OceanLink: Using Patterns for Discovery in EarthCube

Pascal Hitzler
DaSe Lab for Data Semantics
Wright State University
http://www.pascal-hitzler.de

OceanLink Collaborators

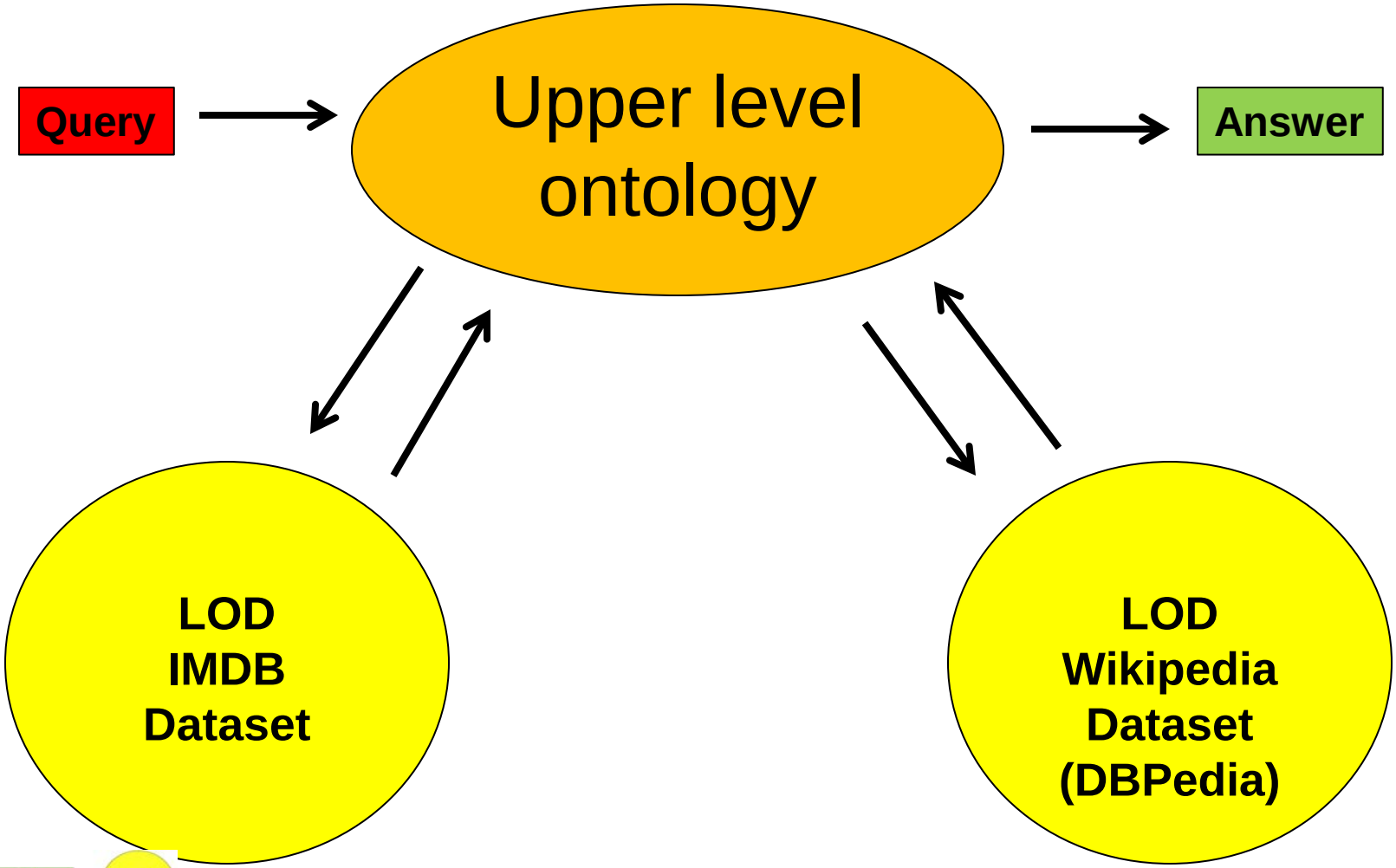
DaSe Lab

Robert Arko, Columbia University
Suzanne Carbotte, Columbia University
Cynthia Chandler, Woods Hole Oceanographic Institution
Michelle Cheatham, Wright State University
Timothy Finin, University of Maryland, Baltimore County
Pascal Hitzler, Wright State University
Krzysztof Janowicz, University of California, Santa Barbara
Adila Krisnadhi, Wright State University
Thomas Narock, Marymount University
Lisa Raymond, Woods Hole Oceanographic Institution
Adam Shepherd, Woods Hole Oceanographic Institution
Peter Wiebe, Woods Hole Oceanographic Institution

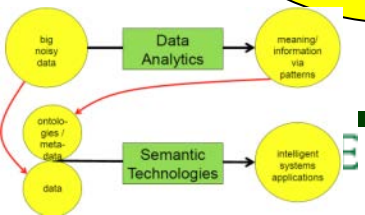
The presented work is part of the NSF *OceanLink* project: EarthCube Building Blocks, Leveraging Semantics and Linked Data for Geoscience Data Sharing and Discovery

WRIGHT STATE UNIVERSITY

Classical ontology-based integration



[ODBASE 2012, JWS 2007]



Example querying LoD

DaSe Lab

"Identify congress members, who have voted "No" on pro environmental legislation in the past four years, with high-pollution industry in their congressional districts."

In principle, all the knowledge is there:

- **GovTrack**
- **GeoNames**
- **DBPedia**
- **US Census**

But even with LoD we cannot answer this query.

Example querying LoD

DaSe Lab

"Identify congress members, who have voted "No" on pro environmental legislation in the past four years, with high-pollution industry in their congressional districts."

Some missing puzzle pieces:

- Where is the data?
  - GovTrack
    GeoNames
    US Census
    requires intimate knowledge of the LoD data sets

Example querying LoD

DaSe Lab

"Identify congress members, who have voted "No" on pro environmental legislation in the past four years, with high-pollution industry in their congressional districts."

