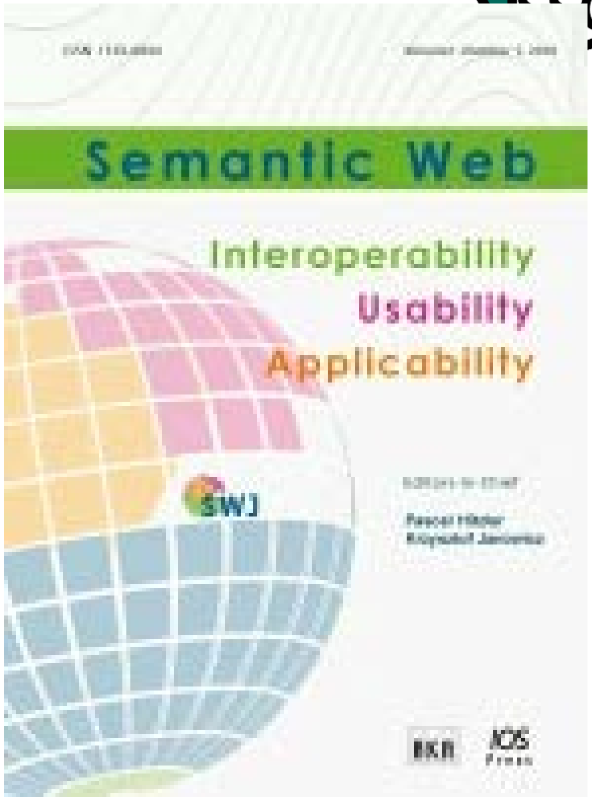
# Ontology Modeling

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

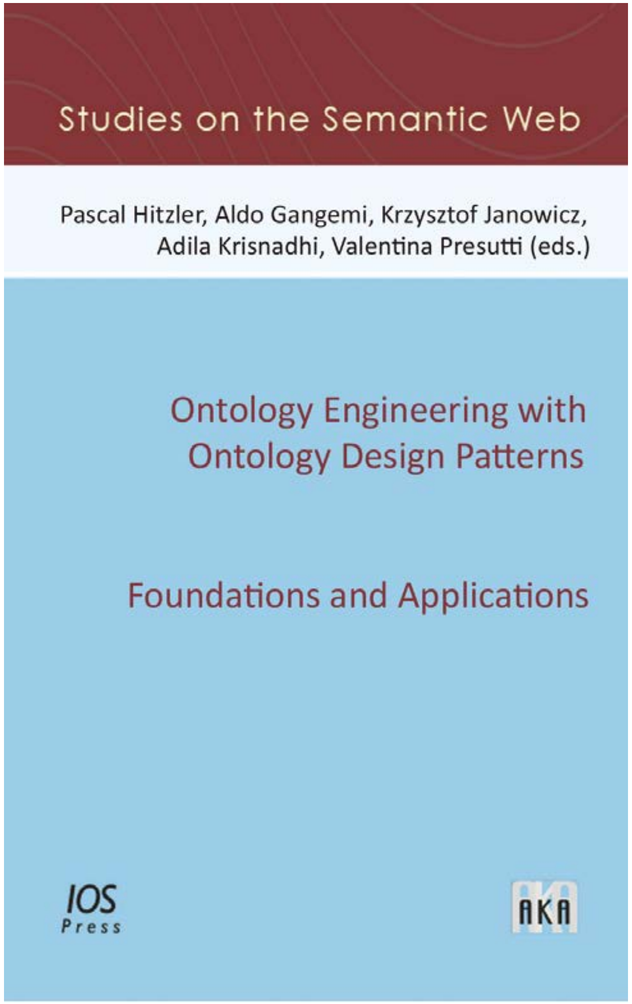
# Semantic Web journal

- **EiCs:** Pascal Hitzler
  Krzysztof Janowicz
- **Funded 2010**
- **2016 Impact factor of 1.786, top 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/**

# Forthcoming book




**Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnathi, Valentina Presutti (eds.), Ontology Engineering with Ontology Design Patterns: Foundations and Applications. Studies on the Semantic Web. IOS Press/AKA Verlag, 2016/2017. To appear.**

