister projects including Wikipedia, Wikivoyage, Wikisource, and others.

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

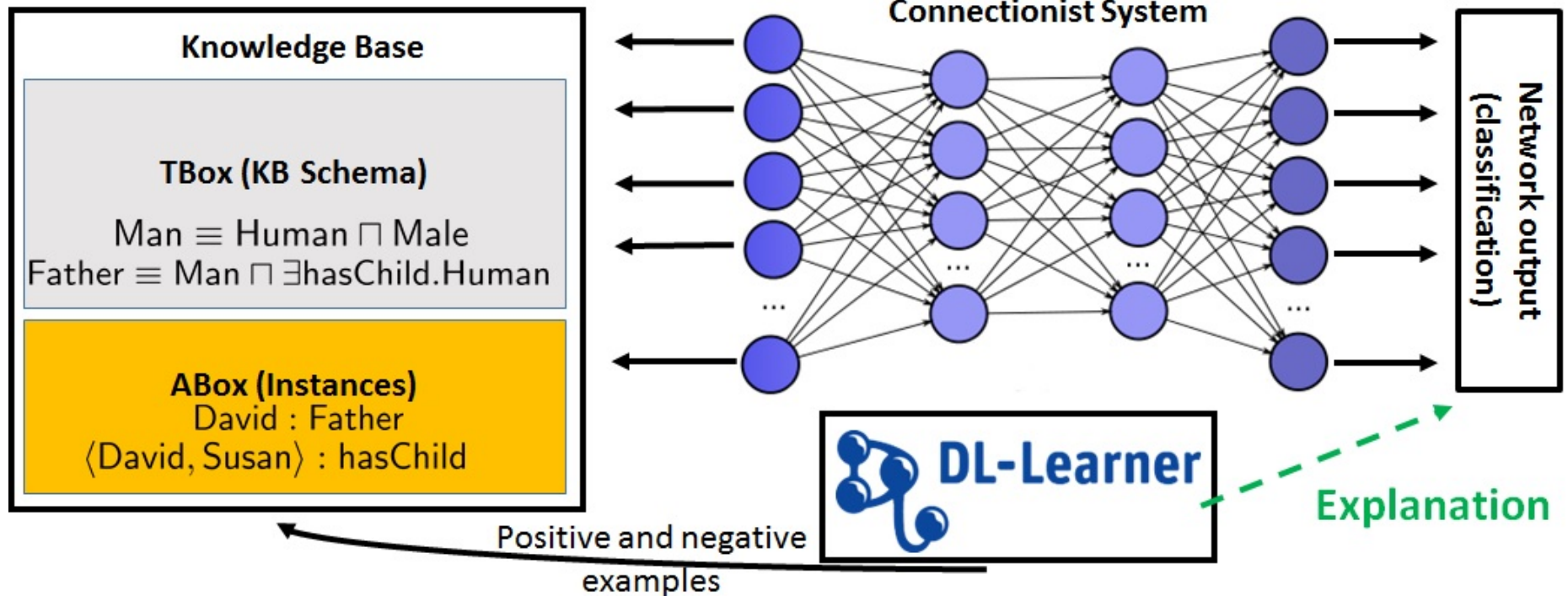
38.606.408.854 triples and counting!



LOD Laundromat



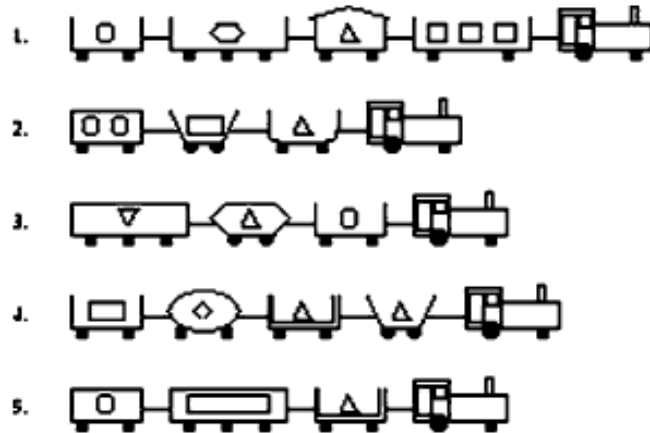
- Explain input-output behavior of trained (deep) NNs.
- Idea:
 - Use background knowledge in the form of linked data and ontologies to help explain.
 - Link inputs and outputs to background knowledge.
 - Use a symbolic learning system (e.g., DL-Learner) to generate an explanatory theory.
- We're just starting on this, I report on very first experiments.