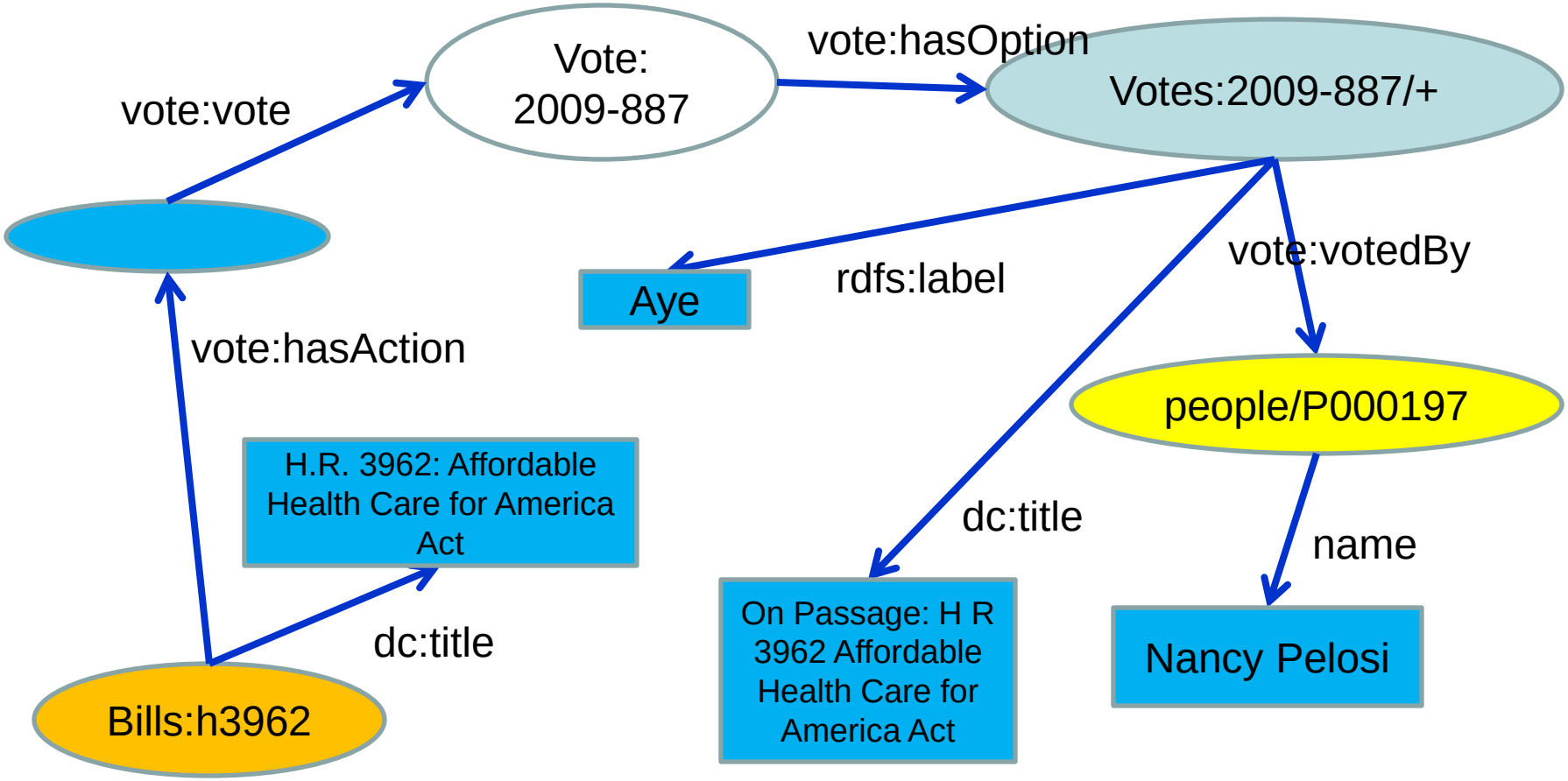
Some missing puzzle pieces:

- Where is the data?
  (smart federation needed)
- Missing background (schema) knowledge.
  (enhancements of the LoD cloud)
- Crucial info still hidden in texts.
  (ontology learning from texts)
- Added reasoning capabilities (e.g., spatial).
  (new ontology language features)