## Motivation: EarthCube and GeoLink

# EarthCube GeoLink Scenario

**The NSF EarthCube Program:**

**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.”**

# EarthCube GeoLink Scenario

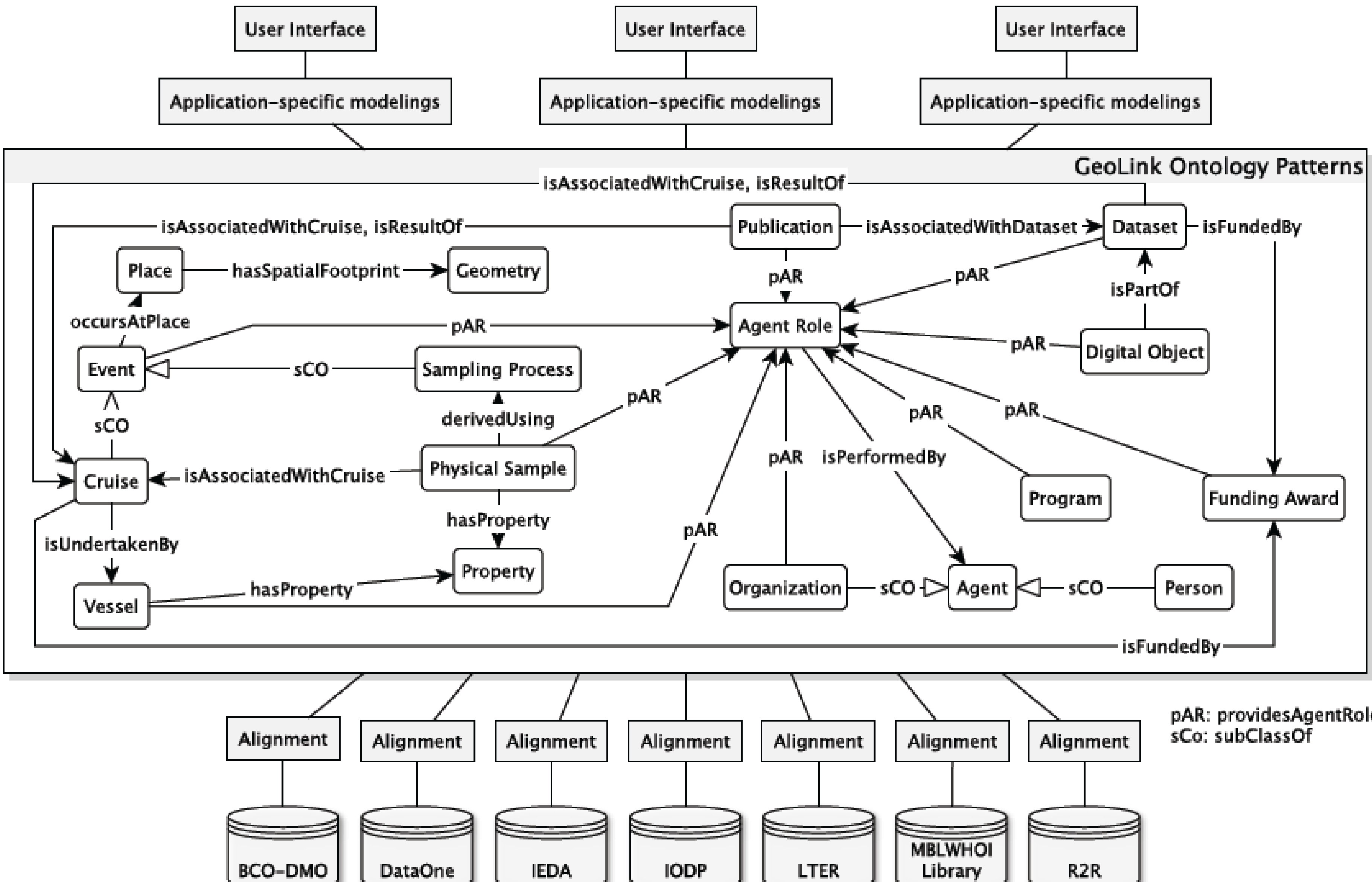
**GeoLink: An EarthCube "Building Block" project (2014-2017)**

**How to realize data search across many large-scale geoscience data repositories, such that**

- **The approach is extendable to new repositories.**
- **The scope can extend across all of the Geosciences.**
- **The search capabilities can be made more fine-grained in the future if desired.**

**Central idea: Use a modular, extendable ontology for the integration of metadata.**

# The GeoLink Framework

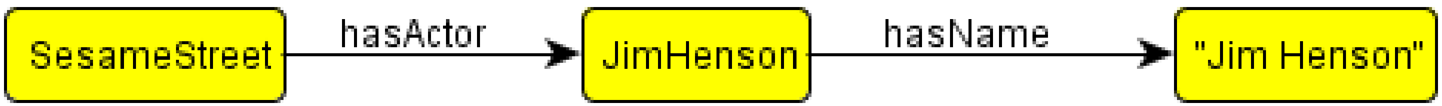



pAR: providesAgentRole
sCo: subClassOf

# Modeling Patterns

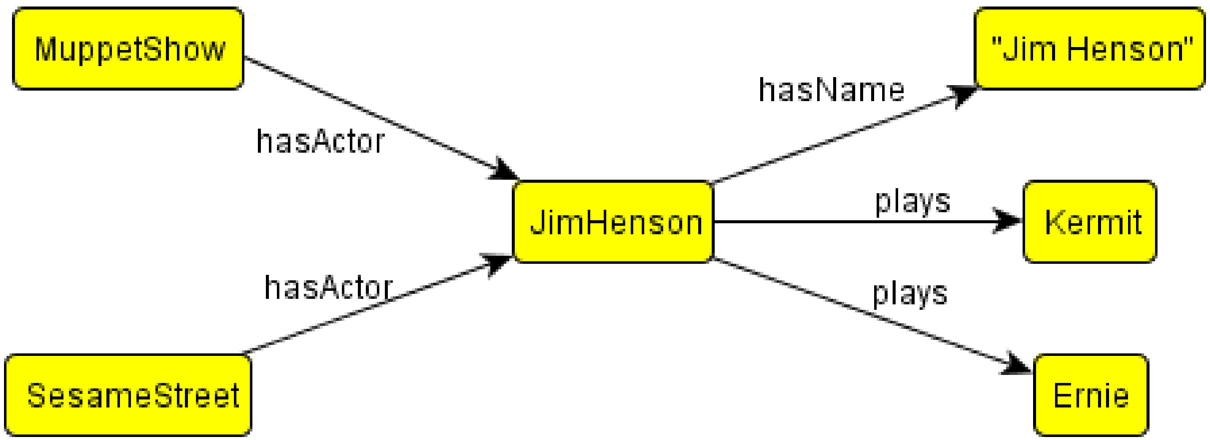
# ABoxes as graphs

| | | |
|---|---|---|
| **SesameStreet** | **has Actor** | **JimHenson .** |
| **JimHenson** | **hasName** | **“Jim Henson” .** |

