Knowledge Graphs, Deep Learning, and What They Have To Do With Each Other

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

Data Semantics Laboratory

DaSe Lab

SSC

Some Data

- From Germany, dual citizen. PhD in Ireland (in Mathematics)
- Wright State University since 2009.
  Assistant Professor 2009-2012
  Associate Professor 2012-2015
  (Full) Professor since 2015
  Endowed NCR Distinguished Professor since 2016
- Over 400 publications
- Over 9,000 Google Scholar citations

- Previous graduate students and postdocs now at (selection):

TU Dresden, Germany
UG Athens, GA
IIT Delhi, India
UN Lisboa, Portugal
Amazon
Nuance
Universitas Indonesia
Southeast University China
U Bonn, Germany
UN Headquarters New York
GE Global Research
etc.

Teaching Highlights

- **Redesigned discrete math sequence for computer scientists with focus on underprepared students, and increased their retention rate from 6% to 24%.**
- **Most of my classes have an additional distance learning section.**
- **I am teaching most of my classes as flipped classrooms.**
- **I received specific funding from my host institution for my teaching innovations.**

Textbook: Syntax & Semantics

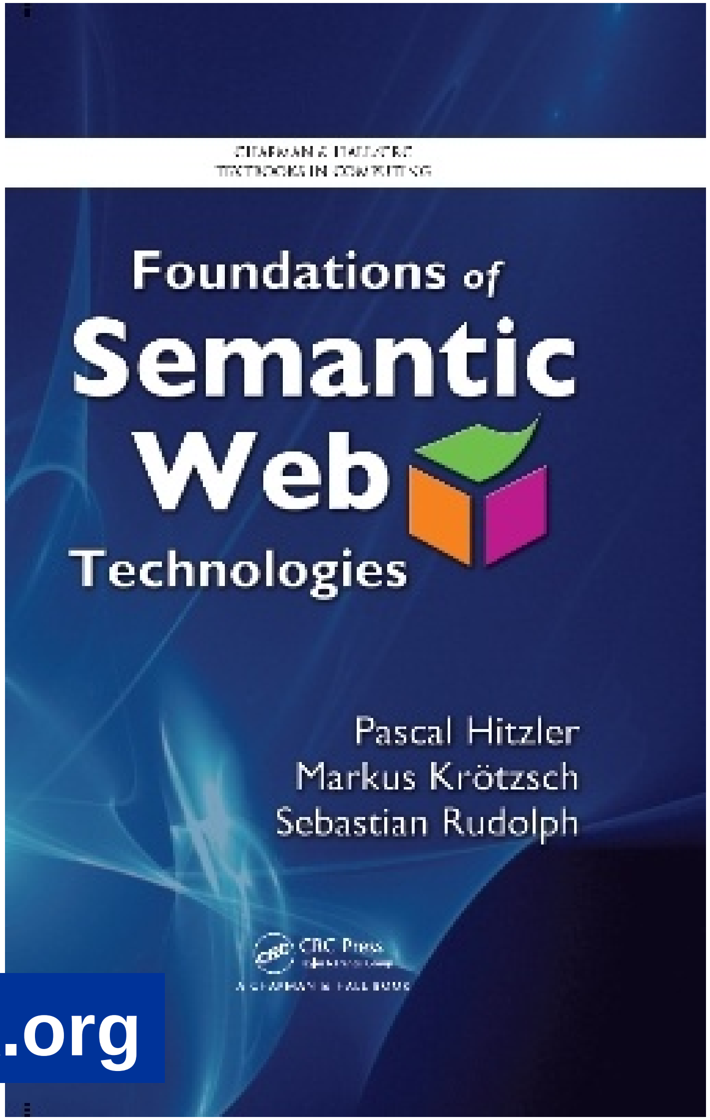
DaSe Lab

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)



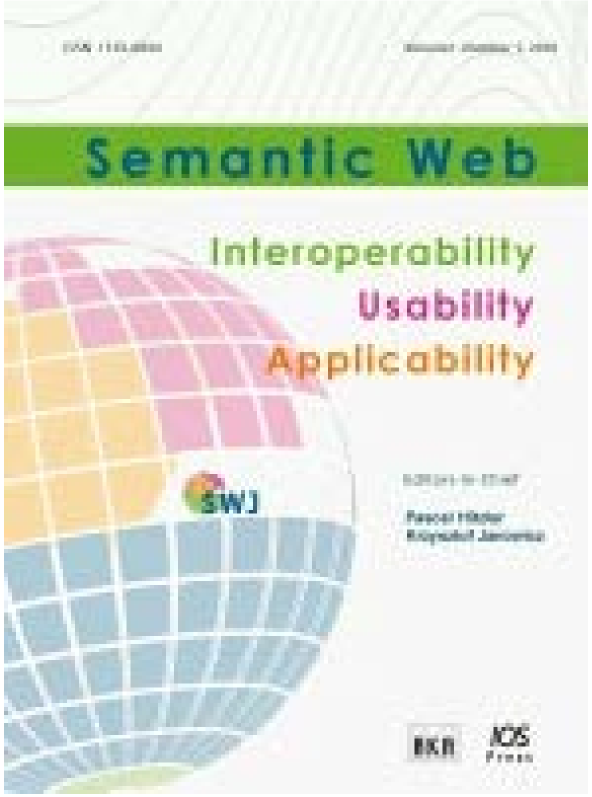
http://www.semantic-web-book.org

Semantic Web journal

DaSe Lab

- **EiCs:** **Pascal Hitzler**
  Krzysztof Janowicz
- **Funded 2010**
- **2018 Impact factor of 2.224, top (with 0.6 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.**



SSC

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

Knowledge Graphs

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+

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

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

J

versity

National Science & Tech Council

DaSe Lab

Open Knowledge Network

https://www.nitrd.gov/nitrdgroups/index.php?title=Open_Knowledge_Network

Report, November 2018:

Conclusion

Artificial intelligence, machine learning, natural language technologies, and robotics are all driving innovation in information systems. Developing the knowledge bases, graphs, and networks that lie at the heart of these systems is expensive and tends to be domain specific, and the largest currently are focused on consumer products (e.g., for web search, advertising placement, and question answering). An open and broad community effort to develop a national-scale data infrastructure—an Open Knowledge Network—would distribute the development expense, be accessible to a broad group of stakeholders, and be domain-agnostic. This infrastructure has the potential to drive innovation across medicine, science, engineering, and finance, and achieve a new round of explosive scientific and economic growth not seen since the adoption of the Internet.

The current hype

- Often, a hype is created because something new has been established.
- In this case, the hype is often over before the technology has really matured to the level of application development.