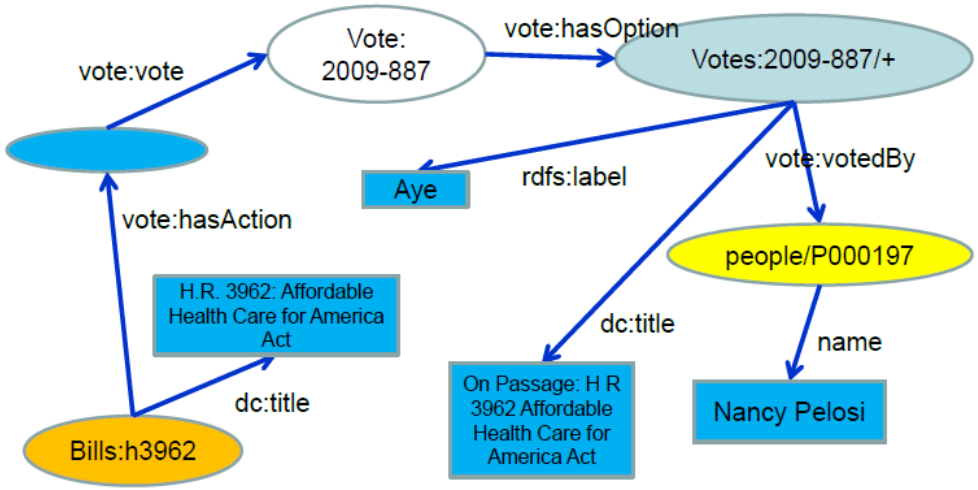
Linked Data: Variety

DaSe Lab

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

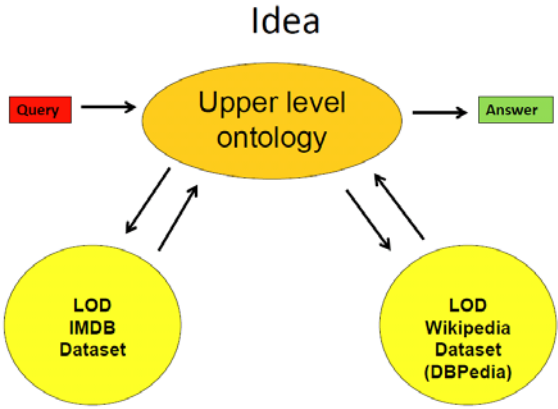


Querying approach

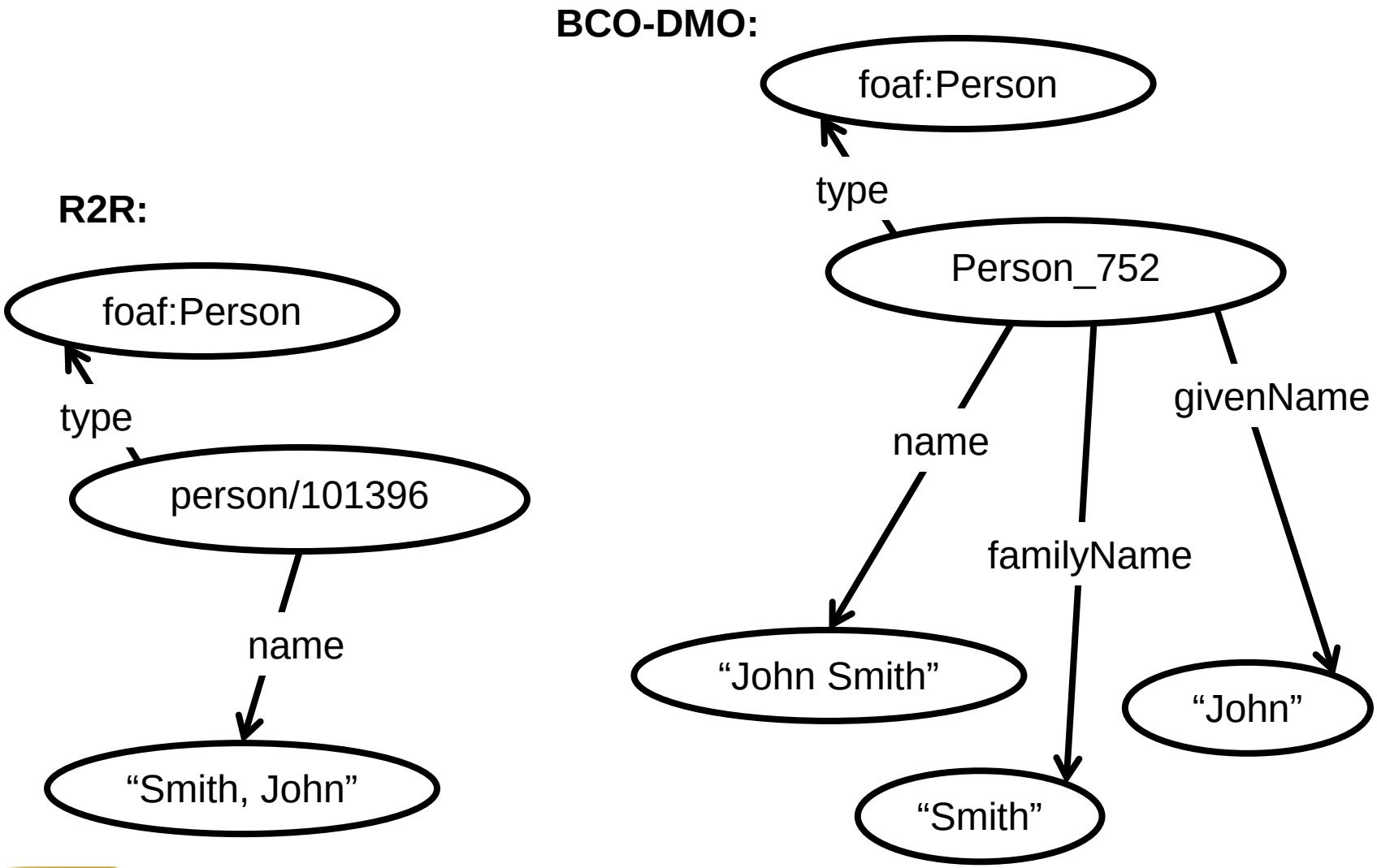


Works very well, but only in some very limited cases.

Cannot deal with graph representations of even very minimal complexity.



Automated federation?



Automated federation?

$$a{:}hasWife \sqsubseteq a{:}hasSpouse$$
$$symmetric(a{:}hasSpouse)$$
$$\exists a{:}hasSpouse.a{:}Female \sqsubseteq a{:}Male$$
$$\exists a{:}hasSpouse.a{:}Male \sqsubseteq a{:}Female$$
$$a{:}hasWife(a{:}john, a{:}mary)$$
$$b{:}Male(a{:}john)$$
$$b{:}Female(a{:}mary)$$
$$a{:}Male \sqcap a{:}Female \sqsubseteq \bot$$

$$symmetric(b{:}hasSpouse)$$
$$b{:}hasSpouse(b{:}mike, b{:}david)$$
$$b{:}Male(b{:}david)$$
$$b{:}Male(b{:}mike)$$
$$b{:}Female(b{:}anna)$$

Ways forward?

How to establish a flexible conceptual architecture using data and ontological modeling?

Ontology Design Patterns

DaSe Lab

“An ontology design pattern is a reusable successful solution to a recurrent modeling problem.”

So-called *content patterns* usually encode specific abstract notions, such as process, event, agent, etc.

Patterns provide modular, reusable, replaceable, pieces.

By agreeing on reuse of generic patterns (but leaving the relationships between the patterns to a specific assembly for a special purpose), we can have reuse while preserving heterogeneity.