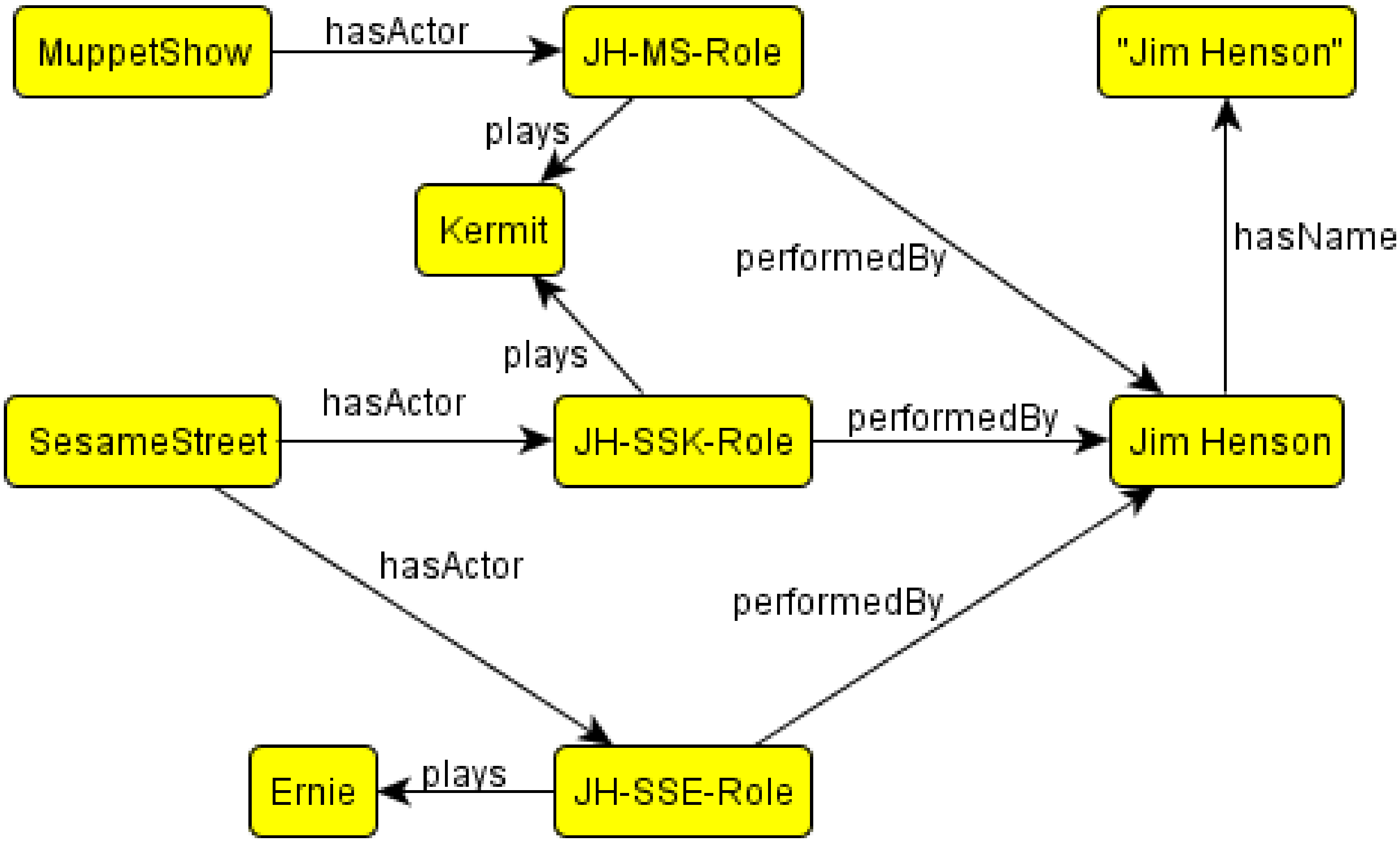


# Problem!

| | | |
|---|---|---|
| **SesameStreet** | **has Actor** | **JimHenson .** |
| **MuppetShow** | **has Actor** | **JimHenson .** |
| **JimHenson** | **plays** | **Kermit .** |
| **JimHenson** | **plays** | **Ernie .** |
| **JimHenson** | **hasName** | **“Jim Henson” .** |