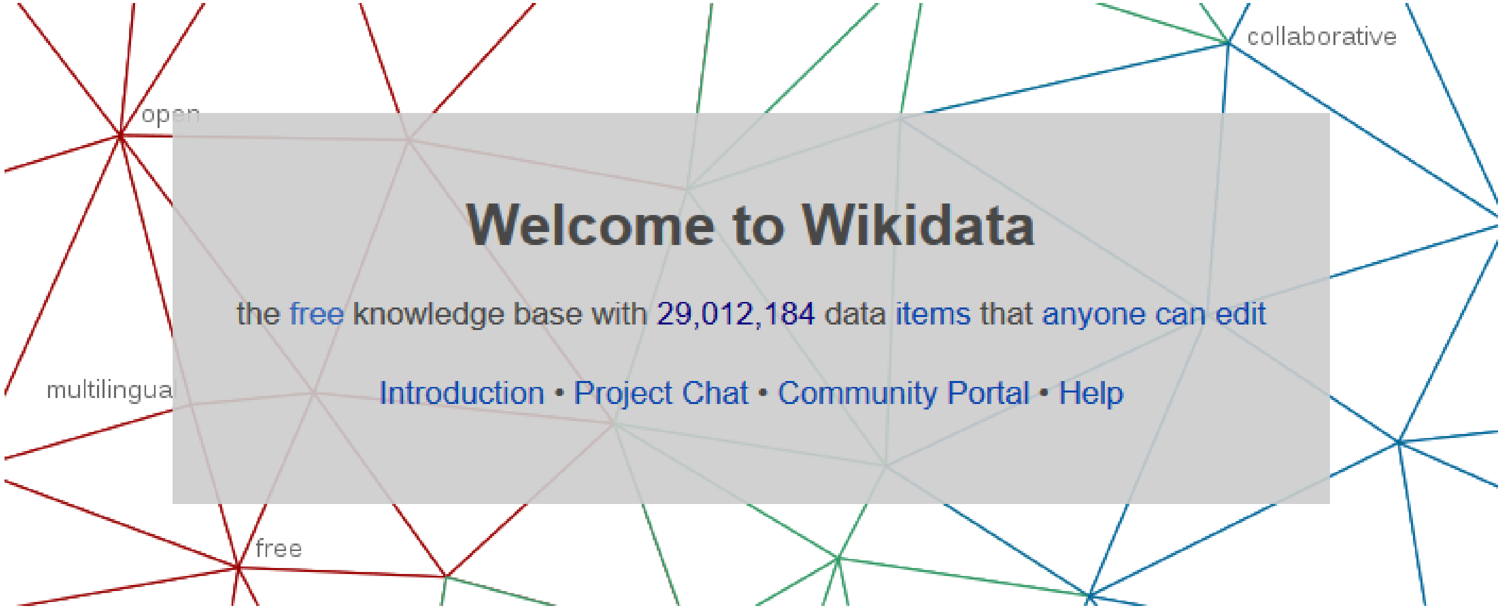
- The current knowledge graph hype is different. Because there was already a pre-maturity hype 15 years ago, under a different name …

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



Welcome to Wikidata

the free knowledge base with 29,012,184 data items that anyone can edit

Introduction • Project Chat • Community Portal • Help

Welcome!

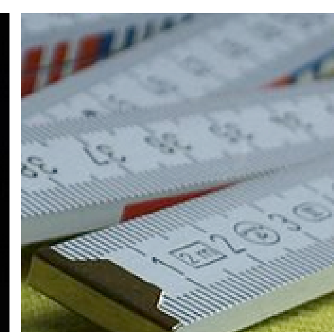
Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

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.



Linked Data

DaSe Lab

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

Number of Datasets in the connected "LOD Cloud"		
	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

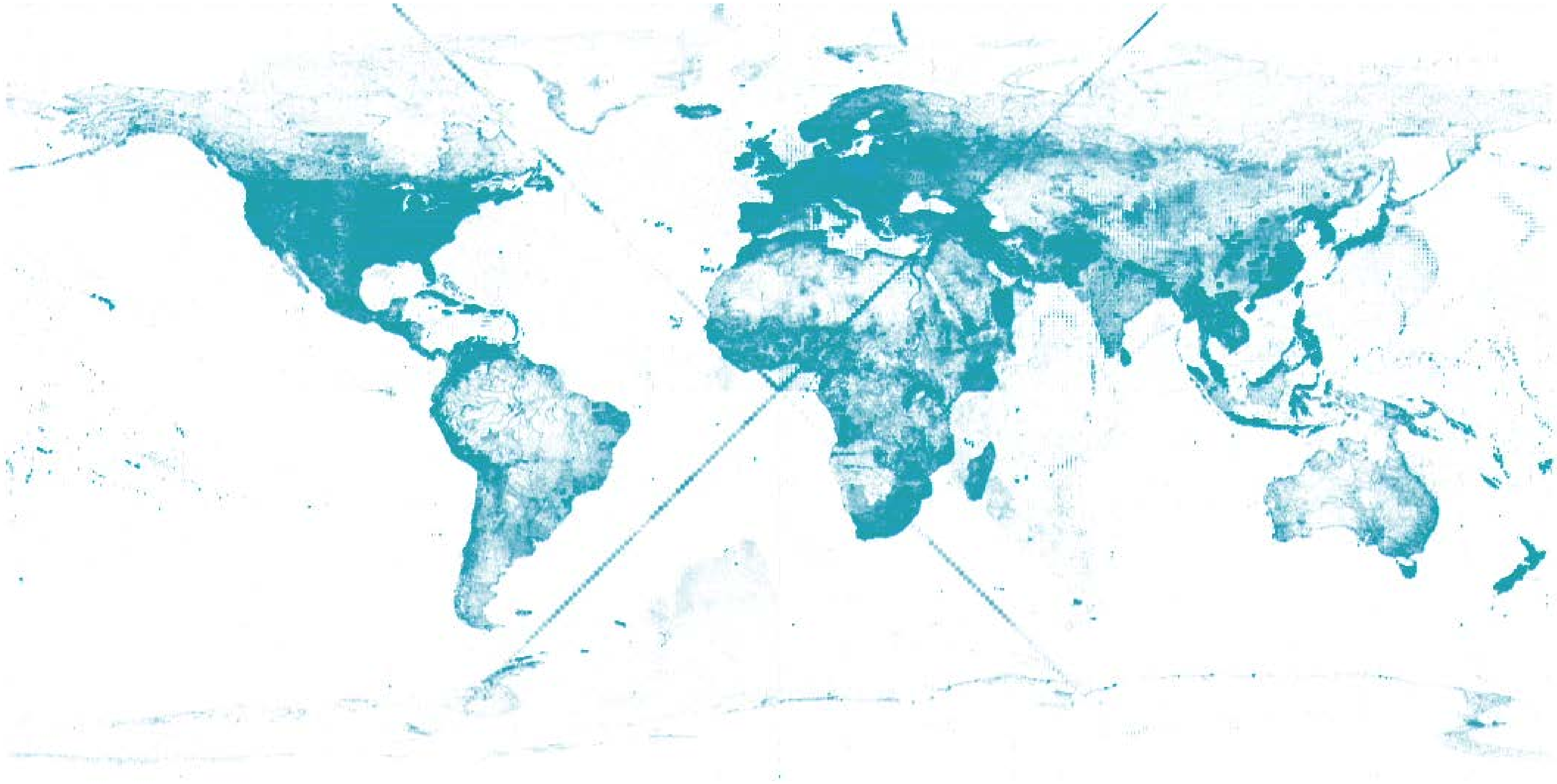
38.606.408.854 triples and counting!

LOD Laundromat

Linked Data: Volume

Geoindexed Linked Data – courtesy of Krzysztof Janowicz, 2012

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



A bit of history

Before the current hype (stimulated by Google), knowledge graphs

- **have been a core artefact of study and deployment since 2007, as part of the maturing "Semantic Web Technologies" field.**
- **have been based on maturing methods and tools around the use of ontologies in the Semantic Web field, since at least 2001.**
- **have even older roots in**
  - **Artificial Intelligence, in particular related to knowledge representation and logical (deductive) reasoning**
  - **the study of terminologies (and ontologies) pre-dating the Semantic Web (and Computer Science) era**

DaSe Lab

What makes a good knowledge graph?