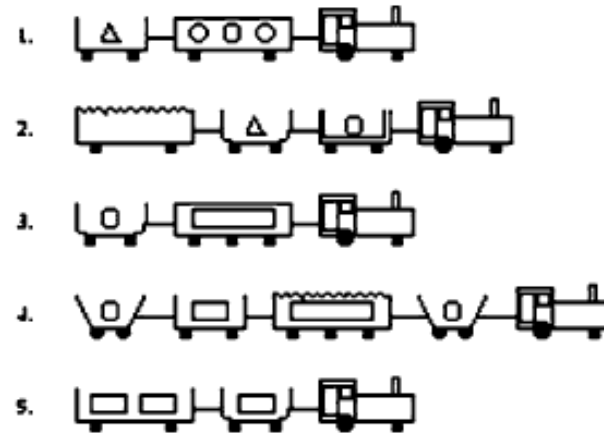


Approach similar to inductive logic programming, but using Description Logics (the logic underlying OWL).

Positive examples:



negative examples:



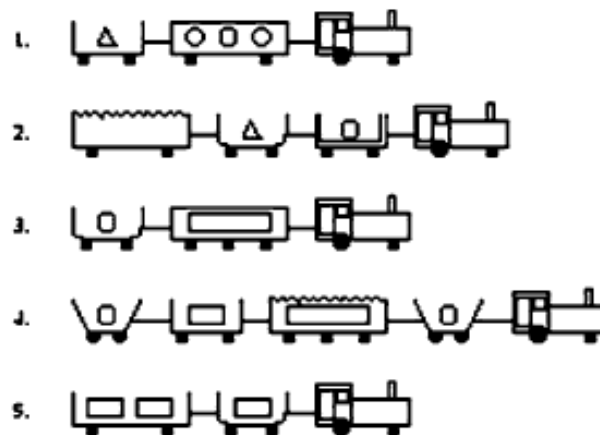
Task: find a class description (logical formula) which separates positive and negative examples.



Positive examples:



negative examples:



DL-Learner result:

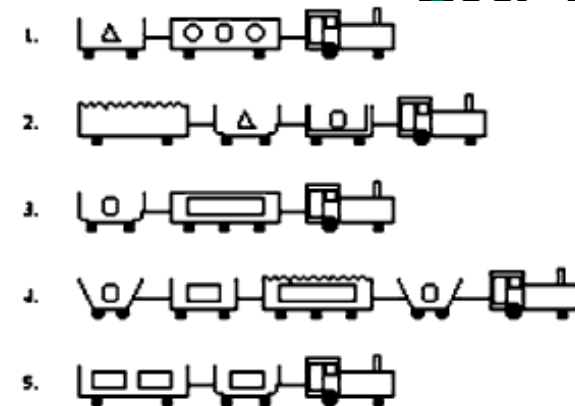
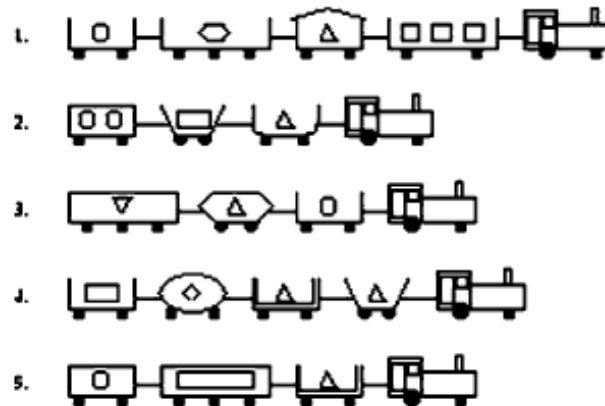
$\exists \text{hasCar} . (\text{Closed} \sqcap \text{Short})$

In FOL:

$$\{x \mid \exists y (\text{hasCar}(x, y) \wedge \text{Closed}(y) \wedge \text{Short}(y))\}$$



DL-Learner uses refinement operators to construct ever better approximations of a solution.