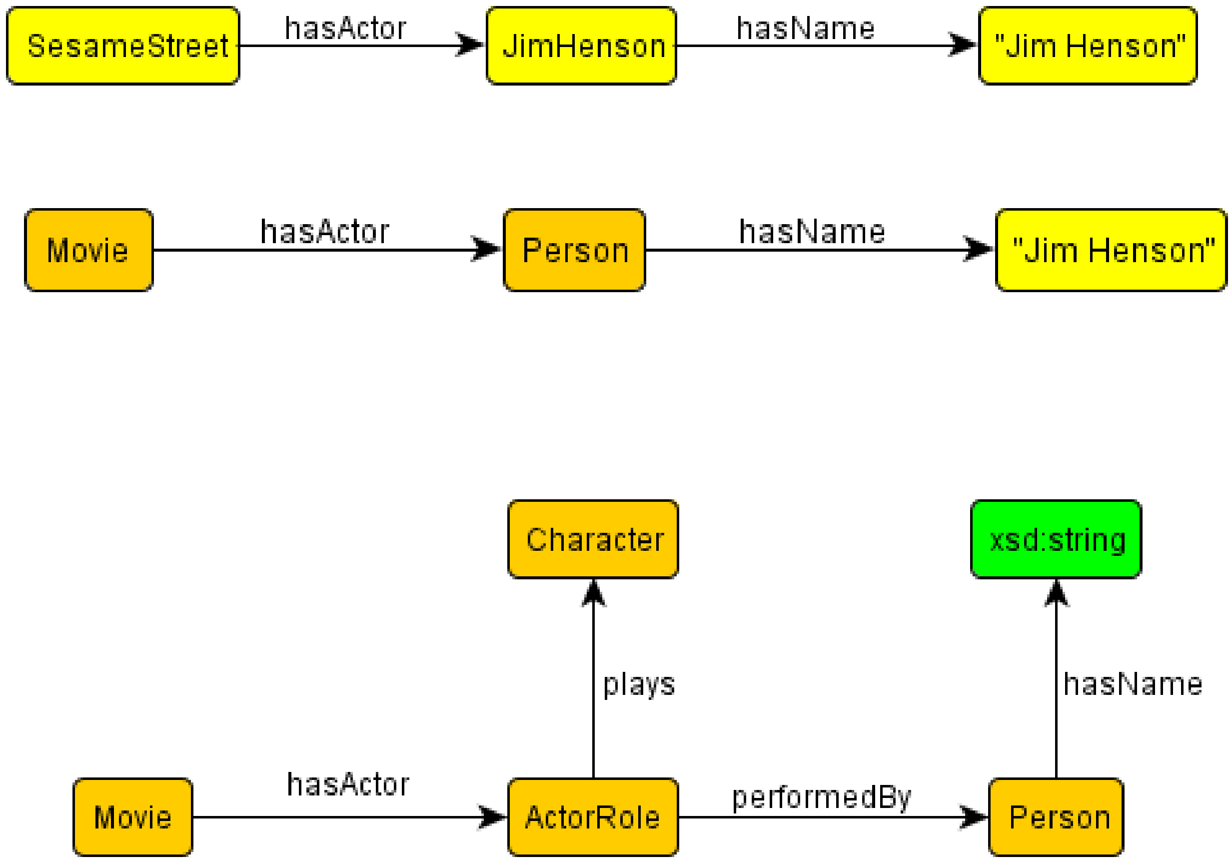


# Solution!

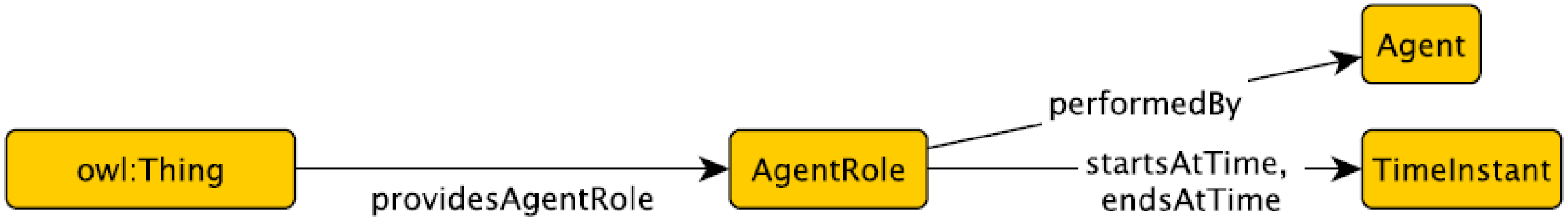


MuppetShow
hasActor
JH-MS-Role
"Jim Henson"
plays
Kermit
performedBy
hasName
plays
SesameStreet
hasActor
JH-SSK-Role
performedBy
Jim Henson
hasActor
performedBy
Ernie
plays
JH-SSE-Role

# Schematically



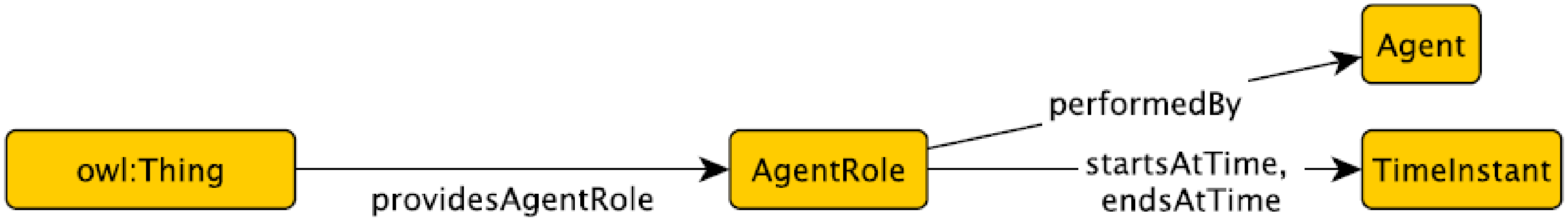
# The AgentRole Pattern




**An *Ontology Design Pattern* (ODP) is a reusable successful solution to a recurrent ontology modeling problem.**

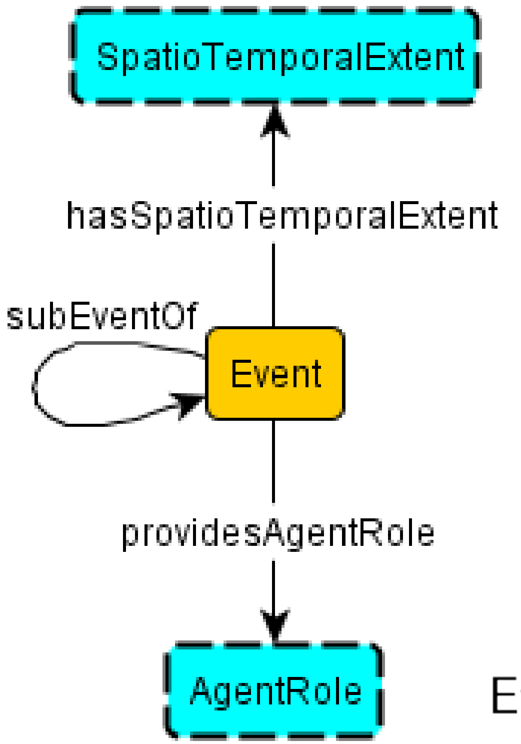
**[Gangemi 2005]**

# Axiomatization




$$\top \sqsubseteq \forall \mathsf{providesAgentRole}.\mathsf{AgentRole}$$
$$\mathsf{AgentRole} \sqsubseteq \forall \mathsf{performedBy}.\mathsf{Agent}$$
$$\exists \mathsf{performedBy}.\mathsf{Agent} \sqsubseteq \mathsf{AgentRole}$$
$$\mathsf{AgentRole} \sqsubseteq \forall \mathsf{startsAtTime}.\mathsf{TimeInstant}$$
$$\mathsf{AgentRole} \sqsubseteq \forall \mathsf{endsAtTime}.\mathsf{TimeInstant}$$
$$\mathsf{AgentRole} \sqsubseteq \exists \mathsf{providesAgentRole}^{-}.\top$$
$$\mathsf{AgentRole} \sqsubseteq {=}1 \mathsf{performedBy}.\mathsf{Agent}$$
$$\mathsf{AgentRole} \sqsubseteq {=}1 \mathsf{startsAtTime}.\mathsf{TimeInstant}$$
$$\mathsf{AgentRole} \sqsubseteq {=}1 \mathsf{endsAtTime}.\mathsf{TimeInstant}$$
$$\mathsf{DisjointClasses}(\mathsf{AgentRole}, \mathsf{Agent}, \mathsf{TimeInstant})$$