Principles

Goal: Easy sharing, discovery, integration, reuse

Key aspects of knowledge graphs:

- **Syntax**
- **Semantics**
- **Graph structure**
- **Tools**

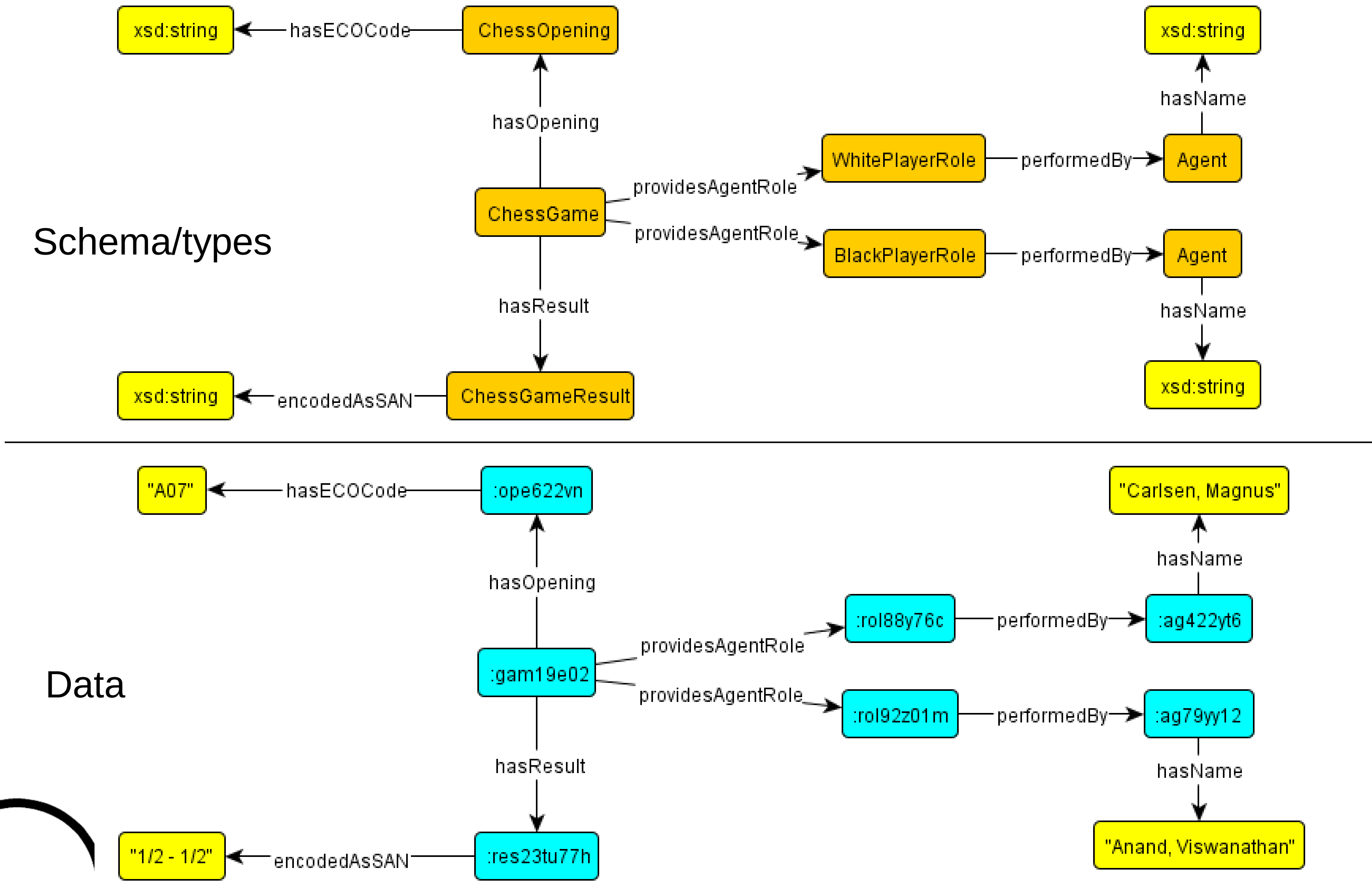
Standards for syntax and semantics have been in place since at least 2004, developed by the World Wide Web Consortium (W3C).

Graph Structure

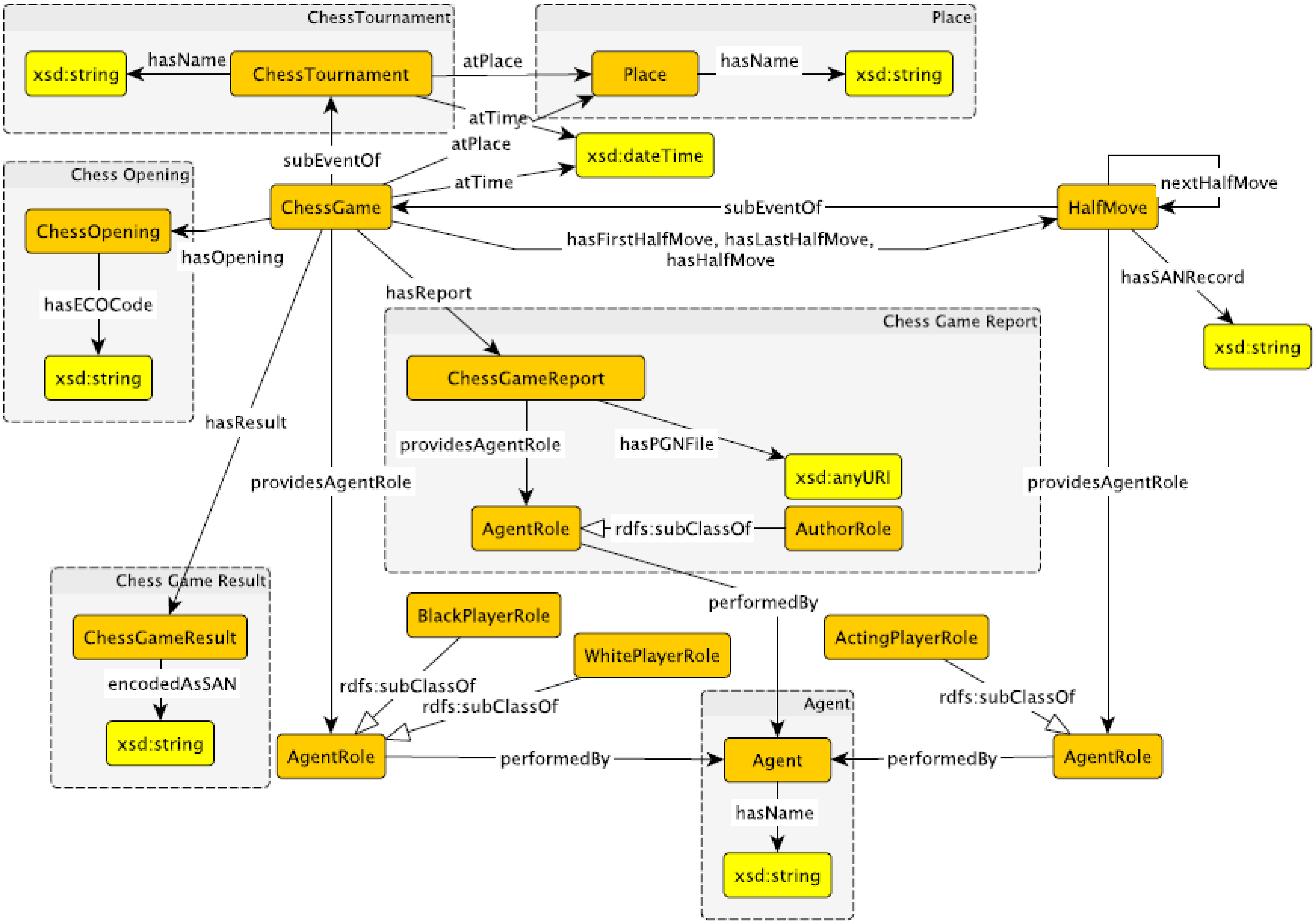
- A schema for a knowledge graph is actually also a knowledge graph, just using more abstract terms, like
  - Classes (or types) of things
    (like, Person, or Material, or Role)
  - Possible relationships between things
    (like, persons may have daughters)
  - Complex relationship assertions
    (like, every cube has 6 sides which are squares).
- A quality schema (or ontology) serves as an intermediary between data/graph structure and human conceptualization.
- A quality schema simplifies understanding and reuseability of the knowledge graph.