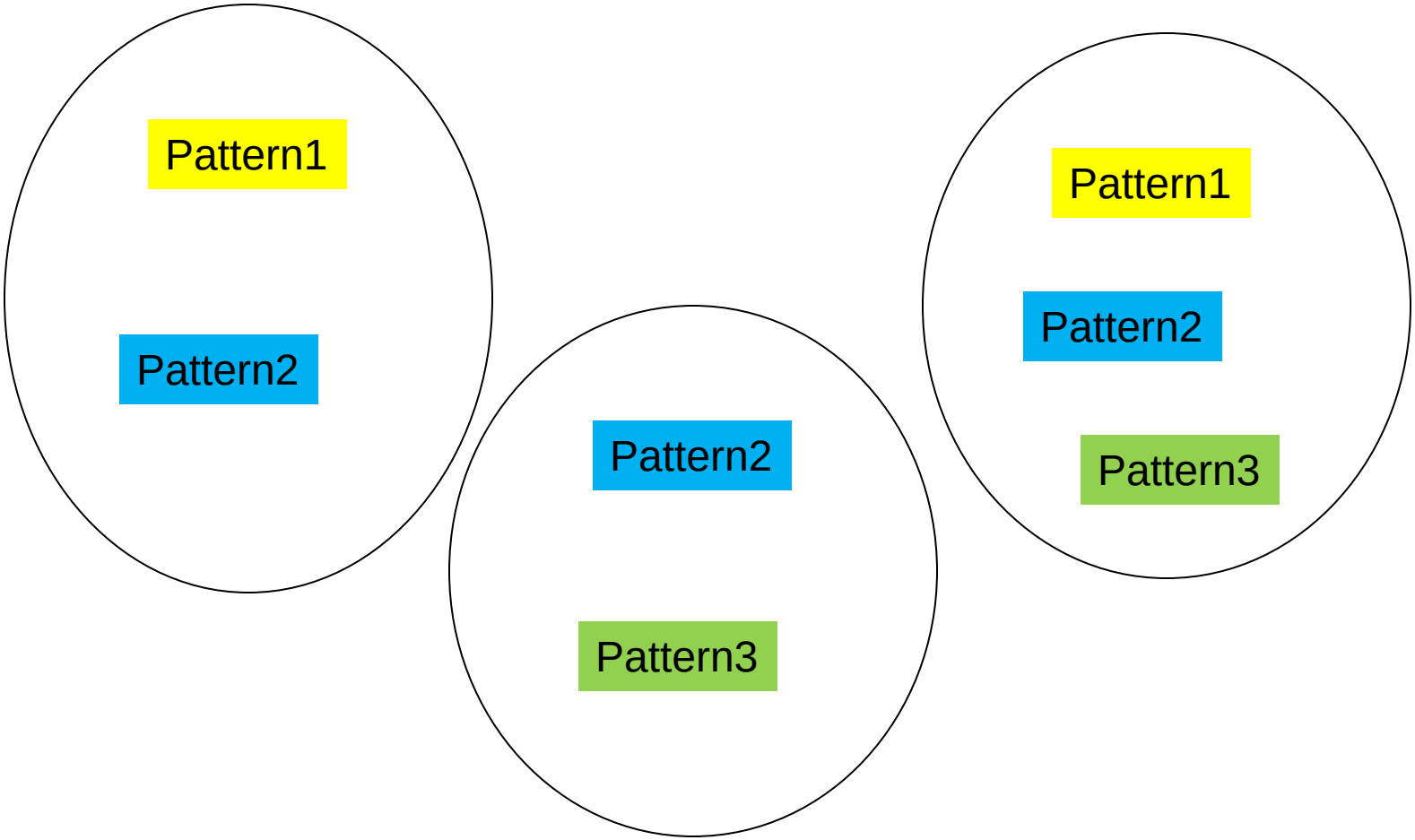
Ontology Design Patterns

DaSe Lab

- **Bottom-up homogenization of data representation.**
- **Avoidance of strong ontological commitments.**
- **Avoidance of standardization of specific modeling details.**
- **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

EarthCube

EarthCube:

Developing a Community-Driven Data and Knowledge Environment for the Geosciences

“concepts and approaches to create integrated data management infrastructures across the Geosciences.”

“EarthCube aims to create a well-connected and facile environment to share data and knowledge in an open, transparent, and inclusive manner, thus accelerating our ability to understand and predict the Earth system.”

OceanLink

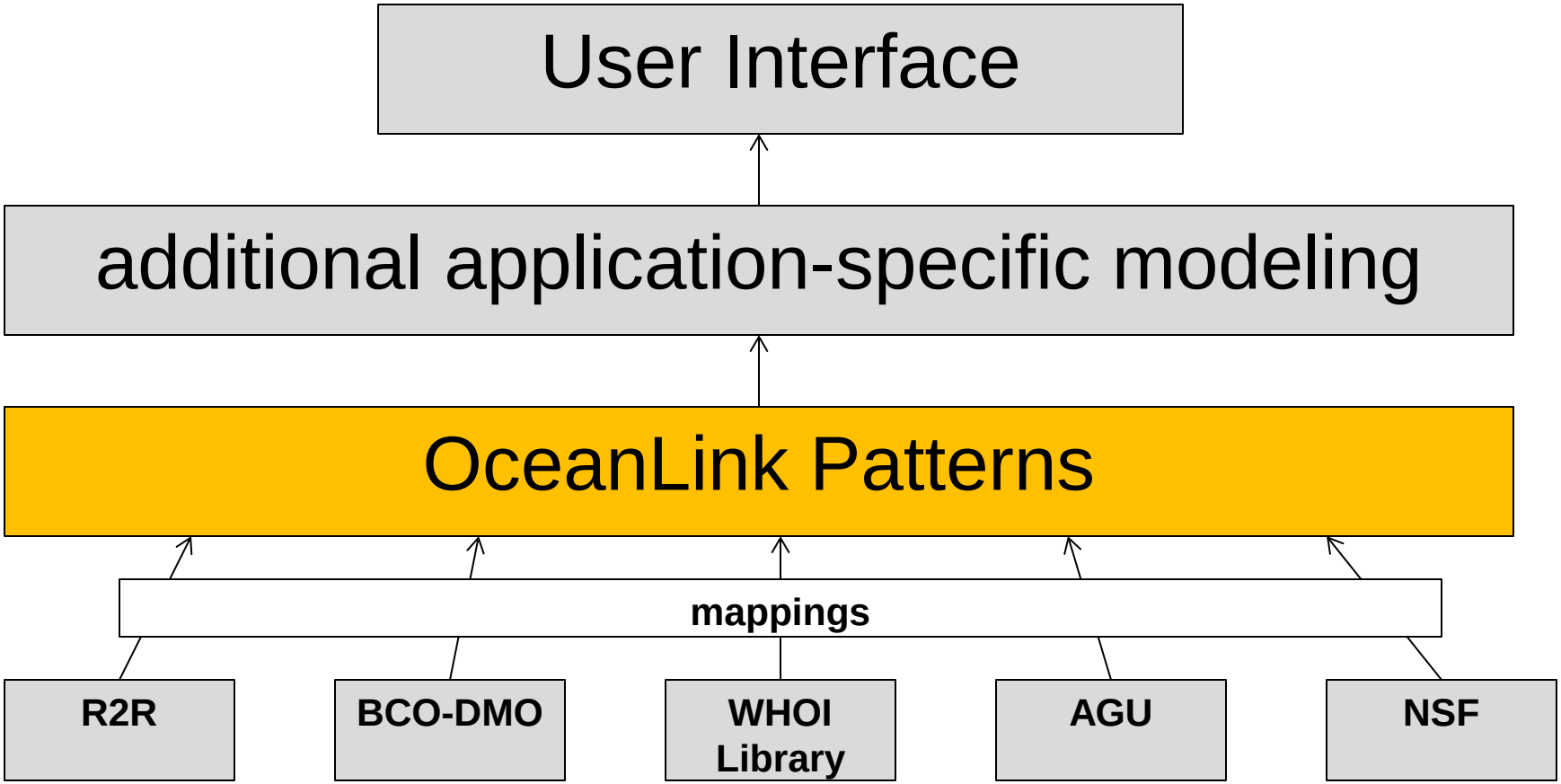
DaSe Lab

NSF EarthCube project “OceanLink”:

- **Integration of existing ocean science data repositories.**
- **For faceted browsing and semantic search.**
- **To be done in a flexible, extendable, modular way.**
- **With minimal effort for additional data providers to integrate their content.**

National Science Foundation award 1354778 "EAGER: Collaborative Research: EarthCube Building Blocks, Leveraging Semantics and Linked Data for Geoscience Data Sharing and Discovery."