# A Minimal Event Pattern




$$\top \sqsubseteq \forall \text{hasSpatioTemporalExtent}.\text{SpatioTemporalExtent}$$

$$\top \sqsubseteq \forall \text{providesAgentRole}.\text{AgentRole}$$

$$\text{Event} \sqsubseteq \exists \text{hasSpatioTemporalExtent}.\text{SpatioTemporalExtent}$$

$$\text{Event} \sqsubseteq \forall \text{subEventOf}.\text{Event}$$

$$\exists \text{subEventOf}.\text{Event} \sqsubseteq \text{Event}$$

$$\text{subEventOf} \circ \text{subEventOf} \sqsubseteq \text{subEventOf}$$

$$\text{DisjointClasses}(\text{Event}, \text{AgentRole}, \text{SpatioTemporalExtent})$$

# Worked Example: Chess

# Worked Example: Chess

- **Establish a searchable repository for chess data.**

- **Starting point are PGN files.**
- **Should be extendable with other information from**
  - **Chess websites**
  - **Wikipedia**
  - **Geographic data**
  - **News**
  - **Etc.**

- **Use an ontology for information integration.**

# GeoVoCamps modeling approach

- Collaborative modeling, group ideally has
  - More than one domain experts.
  - People familiar with the base data.
  - People understanding possible target use cases.
  - An ontology engineer familiar with the modeling approach.
  - Somebody who understands formal semantics of OWL.

- Domain experts are queried as to the main notions for the application domain.
  - E.g. for chess, these would include
    - Chess game; move; opening; tournament; players; commentary

# GeoVoCamps modeling approach

- From available data and from application use cases, devise competency questions, i.e. questions which should be convertible into queries, which in turn should be answerable using the data.

1. Who played against Kasparov in the round 1994 Lineares tournament? Did (s)he play as a white or black player?
2. What is the first move taken by the black player in the Sicilian Defense opening?
3. Find all games in which Bobby Fischer, playing black, lost in the poisoned pawn variation of the Sicilian Defence opening.
4. Are there any recorded games using the Grünfeld Defence from before the 20th century?
5. What did Kasparov say about his opponent's first two moves in his commentary about his game against Topalov in the 1999 Tournament in Wijk ann Zee?
6. Who was the first non-Russian world champion after Fischer?
7. Did Bobby Fischer ever play against a grandmaster in Germany?
8. List all world championship games won by forfeit.

# GeoVoCamps modeling approach

- **Then prioritize which notions to model first. In the chess case, e.g.**

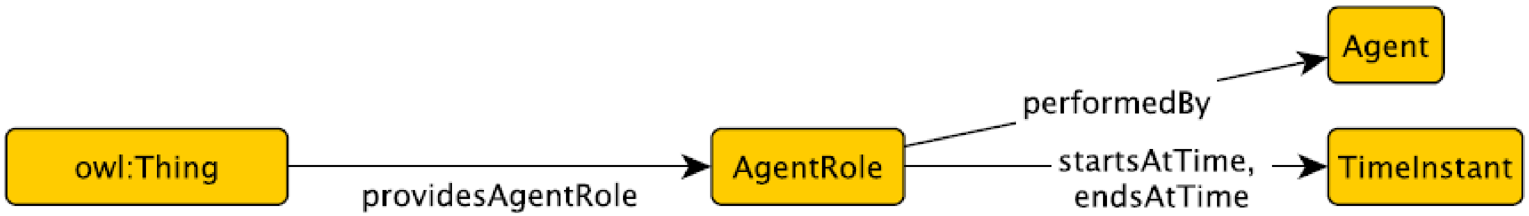
  **chess game**
  **move/half-move**
  **players**
  **opening**
  **tournaments**
  **commentary**