Ontology and Knowledge Graph

DaSe Lab



Schema as a knowledge graph



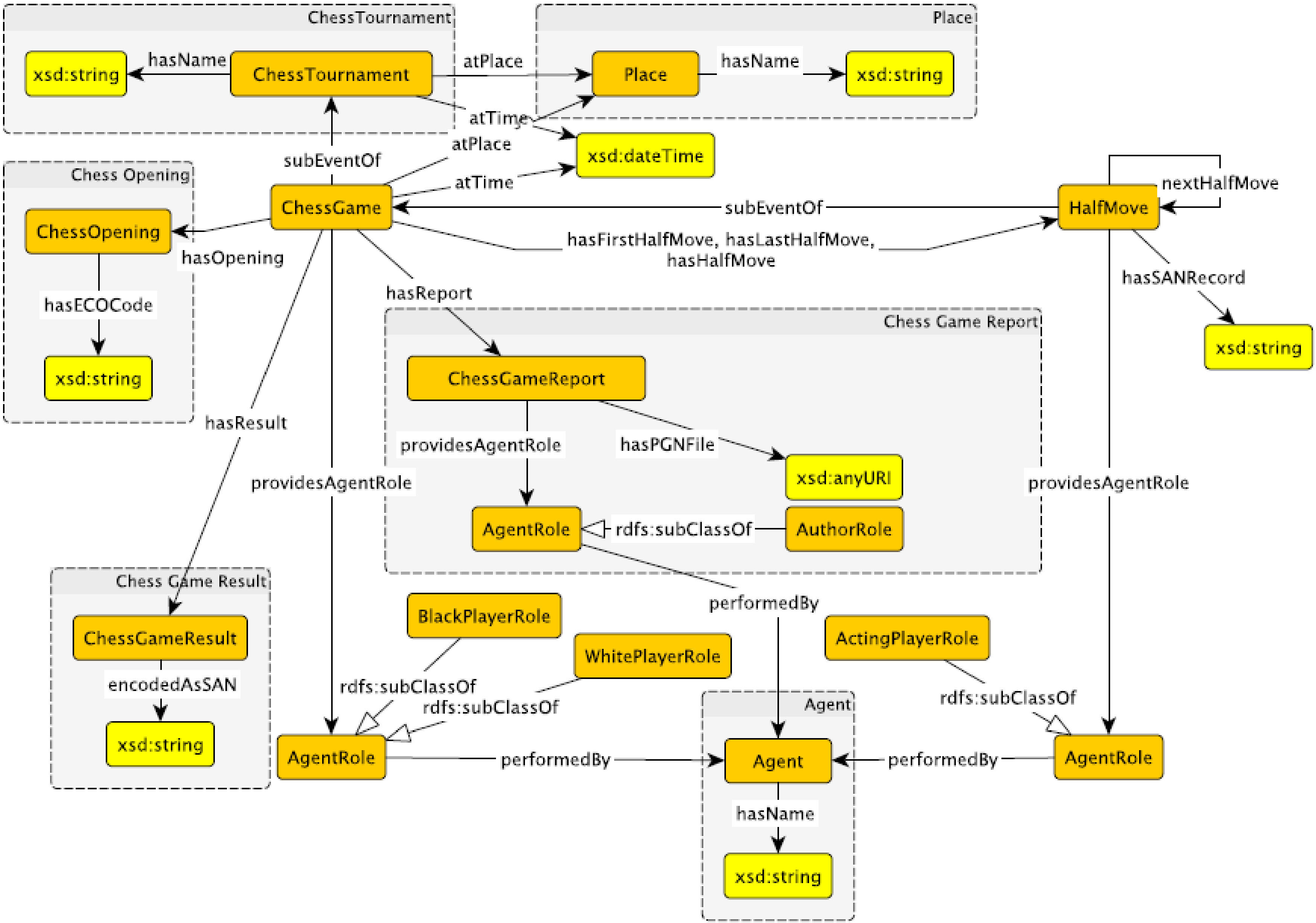
Modular Ontology Architecture

- ***Ontology Design Pattern*: A reusable ontology-piece constituting a high-quality, highly reuseable model for a commonly recurring notion.**
  E.g., “Trajectory”, “Activity”, “Role (of an Agent)”, etc.

- **Use of well-constructed patterns minimizes risk of “naïve” modeling mistakes, thus increases reusability and repurposing of the ontology.**

- **Such ontologies are naturally made up of conceptual “modules” – these make understanding and maintenance of the ontology considerably easier.**

Schema as a knowledge graph



$$\mathsf{AgentRole} \sqsubseteq (=1\ \mathsf{performedBy.Agent}) \sqcap \forall \mathsf{performedBy.Agent} \tag{10.1}$$
$$\exists \mathsf{performedBy.Agent} \sqsubseteq \mathsf{AgentRole} \tag{10.2}$$
$$\top \sqsubseteq \forall \mathsf{pAR.AgentRole} \tag{10.3}$$
$$\mathsf{ChessGame} \sqsubseteq \exists \mathsf{atPlace.Place} \sqcap \forall \mathsf{atPlace.Place} \tag{10.4}$$
$$\mathsf{ChessGame} \sqsubseteq \exists \mathsf{atTime.xsd{:}dateTime} \sqcap \forall \mathsf{atTime.xsd{:}dateTime} \tag{10.5}$$
$$\mathsf{ChessGame} \sqsubseteq \exists \mathsf{pAR.BlackPlayerRole} \sqcap \exists \mathsf{pAR.WhitePlayerRole} \tag{10.6}$$
$$\exists \mathsf{subEventOf.ChessTournament} \sqcup \exists \mathsf{hasOpening.ChessOpening} \sqsubseteq \mathsf{ChessGame} \tag{10.7}$$
$$\exists \mathsf{hasResult.ChessGameResult} \sqcup \exists \mathsf{hasReport.ChessGameReport} \sqsubseteq \mathsf{ChessGame} \tag{10.8}$$
$$\mathsf{ChessGame} \sqsubseteq \forall \mathsf{subEventOf.ChessTournament} \sqcap \forall \mathsf{hasOpening.ChessOpening} \tag{10.9}$$
$$\mathsf{ChessGame} \sqsubseteq \forall \mathsf{hasResult.ChessGameResult} \sqcap \forall \mathsf{hasReport.ChessGameReport} \tag{10.10}$$
$$\mathsf{BlackPlayerRole} \sqcup \mathsf{WhitePlayerRole} \sqsubseteq \mathsf{AgentRole} \sqcap (=1\ \mathsf{pAR}^{-}.\mathsf{ChessGame}) \tag{10.11}$$
$$\mathsf{ChessGame} \sqsubseteq (=1\ \mathsf{hasFirstHalfMove.HalfMove}) \sqcap (=1\ \mathsf{hasLastHalfMove.HalfMove}) \tag{10.12}$$
$$\mathsf{ChessGame} \sqsubseteq (=1\ \mathsf{hasLastHalfMove.HalfMove}) \tag{10.13}$$
$$\mathsf{hasHalfMove} \sqsubseteq \mathsf{subEventOf}^{-} \tag{10.14}$$
$$\mathsf{hasFirstHalfMove} \sqsubseteq \mathsf{hasHalfMove} \tag{10.15}$$
$$\mathsf{hasLastHalfMove} \sqsubseteq \mathsf{hasHalfMove} \tag{10.16}$$
$$\mathsf{HalfMove} \sqsubseteq \mathsf{Event} \sqcap \exists \mathsf{pAR.ActingPlayerRole} \sqcap (=1\ \mathsf{hasHalfMove}^{-}.\mathsf{ChessGame}) \tag{10.17}$$
$$\mathsf{ActingPlayerRole} \sqsubseteq \mathsf{AgentRole} \sqcap (=1\ \mathsf{pAR}^{-}.\mathsf{HalfMove}) \tag{10.18}$$
$$\mathsf{HalfMove} \sqsubseteq (\leq 1\ \mathsf{nextHalfMove.HalfMove}) \sqcap \neg \exists \mathsf{nextHalfMove.Self} \tag{10.19}$$
$$\exists \mathsf{subEventOf.ChessGame} \sqcup \exists \mathsf{nextHalfMove.HalfMove} \sqsubseteq \mathsf{HalfMove} \tag{10.20}$$
$$\exists \mathsf{hasSANRecord.xsd{:}string} \sqsubseteq \mathsf{HalfMove} \tag{10.21}$$
$$\mathsf{HalfMove} \sqsubseteq \forall \mathsf{subEventOf.ChessGame} \sqcap \forall \mathsf{nextHalfMove.HalfMove} \tag{10.22}$$