OceanLink setup



OceanLink patterns

DaSe Lab

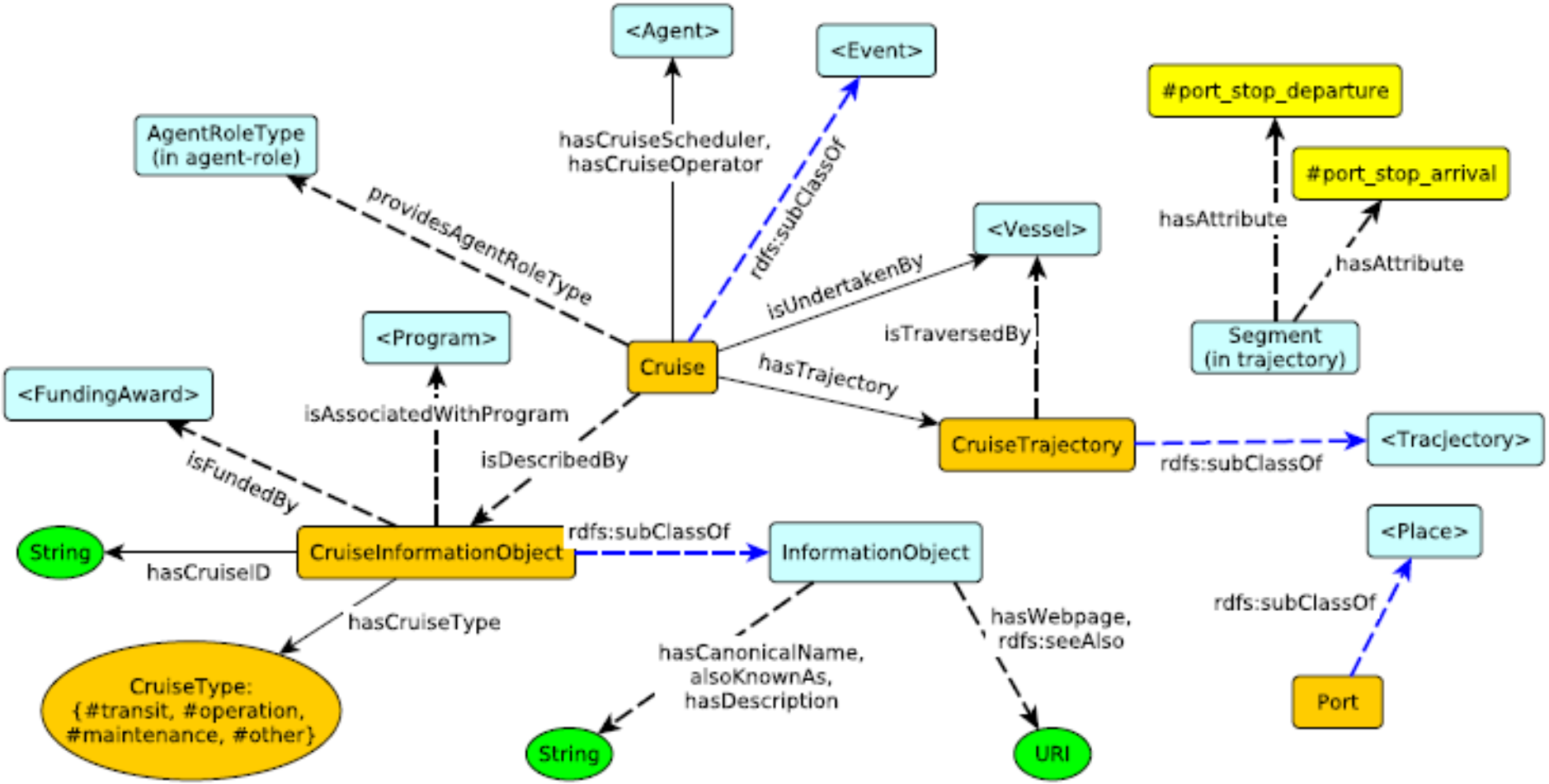
Some central patterns:

- **Cruise**
- **Trajectory**
- **Person**
- **Organization**
- **Roles of Agents**
- **Repository Object**
- **Data Set**
- **Document**

We’re not starting from zero of course.

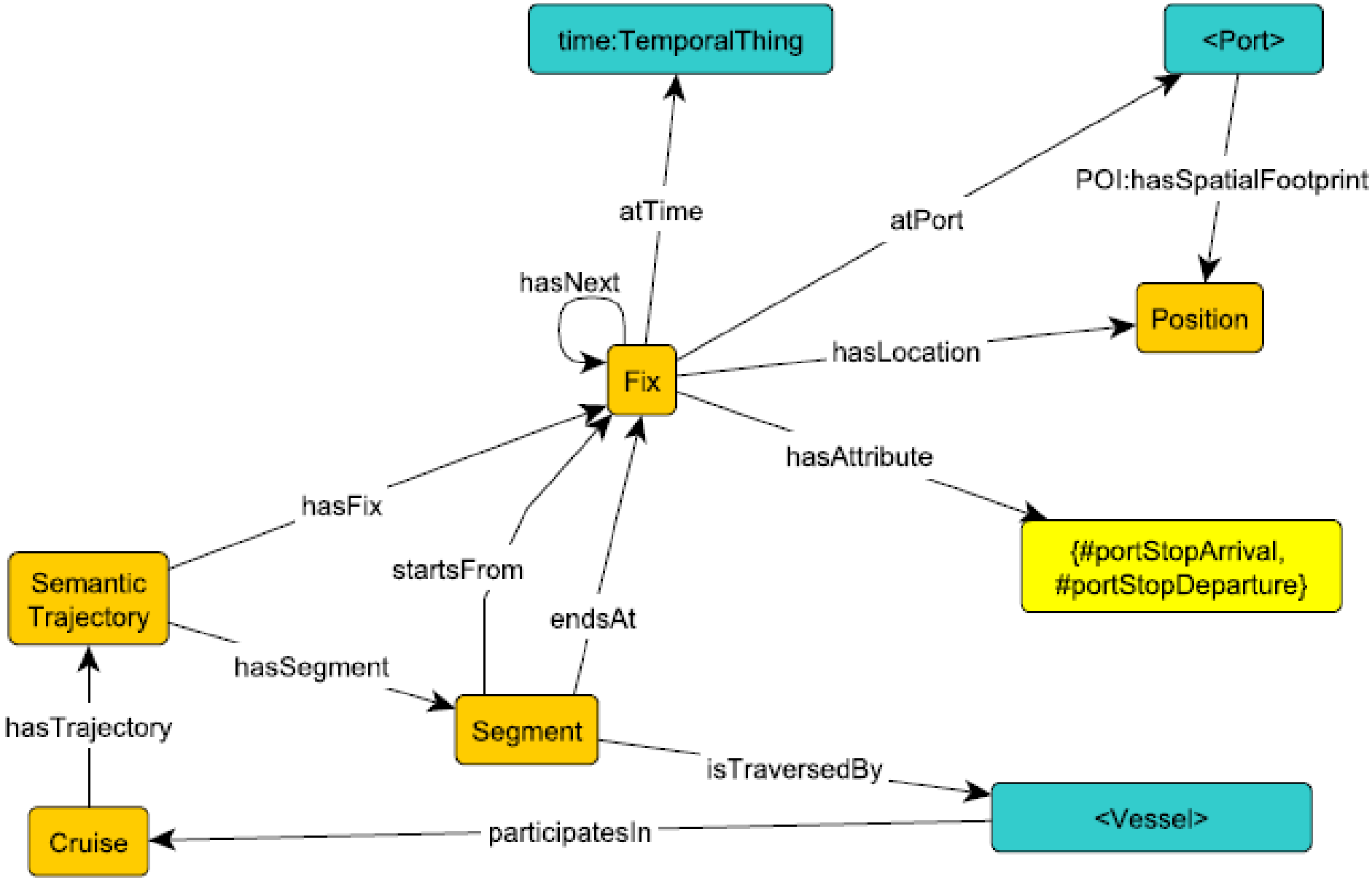
Ocean Science Cruise (draft)