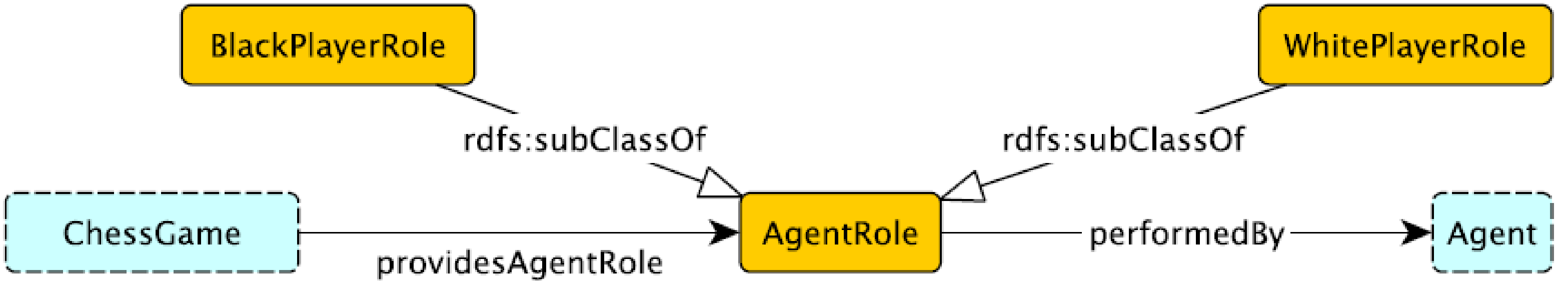
# GeoVoCamps modeling approach

- **Understand the nature of the things you are modeling.**

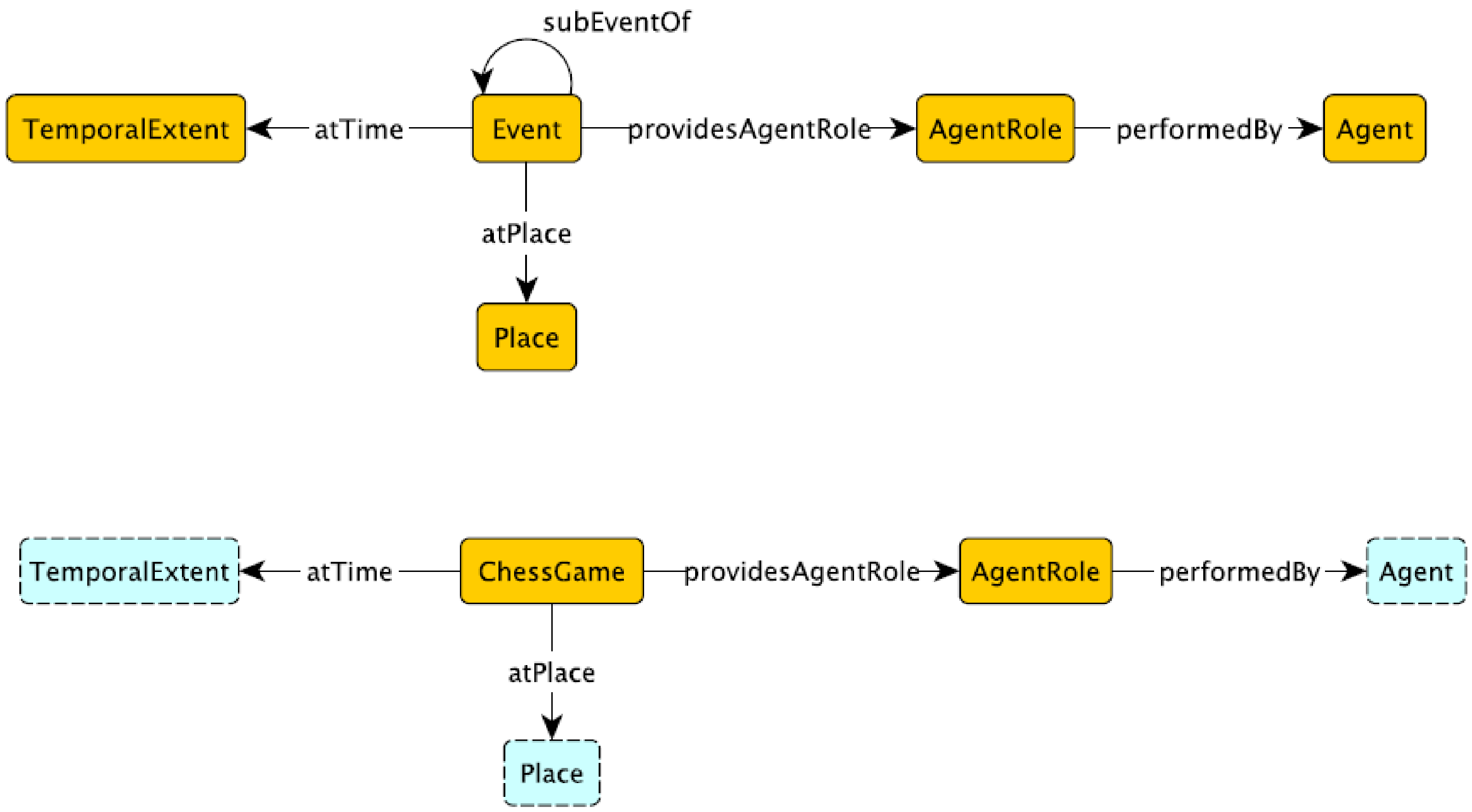
| | | |
|---|---|---|
| **Chess game** | **...** | **An Event** |
| **Half-move** | **…** | **A Subevent of a chess game** |
| **Player** | **…** | **The Role of an Agent** |
| **Opening** | **…** | **this is probably complex** |
| **tournaments** | **…** | **Events** |
| **commentary** | **…** | **this is again more complex** |