Research Direction

- High-Quality Ontology Engineering process well understood by some experts.
- But this is “soft” knowledge. Some missing pieces:
  - Systematic exploration and evaluation of the methodology
  - Providing a powerful tool landscape supporting the methodology – plus evaluations of their effectiveness.
  - Writing it up in tutorials and textbooks, and disseminate.

- Our methods development was supported primarily through two NSF GEO projects. Currently, it is supported through a $1.8M AFOSR project on cognitive agents.

- Goal: Practical methods and tools for high-quality knowledge graph schema development.

Promise

- Data Management (DM) is central for cost-effective / efficient data-intensive solutions, for many application areas and scenarios.
- DM easily takes 80% of the time when data analytics is done.

- Knowledge Graphs are quickly becoming a central DM tool in industry and academia.
- Our methods target lowering the cost of Data Management with Knowledge Graphs.

- I can contribute to large methods- or application-oriented projects which have Data Management components.
- There is also high potential for a company spin-off.

Studies on the Semantic Web

Ontology Engineering with Ontology Design Patterns

Foundations and Applications

Pascal Hitzler, Aldo Gangemi,
Krzysztof Janowicz, Adila Krisnadhi,
Valentina Presutti (Eds.)

IOS Press

published 2016

AKA

Studies on the Semantic Web

Karl Hammar, Pascal Hitzler, Adila Krisnadhi,
Agnieszka Ławrynowicz, Andrea Giovanni Nuzzolese,
Monika Solanki (Editors)

Advances in Ontology Design and Patterns

IOS Press

published 2017

AKA

Other Aspects

Other Aspects of Knowledge Graph management we are (or have recently been) investigating:

- **Data/schema merging and integration**
- **Formal logic as schema representation language**
- **Deductive (logical) reasoning as KG engineering tool**
- **Efficient algorithms for deductive reasoning, including cloud-based**
- **KG compression**
- **Other aspects of KG quality**
- **Benchmark generation for different KG tools**

Target funding agencies

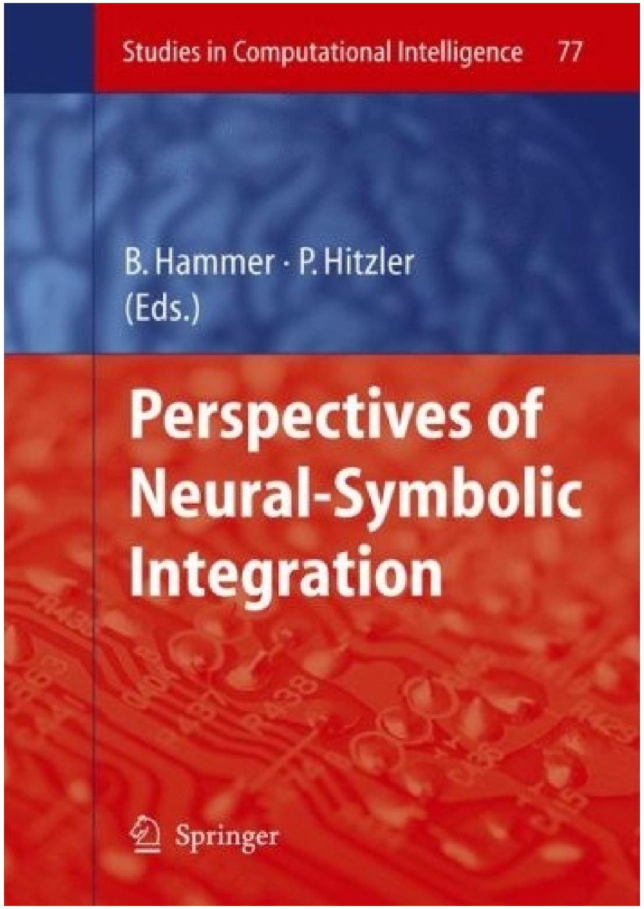
- NSF for core new methods projects
- Intelligence/Defense for application-oriented projects with data management component, where effort can be used to improve and evaluate existing methods and tools. E.g. DARPA "Knowledge-directed Artificial Intelligence Reasoning Over Schemas (KAIROS)" – proposer's day is today (Jan 9, 2018).
- NIH similar, but haven't tapped into this yet.
- Potential sources also on application domains such as smart cities, data privacy and security, library science, human performance improvements, etc. E.g. we're part of a $1.8M Mellon Foundation project on the history of the slave trade.
- I keep watching the Open Knowledge Network initiative.

Deep Learning
and Knowledge Graphs

– selected efforts in Neural-Symbolic Integration

Some Background

Workshop Series on Neural-Symbolic Learning and Reasoning Since 2005.
http://neural-symbolic.org/



Perspectives on Neural-Symbolic Integration

Barbara Hammer and Pascal Hitzler (eds)

Springer, 2007

Neural-Symbolic Learning and Reasoning: A Survey and Interpretation

Tarek R. Besold, Artur d'Avila Garcez, Sebastian Bader, Howard Bowman, Pedro Domingos, Pascal Hitzler, Kai-Uwe Kuehnberger, Luis C. Lamb, Daniel Lowd, Priscila Machado Vieira Lima, Leo de Penning, Gadi Pinkas, Hoifung Poon, Gerson Zaverucha

https://arxiv.org/abs/1711.03902 **(2017)**

Neural-Symbolic

Computer Science perspective:

- **Connectionist machine learning systems are**
  - **very powerful for some machine learning problems**
  - **robust to data noise**
  - **very hard to understand or explain**
  - **really poor at symbol manipulation**
  - **unclear how to effectively use background (domain) knowledge**
- **Symbolic systems are**
  - **Usually rather poor regarding machine learning problems**
  - **Intolerant to data noise**
  - **Relatively easy to analyse and understand**
  - **Really good at symbol manipulation**
  - **Designed to work with other (background) knowledge**

Neural-Symbolic

Computer Science perspective:

- **Let's try to get the best of both worlds:**
  - **very powerful machine learning paradigm**
  - **robust to data noise**
  - **easy to understand and assess by humans**
  - **good at symbol manipulation**
  - **work seamlessly with background (domain) knowledge**

- **How to do that?**
  - **Endow connectionist systems with symbolic components?**
  - **Add connectionist learning to symbolic reasoners?**

The Interface Issue

- Symbolic knowledge comes as logical theories (sets of formulas over a logic)
- Subsymbolic systems process tuples of real/float numbers (vectors, matrices, tensors)

- How do you interface?
- How do you map between the symbolic world and the subsymbolic world?

Some key problems that need to be overcome:

- Logic is full of highly structured objects, how to represent them in Real Space?
- How to represent variable bindings in a distributed setting?
- The required length of logical deduction chain is not known up front.