DaSe Lab



Cruise trajectory (draft)

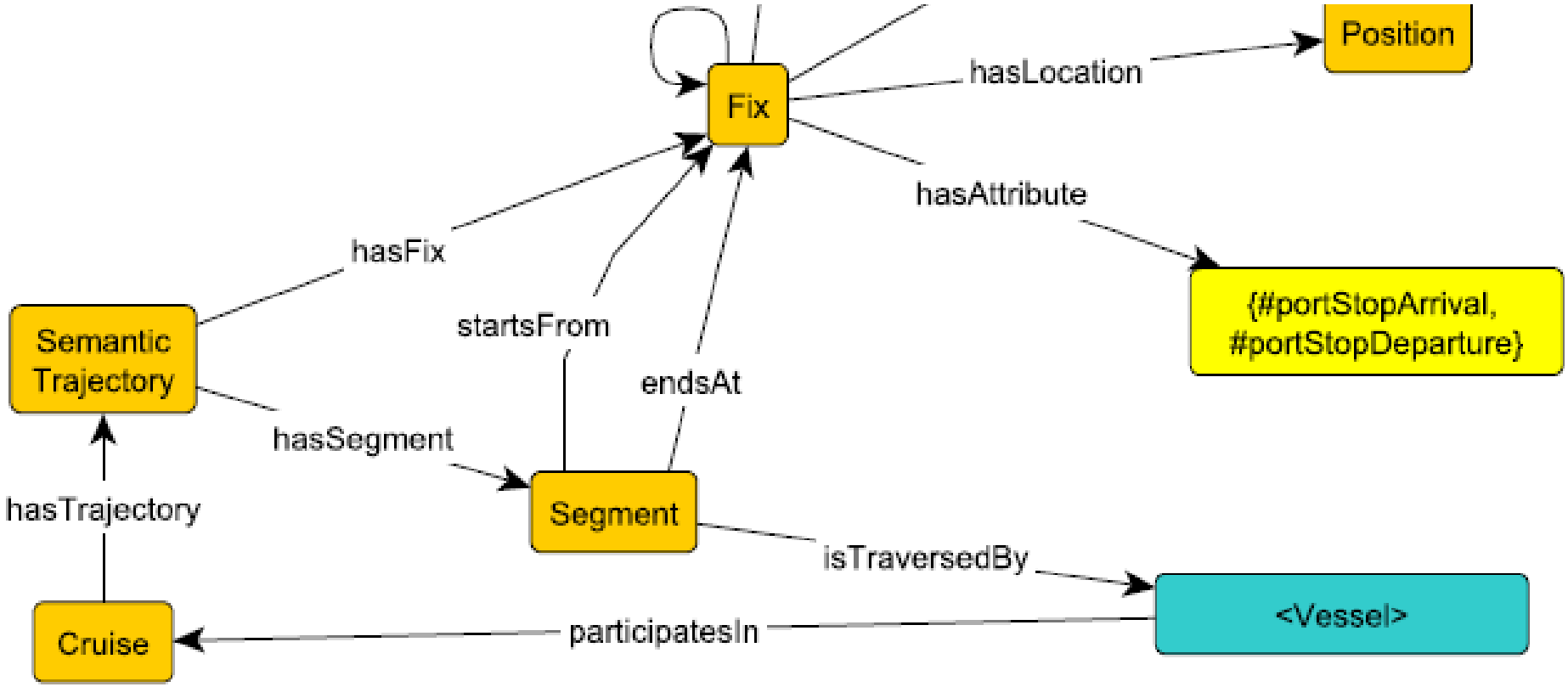
DaSe Lab



Cruise trajectory

DaSe Lab

$$\text{Cruise}(x) \wedge \text{hasTrajectory}(x, y)$$
$$\wedge\, \text{hasSegment}(y, z) \wedge \text{isTraversedBy}(z, v)$$
$$\rightarrow \text{participatesIn}(v, z)$$



Cruise trajectory

DaSe Lab

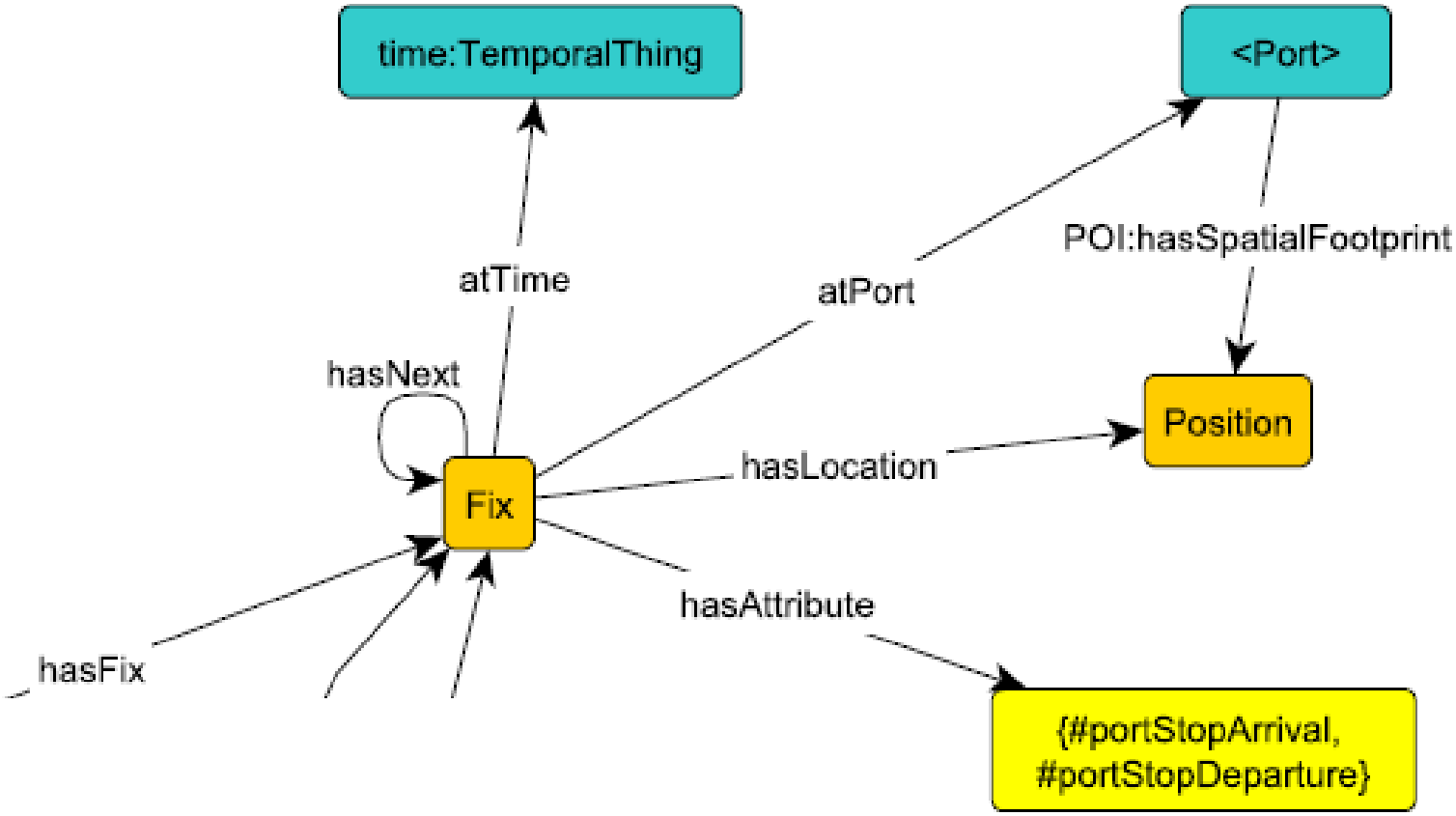
$$\text{Cruise}(x) \wedge \text{hasTrajectory}(x, y) \wedge \text{hasSegment}(y, z) \wedge \text{isTraversedBy}(z, v) \rightarrow \text{participatesIn}(v, z)$$