# Player as AgentRole

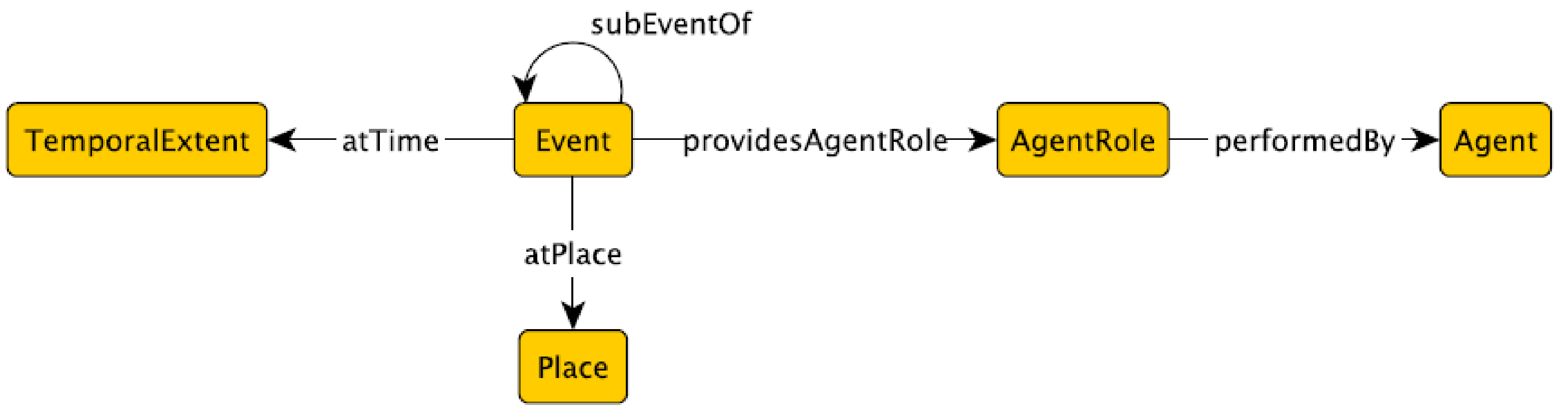


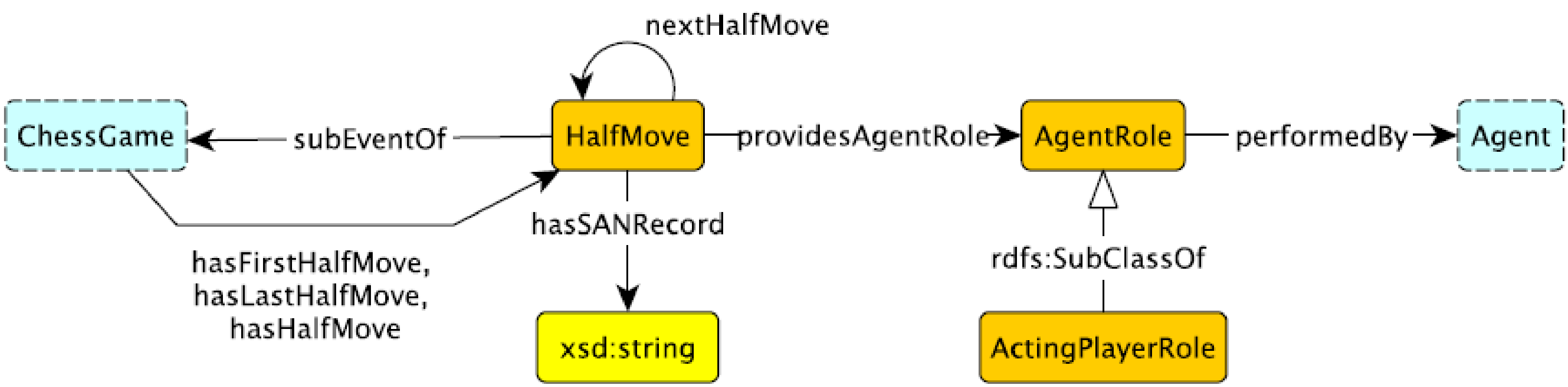



# ChessGame as Event



# Half-moves


subEventOf
TemporalExtent
atTime
Event
providesAgentRole
AgentRole
performedBy
Agent
atPlace
Place

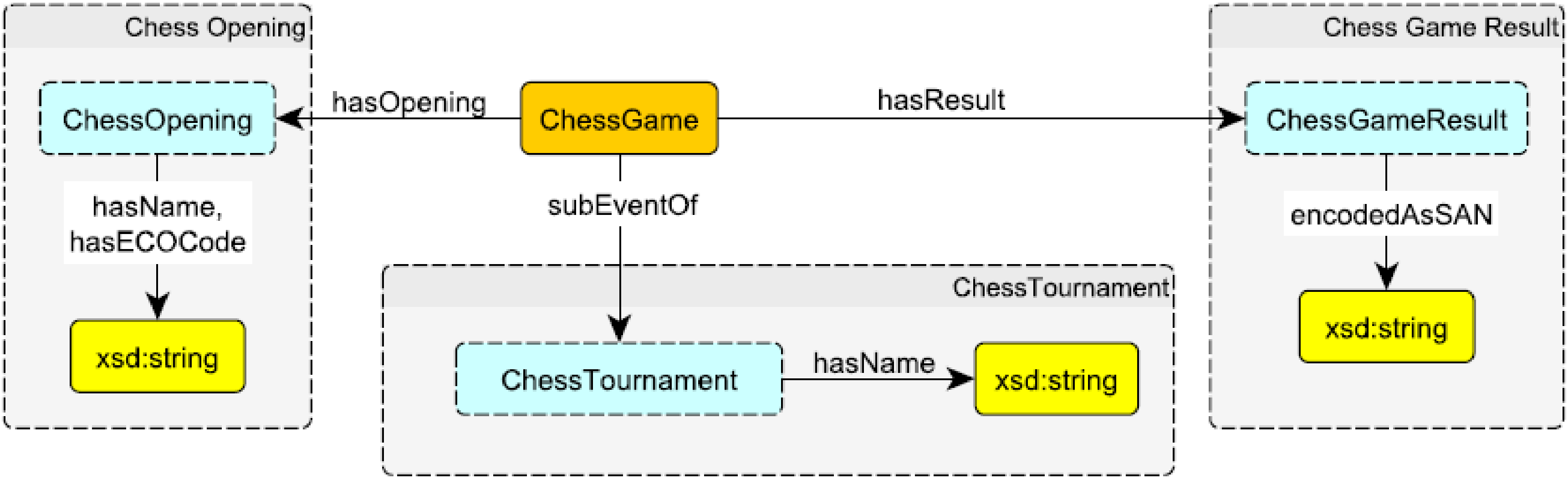

nextHalfMove
ChessGame
subEventOf
HalfMove
providesAgentRole
AgentRole
performedBy
Agent
hasSANRecord
hasFirstHalfMove,
hasLastHalfMove,
hasHalfMove
xsd:string
rdfs:SubClassOf
ActingPlayerRole

# Opening, game result, etc.

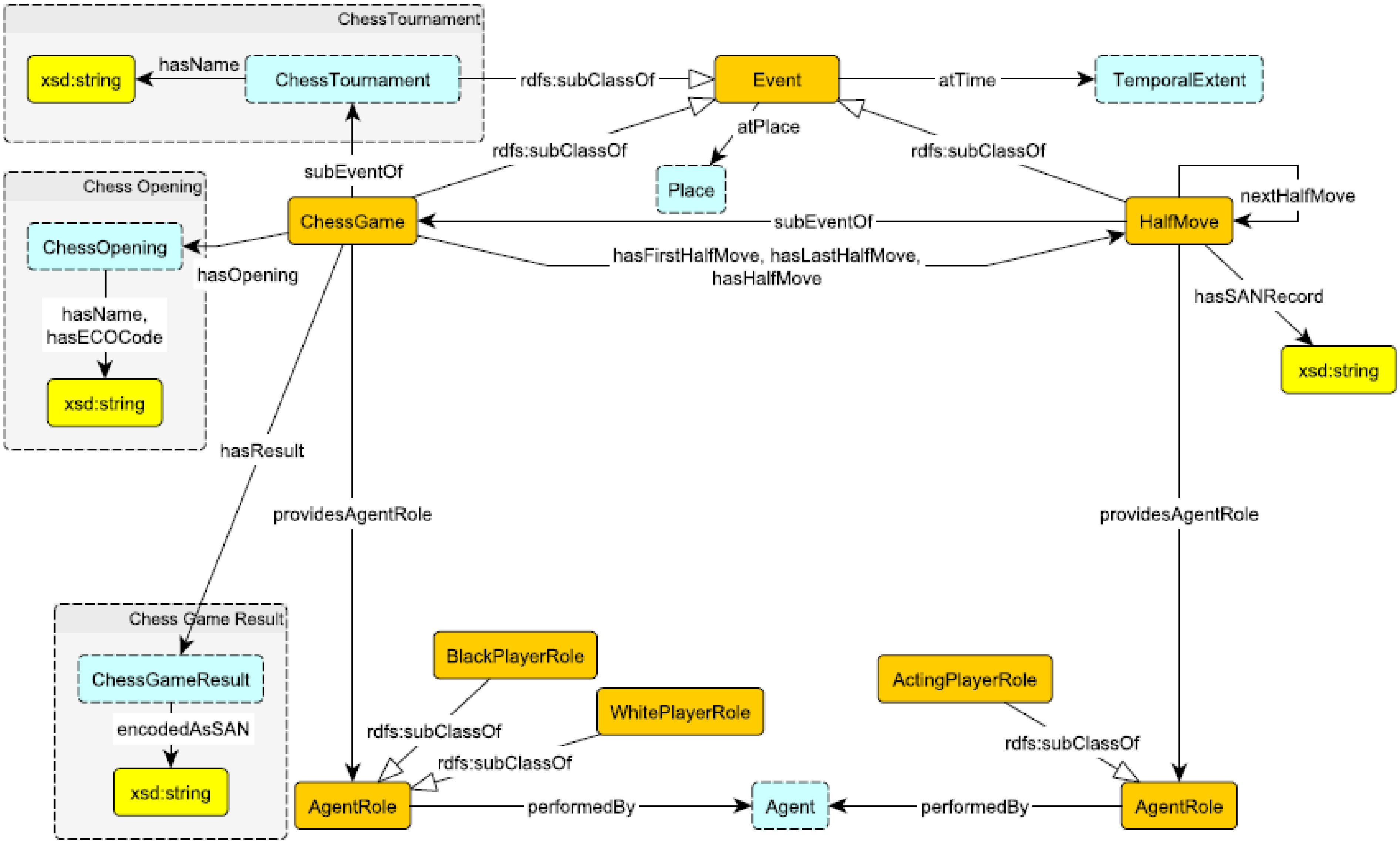
We call these "stubs".