RDFS Deductive Reasoning via Deep Memory Networks

RDF reasoning

- **[Note: RDF is one of the simplest useful knowledge representation languages beyond propositional logic.]**

- **Think knowledge graph.**
- **Think node-edge-node triples such as**

BarackObama	**rdf:type**	**President**
BarackObama	**husbandOf**	**MichelleObama**
President	**rdfs:subClassOf**	**Human**
husbandOf	**rdfs:subPropertyOf**	**spouseOf**

- **Then there is a (fixed, small) set of inference rules, such as**

  rdf:type(x,y) AND rdfs:subClassOf(y,z)THEN rdf:type(x,z)

Representation

- **Goal is to be able to reason over unseen knowledge graphs. I.e. the out-of-vocabulary problem needs addressing.**

- **Normalization of vocabulary (i.e., it becomes shared vocabulary across all input knowledge graphs.**
- **One vocabulary item becomes a one-hot vector (dimension d, number of normalized vocabulary terms)**
- **One triple becomes a 3 x d matrix.**
- **The knowledge graph becomes an n x 3 x d tensor. (n is the number of knowledge graph triples)**

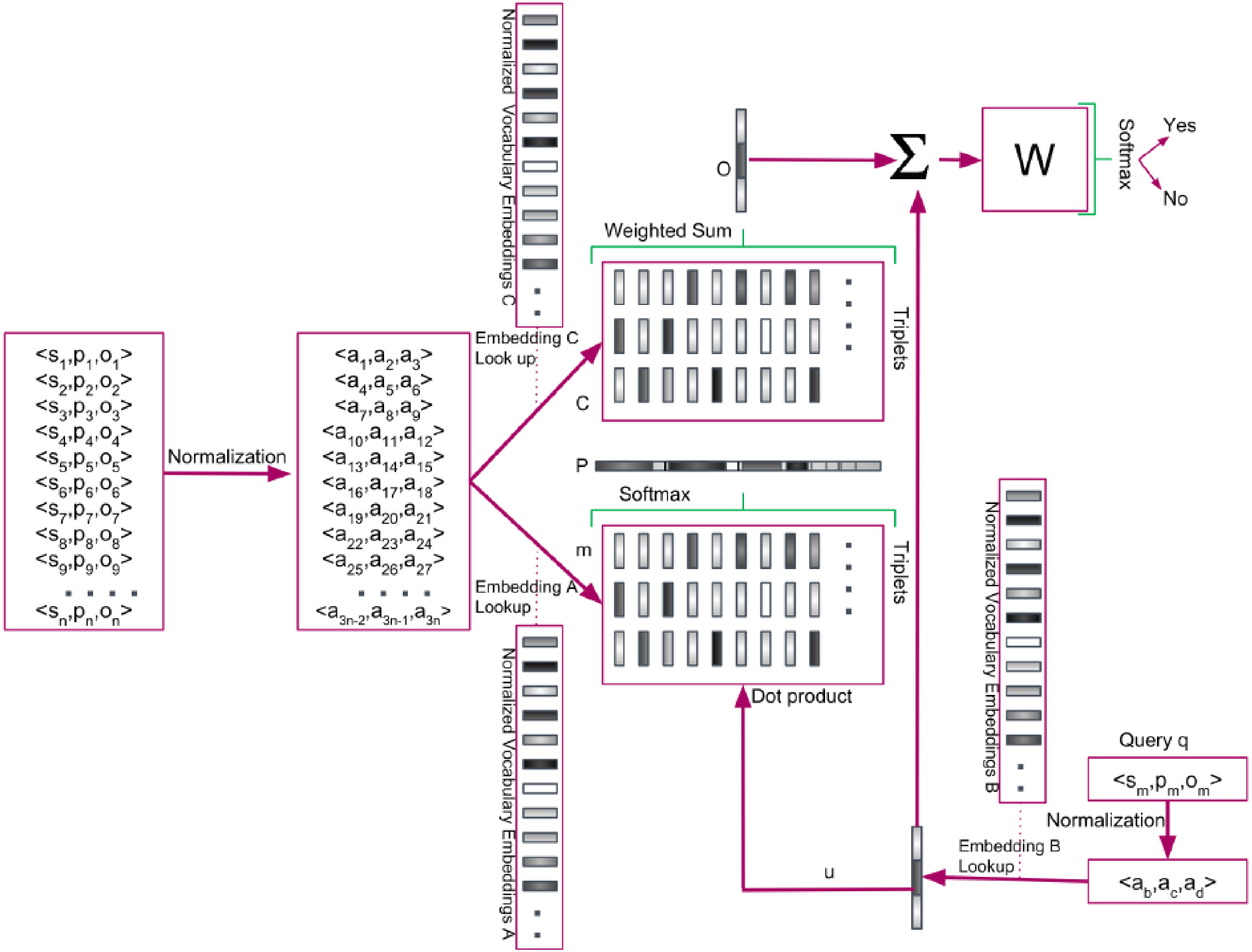
- **Knowledge graph is stored in “memory.”**

Mechanics

- An attention mechanism retrieves momory slots useful for finding the correct answer to a query.
- These are combined with the query and run through a (learned) matrix to retrieve a new (processed) query.

- This is repeated (in our experiment with 10 “hops”).
- The final out put is a yes/no answer to the query.

Memory Network based on MemN2N



Experiments: Performance

DaSe Lab

Test Dataset	#KG	Base						Inferred						Invalid
		#Facts	#Ent.	%Class	%Indv	%R.	%Axiom.	#Facts	#Ent.	%Class	%Indv	%R.	%Axiom.	#Facts
OWL-Centric	2464	996	832	14	19	3	0	494	832	14	0.01	1	20	462
Linked Data	20527	999	787	3	22	5	0	124	787	3	0.006	1	85	124
OWL-Centric Test Set	21	622	400	36	41	3	0	837	400	36	3	1	12	476
Synthetic Data	2	752	506	52	0	1	0	126356	506	52	0	1	0.07	700

Table 2: Statistics of various datasets used in experiments

Baseline: non-normalized embeddings, same architecture

Training Dataset	Test Dataset	Valid Triples Class			Invalid Triples Class			Accuracy
		Precision	Recall /Sensitivity	F-measure	Precision	Recall /Specificity	F-measure	
OWL-Centric Dataset	Linked Data	93	98	96	98	93	95	**96**
OWL-Centric Dataset (90%)	OWL-Centric Dataset (10%)	88	91	89	90	88	89	**90**
OWL-Centric Dataset	OWL-Centric Test Set [b]	79	62	68	70	84	76	**69**
OWL-Centric Dataset	Synthetic Data	65	49	40	52	54	42	**52**
OWL-Centric Dataset	Linked Data [a]	54	98	70	91	16	27	86
OWL-Centric Dataset [a]	Linked Data [a]	62	72	67	67	56	61	91
OWL-Centric Dataset(90%) [a]	OWL-Centric Dataset(10%) [a]	79	72	75	74	81	77	80
OWL-Centric Dataset	OWL-Centric Test Set [ab]	58	68	62	62	50	54	58
OWL-Centric Dataset [a]	OWL-Centric Test Set [ab]	77	57	65	66	82	73	73
OWL-Centric Dataset	Synthetic Data [a]	70	51	40	47	52	38	51
OWL-Centric Dataset [a]	Synthetic Data [a]	67	23	25	52	80	62	50
Baseline								
OWL-Centric Dataset	Linked Data	73	98	83	94	46	61	43
OWL-Centric Dataset (90%)	OWL-Centric Dataset (10%)	84	83	84	84	84	84	82
OWL-Centric Dataset	OWL-Centric Test Set [b]	62	84	70	80	40	48	61
OWL-Centric Dataset	Synthetic Data	35	41	32	48	55	45	48

[a] More Tricky Nos & Balanced Dataset
[b] Completely Different Domain.