$$\text{Cruise} \equiv \exists \text{cruise}.\text{Self}$$

$$\text{cruise} \circ \text{hasTrajectory} \circ \text{hasSegment} \circ \text{isTraversedBy} \sqsubseteq \text{hasParticipant}$$

$$\text{hasParticipant} \equiv \text{participatesIn}^{-}$$

Cruise trajectory



$$\mathrm{Fix}(x) \wedge \mathrm{hasAttribute}(x, \#\mathrm{portStopArrival})$$
$$\wedge\, \mathrm{atPort}(x, y) \wedge \mathrm{hasSpatialFootprint}(y, z)$$
$$\wedge\, \mathrm{hasLocation}(x, w) \rightarrow \mathrm{locatedIn}(w, z)$$

Cruise trajectory

$$\mathrm{Fix}(x) \wedge \mathrm{hasAttribute}(x, \#\mathrm{portStopArrival}) \wedge \mathrm{atPort}(x, y) \wedge \mathrm{hasSpatialFootprint}(y, z) \wedge \mathrm{hasLocation}(x, w) \rightarrow \mathrm{locatedIn}(w, z)$$

$$\mathrm{Fix} \wedge \exists \mathrm{hasTrajectory}.\{\#\mathrm{portStopArrival}\} \equiv \exists \mathrm{fixps}.\mathrm{Self}$$

$$\mathrm{hasLocation}^{-} \circ \mathrm{fixps} \circ \mathrm{atPort} \circ \mathrm{hasSpatialFootprint} \sqsubseteq \mathrm{locatedIn}$$

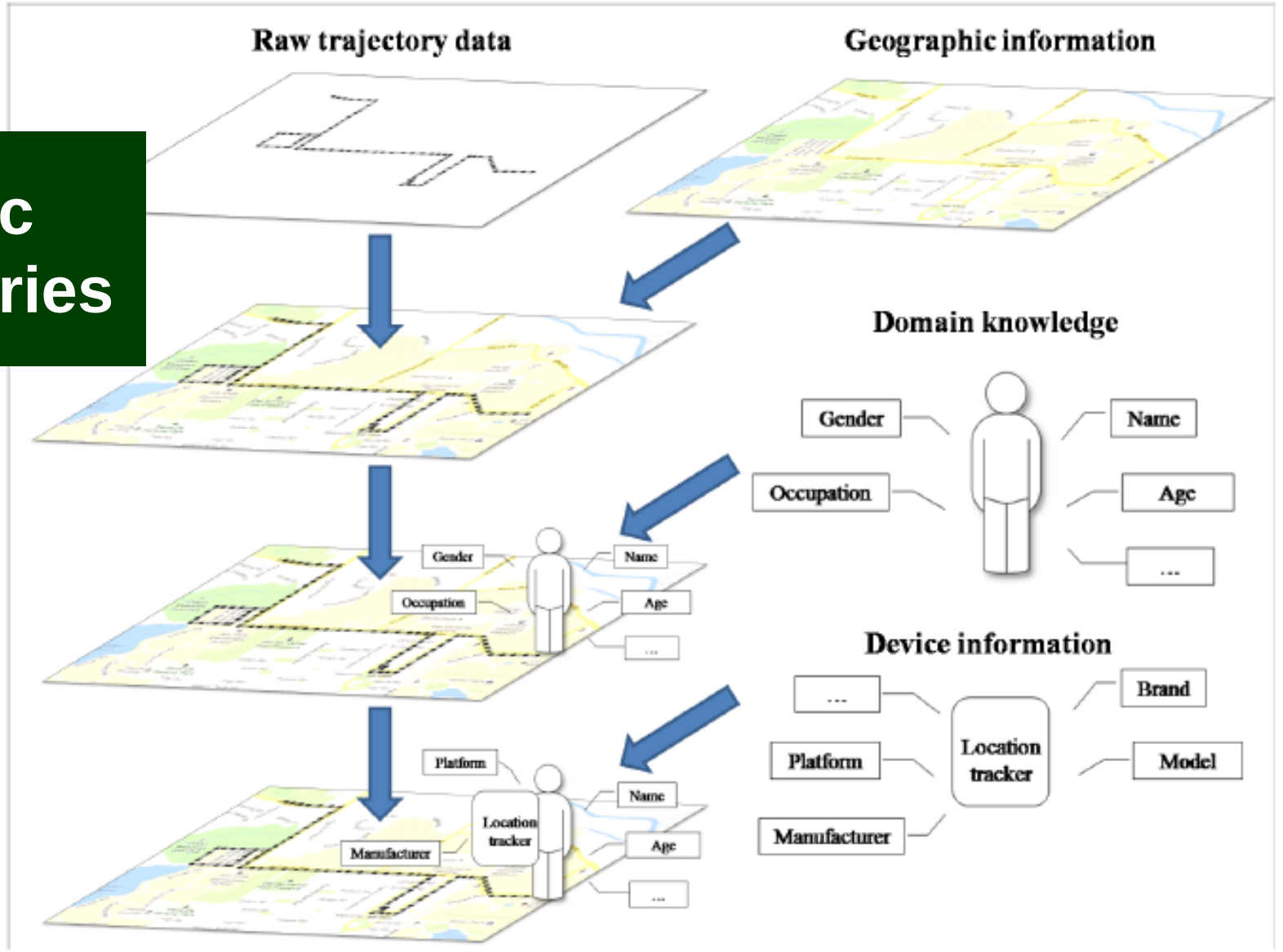
Ways forward

- **Establish a flexible conceptual architecture using data and ontological modeling.**
- **A principled use of patterns, including**
  - **the development of a theory of patterns and**
  - **the provision of a critical amount of central patterns**

  may provide a primary path forward.