\top

Train – covers all examples.

$\exists \text{hasCar}.\top$

$\exists \text{hasCar}.\text{Closed}$ – covers all positives, two negatives

$\exists \text{hasCar}(\text{Closed} \sqcap \text{Short})$ – solution

Proof of Concept Experiment

Positive:



Negative:





Come from the MIT ADE20k dataset

<http://groups.csail.mit.edu/vision/datasets/ADE20K/>

They come with annotations of objects in the picture:

```
001 # 0 # 0 # sky # sky # ""
002 # 0 # 0 # road, route # road # ""
005 # 0 # 0 # sidewalk, pavement # sidewalk # ""
006 # 0 # 0 # building, edifice # building # ""
007 # 0 # 0 # truck, motortruck # truck # ""
008 # 0 # 0 # hovel, hut, hutch, shack, shanty # hut # ""
009 # 0 # 0 # pallet # pallet # ""
011 # 0 # 0 # box # boxes # ""
001 # 1 # 0 # door # door # ""
002 # 1 # 0 # window # window # ""
009 # 1 # 0 # wheel # wheel # ""
```



Mapping to SUMO

Simple approach: for each known object in image, create an individual for the ontology which is in the appropriate SUMO class:

- contains road1**
- contains window1**
- contains door1**
- contains wheel1**
- contains sidewalk1**
- contains truck1**
- contains box1**
- contains building1**





- Suggested Merged Upper Ontology
<http://www.adampease.org/OP/>
- Approx. 25,000 common terms covering a wide range of domains
- Centrally, a relatively naïve class hierarchy.
- Objects in image annotations became individuals (constants), which were then typed using SUMO classes.



Positive:

img1: road, window, door, wheel, sidewalk, truck, box, building

img2: tree, road, window, timber, building, lumber

img3: hand, sidewalk, clock, steps, door, face, building, window, road

Negative:

img4: shelf, ceiling, floor

img5: box, floor, wall, ceiling, product

img6: ceiling, wall, shelf, floor, product

DL-Learner results include:

\exists contains.Transitway

\exists contains.LandArea

Proof of Concept Experiment

Positive:



Negative:



\exists contains.Transitway

\exists contains.LandArea



- | | | | |
|--|-----|------------------------------------|------|
| \exists contains.Window | (1) | \exists contains.LandTransitway | (6) |
| \exists contains.Transitway | (2) | \exists contains.LandArea | (7) |
| \exists contains.SelfConnectedObject | (3) | \exists contains.Building | (8) |
| \exists contains.Roadway | (4) | \forall contains. \neg Floor | (9) |
| \exists contains.Road | (5) | \forall contains. \neg Ceiling | (10) |

Experiment 2

Positive (selection):



Negative (selection):



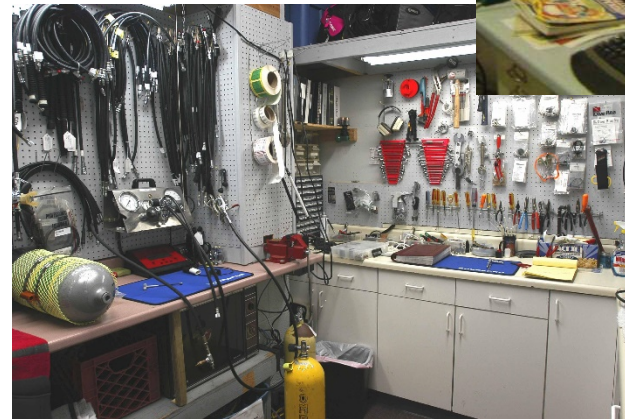
∃contains. (DurableGood \sqcap \neg ForestProduct)