Table 3: Experimental results of proposed model

Experiments: Reasoning Depth

Test Dataset	Hop 0			Hop 1			Hop 2			Hop 3			Hop 4			Hop 5			Hop 6			Hop 7			Hop 8			Hop 9			Hop 10		
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
Linked Data[a]	0	0	0	80	99	88	89	97	93	77	98	86	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Linked Data[b]	2	0	0	82	91	86	89	98	93	79	100	88	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
OWL-Centric[c]	19	5	9	31	75	42	78	80	78	48	47	44	4	34	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Synthetic	32	46	33	31	87	38	66	55	44	25	45	32	29	46	33	26	46	33	25	46	33	25	46	33	24	43	31	25	43	31	22	36	28

[a] LemonUby Ontology
[b] Agrovoc Ontology
[c] Completely Different Domain

Table 4: Experimental results over each reasoning hop

Dataset	Hop 1	Hop 2	Hop 3	Hop 4	Hop 5	Hop 6	Hop 7	Hop 8	Hop 9	Hop 10
OWL-Centric [a]	8%	67%	24%	0.01%	0%	0%	0%	0%	0%	0%
Linked Data[b]	31%	50%	19%	0%	0%	0%	0%	0%	0%	0%
Linked Data[c]	34%	46%	20%	0%	0%	0%	0%	0%	0%	0%
OWL-Centric[d]	5%	64%	30%	1%	0%	0%	0%	0%	0%	0%
Synthetic Data	0.03%	1.42%	1%	1.56%	3.09%	6.03%	11.46%	20.48%	31.25%	23.65%

[a] *Training Set*
[b] LemonUby Ontology
[c] Agrovoc Ontology
[d] Completely Different Domain

Table 5: Data distribution per knowledge graph over each reasoning hop

Training time: just over a full day

Next steps

- More complex logics (we're optimistic that methods carry over at least to some proper description logics).
- Applications to commonsense (natural language) reasoning.
- Investigating reasoning robustness and efficiency.
- We have confirmed industry interest already.

Explaining Deep Learning via Symbolic Background Knowledge

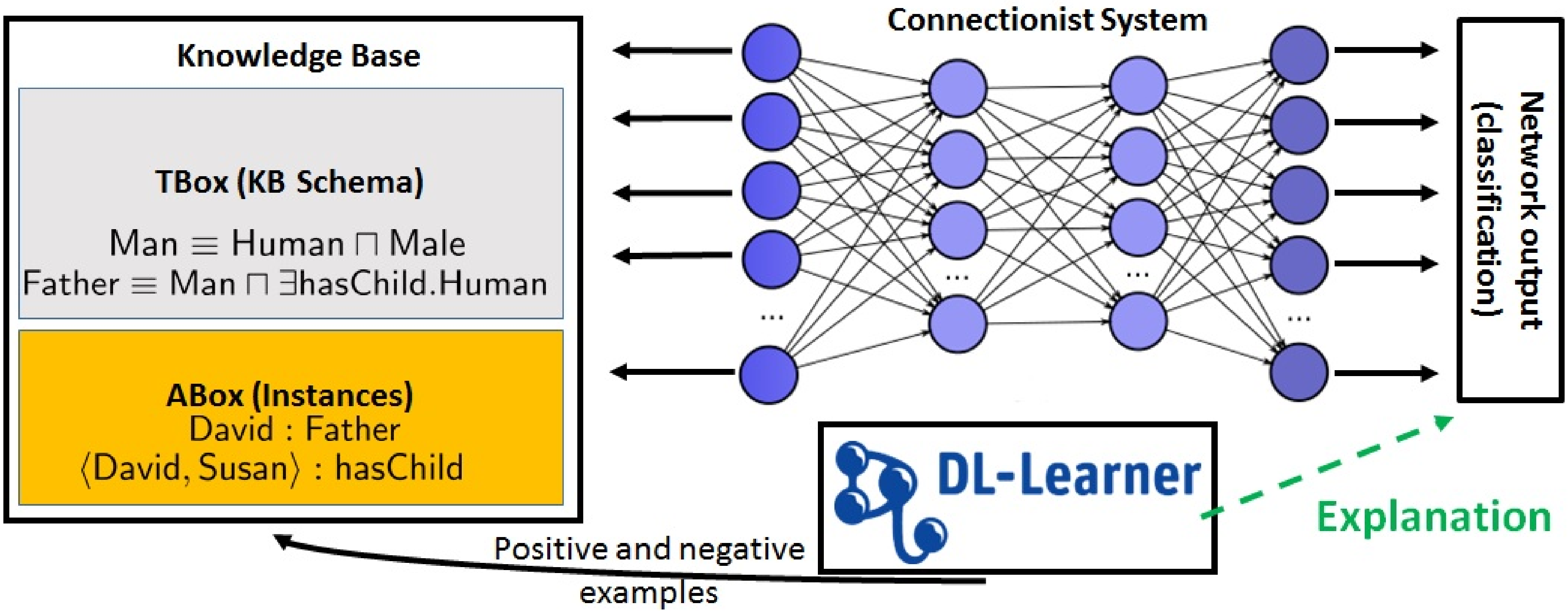
Explainable AI

- Explain 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 first experiments.

Explainable AI



Using SUMO

Testing on ADE20k image dataset / scene recognition.

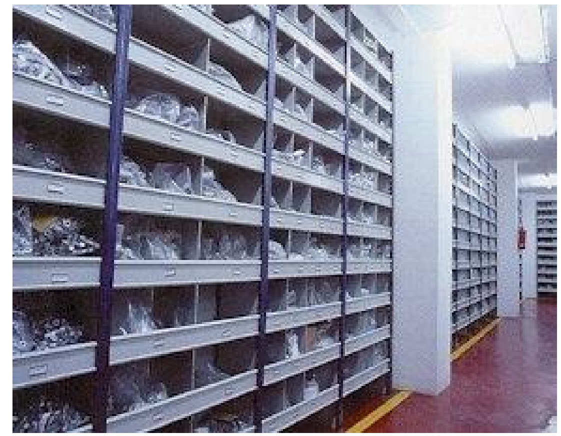
Workshop paper at NeSy'2017 with preliminary results.

Proof of Concept Experiment

DaSe Lab

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

SUMO

- **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.**

DL-Learner input

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:

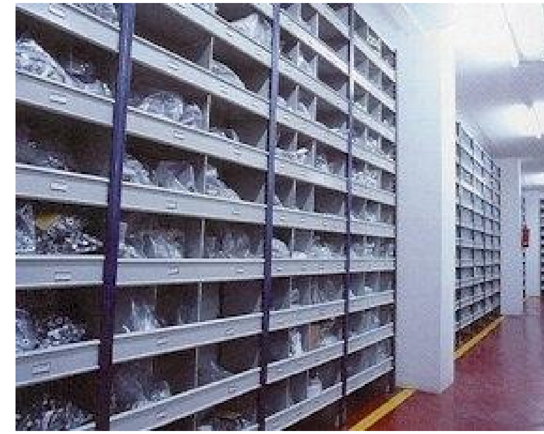
$$\exists \mathrm{contains}.\mathrm{Transitway}$$

$$\exists \mathrm{contains}.\mathrm{LandArea}$$

Proof of Concept Experiment

Positive:

Negative:



$$\exists \mathrm{contains}.\mathrm{Transitway}$$

$$\exists \mathrm{contains}.\mathrm{LandArea}$$

Experiment 2

Positive (selection):

Negative (selection):

$$\exists \mathrm{contains}.(\mathrm{DurableGood} \sqcap \neg \mathrm{ForestProduct})$$

Experiment 4

Positive (selection):

Negative (selection):

$\exists \text{contains.SentientAgent}$

Experiment 5

Positive:




Negative (selection):



$\exists \mathrm{contains.BodyOfWater}$

DL-Learner efficiency problem

DaSe Lab

- **DL-Learner was too slow – we needed several hours for each computation, and couldn't explore and/or scale up.**
- **We thus implemented our own system, ECII (Efficient Concept Induction from Instances) which trades some correctness for speed. [Sarker, Hitzler, AAAI-19, to appear]**