Thanks!

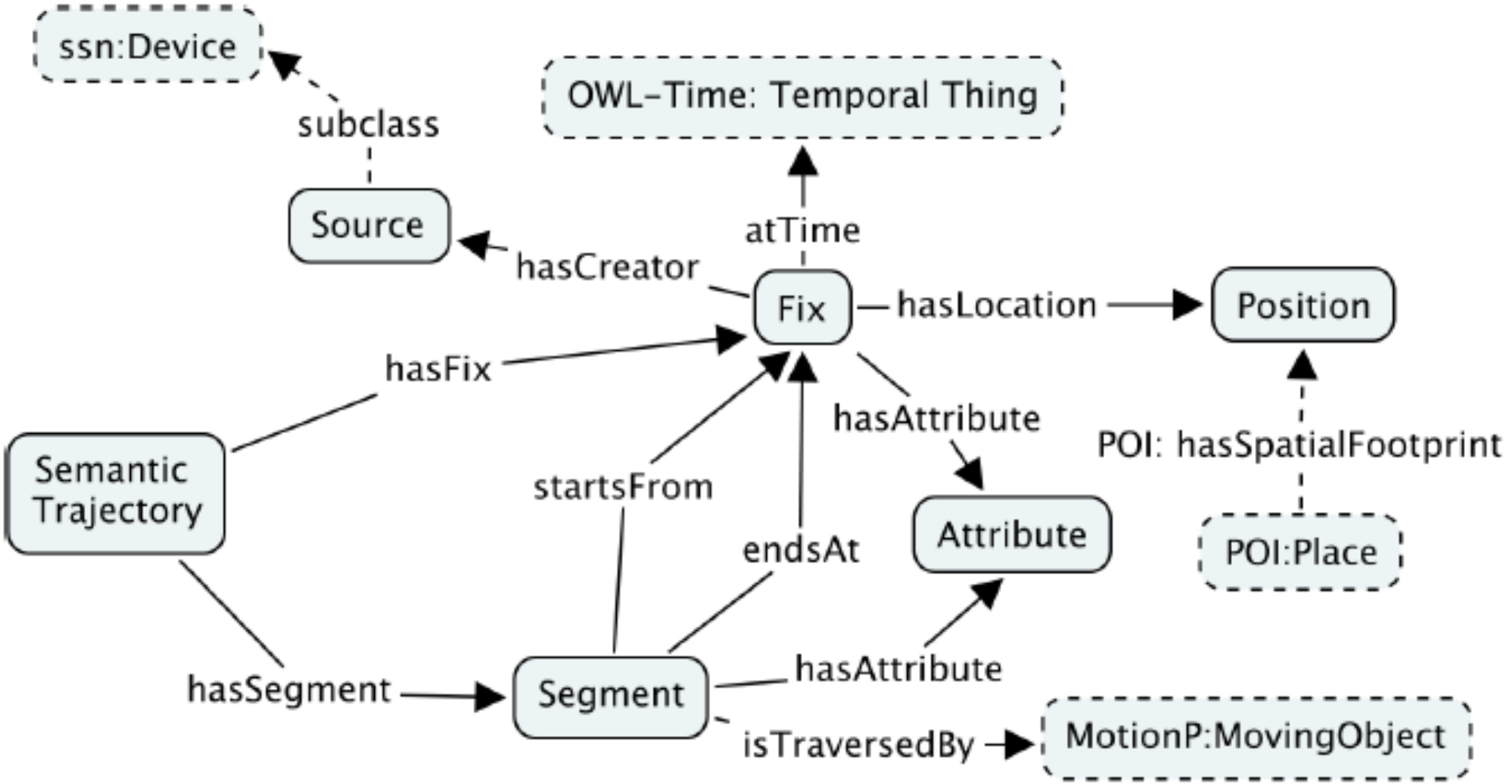
Semantic Trajectories



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

Semantic Trajectories

DaSe Lab



Semantics in OWL

DaSe Lab

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

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

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

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

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

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

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

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

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

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

Semantics in OWL

DaSe Lab

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

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

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

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

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

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

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

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

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

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

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

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

- D. Oberle, A. Ankolekar, P. Hitzler, P. Cimiano, M. Sintek, M. Kiesel, B. Mougouie, S. Vembu, S. Baumann, M. Romanelli, P. Buitelaar, R. Engel, D. Sonntag, N. Reithinger, B. Loos, R. Porzel, H.-P. Zorn, V. Micelli, C. Schmidt, M. Weiten, F. Burkhardt, J. Zhou, DOLCE ergo SUMO: On Foundational and Domain Models in the SmartWeb Integrated Ontology (SWIntO). Journal of Web Semantics: Science, Services and Agents on the World Wide Web 5 (3), 2007, 156-174.
- Adila Krisnadhi, Robert Arko, Suzanne Carbotte, Cynchia Chandler, Michelle Cheatham, Timothy Finin, Pascal Hitzler, Krzysztof Janowicz, Thomas Narock, Lisa Raymond, Adam Shepherd, Peter Wiebe, An Ontology Pattern for Oceanograhic Cruises: Towards an Oceanograhper's Dream of Integrated Knowledge Discovery. OceanLink Technical Report 2014.1.

References

- Sebastian Rudolph, Markus Krötzsch, Pascal Hitzler, Cheap Boolean Role Constructors for Description Logics. In: Steffen Hölldobler and Carsten Lutz and Heinrich Wansing (eds.), Proceedings of 11th European Conference on Logics in Artificial Intelligence (JELIA), volume 5293 of LNAI, pp. 362-374. Springer, September 2008.
- Adila Alfa Krisnadhi, Frederick Maier, Pascal Hitzler, OWL and Rules. In: A. Polleres, C. d'Amato, M. Arenas, S. Handschuh, P. Kroner, S. Ossowski, P.F. Patel-Schneider (eds.), Reasoning Web. Semantic Technologies for the Web of Data. 7th International Summer School 2011, Galway, Ireland, August 23-27, 2011, Tutorial Lectures. Lecture Notes in Computer Science Vol. 6848, Springer, Heidelberg, 2011, pp. 382-415.