Experiment 3

Positive:



Negative:



$\forall \text{contains.} (\neg \text{Furniture} \sqcap \neg \text{IndustrialSupply})$

Experiment 4

Positive (selection):



Negative (selection):



∄contains.SentientAgent

Experiment 5



Positive:



Negative (selection):



\exists contains.BodyOfWater



- Utilize more sophisticated ontology.
- Utilize more sophisticated mappings.
- Explain hidden neurons.
- Tune DL-Learner better to the specific task.

Collaborators Derek Doran and Ning Xie (Web and Complex Systems Lab)



They explore how to determine groups of hidden neurons which often fire together and thus may indicate the “detection” of certain features.

We plan to apply the above mentioned DL-Learner approach also to these groups of hidden neurons, in order to determine which features they detect.

Thanks!

**Joint work with
Md Kamruzzaman Sarker, Derek Doran, Ning Xie, Mike Raymer**





- Maryam Labaf, Pascal Hitzler, Anthony B. Evans, Propositional Rule Extraction from Neural Networks under Background Knowledge. In: Proceedings of the Twelfth International Workshop on Neural-Symbolic Learning and Reasoning, NeSy'17, London, UK, July 2017.
- Md. Kamruzzaman Sarker, Ning Xie, Derek Doran, Michael Raymer, Pascal Hitzler, Explaining Trained Neural Networks with Semantic Web Technologies: First Steps. In: Proceedings of the Twelfth International Workshop on Neural-Symbolic Learning and Reasoning, NeSy'17, London, UK, July 2017.
- Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph, *Foundations of Semantic Web Technologies*. Chapman and Hall/CRC Press, 2010.
- Krzysztof Janowicz, Frank van Harmelen, James A. Hendler, Pascal Hitzler, Why the Data Train Needs Semantic Rails. *AI Magazine* 26 (1), 2015, 5-14.

- Wouter Beek, Laurens Rietveld, Stefan Schlobach, Frank van Harmelen, LOD Laundromat: Why the Semantic Web Needs Centralization (Even If We Don't Like It). *IEEE Internet Computing* 20(2): 78-81 (2016)
- Denny Vrandečić, Markus Krötzsch. Wikidata: A Free Collaborative Knowledgebase. In *Communications of the ACM* 57:10. ACM 2014.
- Ramanathan V. Guha, Dan Brickley, Steve Macbeth: Schema.org: Evolution of Structured Data on the Web. *ACM Queue* 13(9): 10 (2015)
- Jens Lehmann, Pascal Hitzler, Concept Learning in Description Logics Using Refinement Operators. *Machine Learning* 78 (1-2), 203-250, 2010.
- B. Hammer, P. Hitzler (eds.). Perspectives of Neural-Symbolic Integration. *Studies in Computational Intelligence*, Vol. 77. Springer, 2007, ISBN 978-3-540-73952-1.



- **S. Bader and P. Hitzler, Dimensions of neural-symbolic integration – a structured survey. In: S. Artemov et al. (eds). We Will Show Them: Essays in Honour of Dov Gabbay, Volume 1. College Publications, London, 2005, pp. 167-194.**
- **Artur d'Avila Garcez, Tarek R. Besold, Luc de Raedt, Peter Földiak, Pascal Hitzler, Thomas Icard, Kai-Uwe Kühnberger, Luis C. Lamb, Risto Miikkulainen, Daniel L. Silver, Neural-Symbolic Learning and Reasoning: Contributions and Challenges. In: Andrew McCallum, Evgeniy Gabrilovich, Ramanathan Guha, Kevin Murphy (eds.), Proceedings of the AAI 2015 Spring Symposium on Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches. Technical Report SS-15-03, AAAI Press, Palo Alto, 2015.**