Experiment Name	Number of Logical Axioms	Runtime (sec)					Accuracy (α_3)		Accuracy α_2			
		DL[a]	DL FIC(1)[b]	DL FIC(2)[c]	ECII DF[d]	ECII KCT[e]	DL[a]	ECII DF[d]	DL FIC(1)[b]	DL FIC(2)[c]	ECII DF[d]	ECII KCT[e]
Yinyang_examples	157	0.065	0.0131	0.019	0.089	0.143	1.000	0.610	1.000	1.000	0.799	1.000
Trains	273	0.01	0.020	0.047	0.05	0.095	1.000	1.000	1.000	1.000	1.000	1.000
Forte	341	2.5	1.169	6.145	0.95	0.331	0.965	0.642	0.875	0.875	0.733	1.000
Poker	1,368	0.066	0.714	0.817	1	0.281	1.000	1.000	0.981	0.984	1.000	1.000
Moral Reasoner	4,666	0.1	3.106	4.154	5.47	6.873	1.000	0.785	1.000	1.000	1.000	1.000
ADE20k I	4,714	577.3[f]	4.268	31.887	1.966	23.775	0.926	0.416	0.263	0.814	0.744	1.000
ADE20k II	7,300	983.4[f]	16.187	307.65	20.8	293.44	1.000	0.673	0.413	0.413	0.846	0.900
ADE20k III	12,193	4,500[g]	13.202	263.217	51	238.8	0.375	0.937	0.375	0.375	0.930	0.937
ADE20k IV	47,468	4,500[g]	93.658	523.673	116	423.349	0.375	NA	0.608	0.608	0.660	0.608

[a] DL : DL-Learner
[b] DL FIC (1) : DL-Learner fast instance check with runtime capped at execution time of ECII DF
[c] DL FIC (2) : DL-Learner fast instance check with runtime capped at execution time of ECII KCT
[d] ECII DF : ECII default parameters
[e] ECII KCT : ECII keep common types and other default parameters
[f] Runtimes for DL-Learner were capped at 600 seconds.
[g] Runtimes for DL-Learner were capped at 4,500 seconds.

SSC

ECII vs. DL-Learner

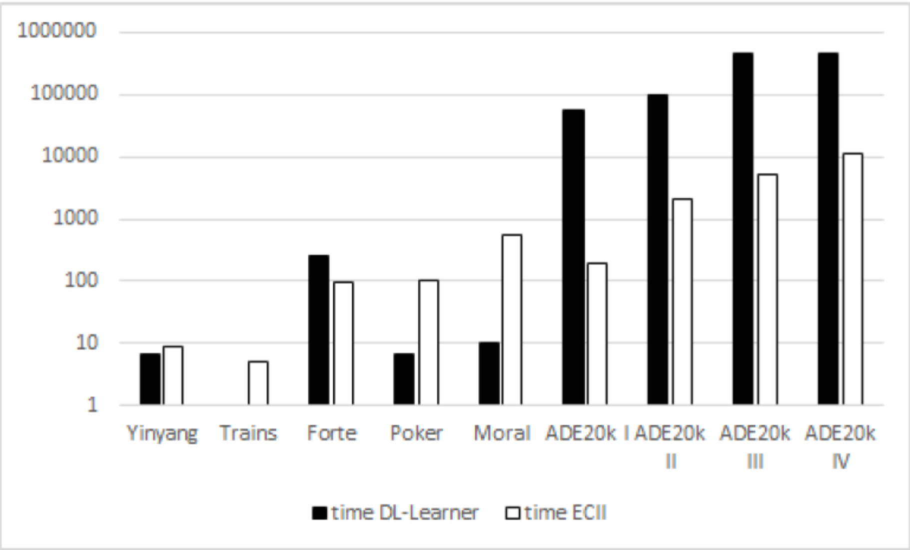


Figure 1: Runtime comparison between DL-Learner and ECII. The vertical scale is logarithmic in hundredths of seconds, and note that DL-Learner runtime has been capped at 4,500 seconds for ADE20k III and IV. For ADE20k I it was capped at each run at 600 seconds.

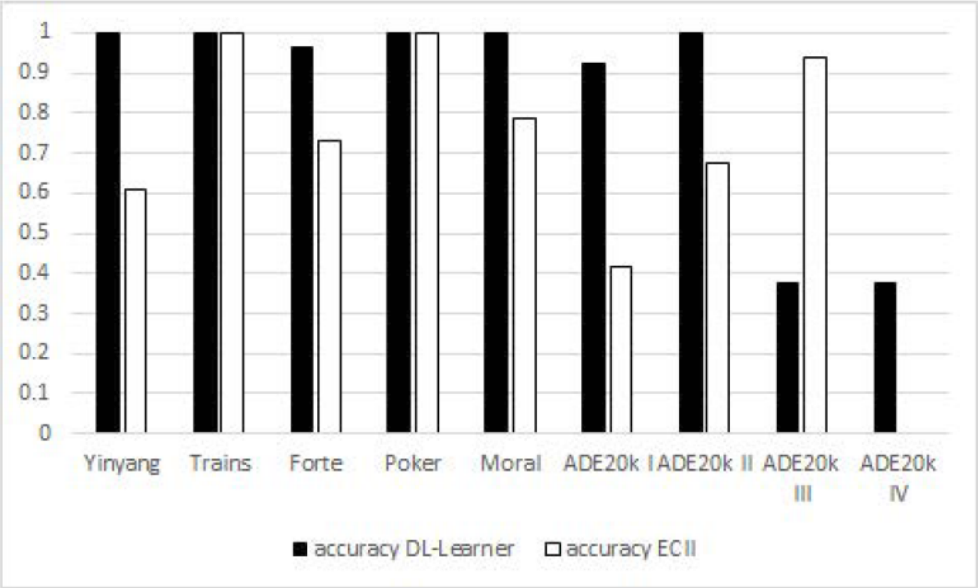


Figure 2: Accuracy (α_3) comparison between DL-Learner and ECII. For ADE20k IV it was not possible to compute an accuracy score within 3 hours for ECII as the input ontology was too large.

Next:

- We’re just now starting to run full-scale experiments with ECII in the described setting.
(The main PhD student on this topic just departed on an internship to Intel.)

Uniqueness of approach

- **To the best of our knowledge, nobody else is pursuing explainable deep learning though background knowledge.**
- **To the best of our knowledge, nobody funded under the DARPA XAI program is pursuing explainable deep learning through background knowledge.**

Other Aspects

Other Aspects of Deep Learning we are investigating include:

- Deep Learning methods for data integration
- Deep Learning for text analysis
- Deep Learning algorithms to support KG engineering tools (e.g., graph completion)

- Explaining other statistical approaches (not only deep learning) by transferring our methods.

Target funding agencies

- **NSF for core new methods projects**

- **Intelligence/Defense for application-oriented projects regarding the use of explainable deep learning.**

- **NIH similar, but haven't tapped into this yet. We just started a collaboration with IBM TJ Watson on applying our method to drug-drug-interaction and are in talks with Bosch to receive direct industry funding.**

- **Potential sources also on any current or emerging application domain of deep learning, including security, social media analysis, intelligence analysis, etc.**

Summary

- **My work has many facets.**
- **My work is in synch with several current trends, including**
  - **Knowledge Graphs**
  - **Deep Learning**
  - **Big Data**
  - **Data Science**
- **Covering methods/foundations and applications; and the transfer between them.**

- **Broad options for obtaining research funding.**
- **Easier because of already significant visibility and standing. Since I became a US citizen (summer 2017) I also made significant inroads for defense funding, in particular establishing a network of contacts.**

Thanks!