I.e. we're aware that more fine-grained modeling will be needed for some use cases.

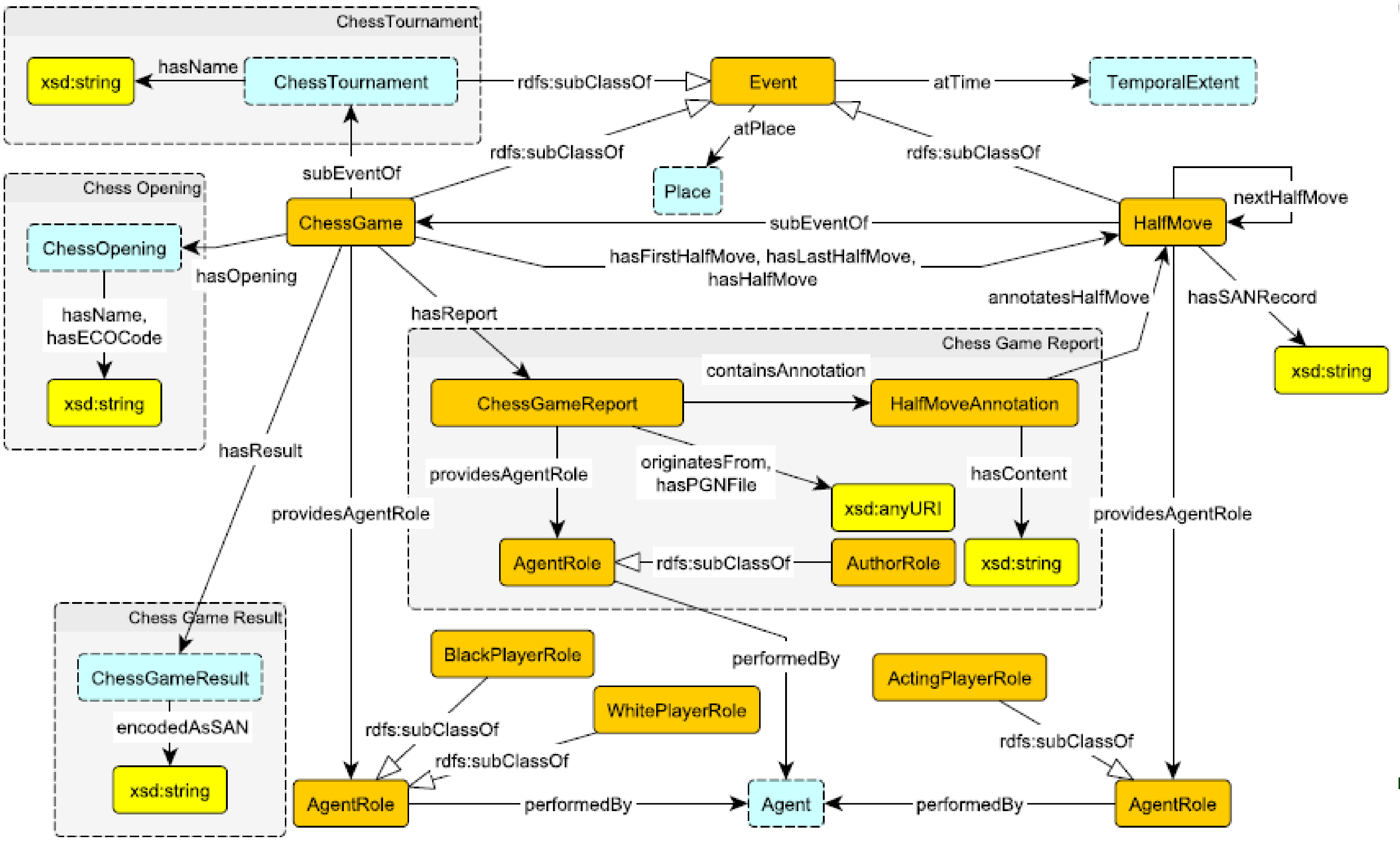
But currently there's no reason to do it (not in use case, no data), so we only provide "hooks" for future development of the ontology.



# Putting things together



# Adding commentaries



# Adequacy check

- Triplify sample data using the ontology.
  Does it work?

- Check if competency questions can be answered.

- Add axioms as appropriate (the graph is only for intuition, the OWL axioms are the actual ontology).

- (there are more post-hoc details to be taken care of, but let's leave it at that)

# Axioms

**Axioms in this case are mostly straightforward:**

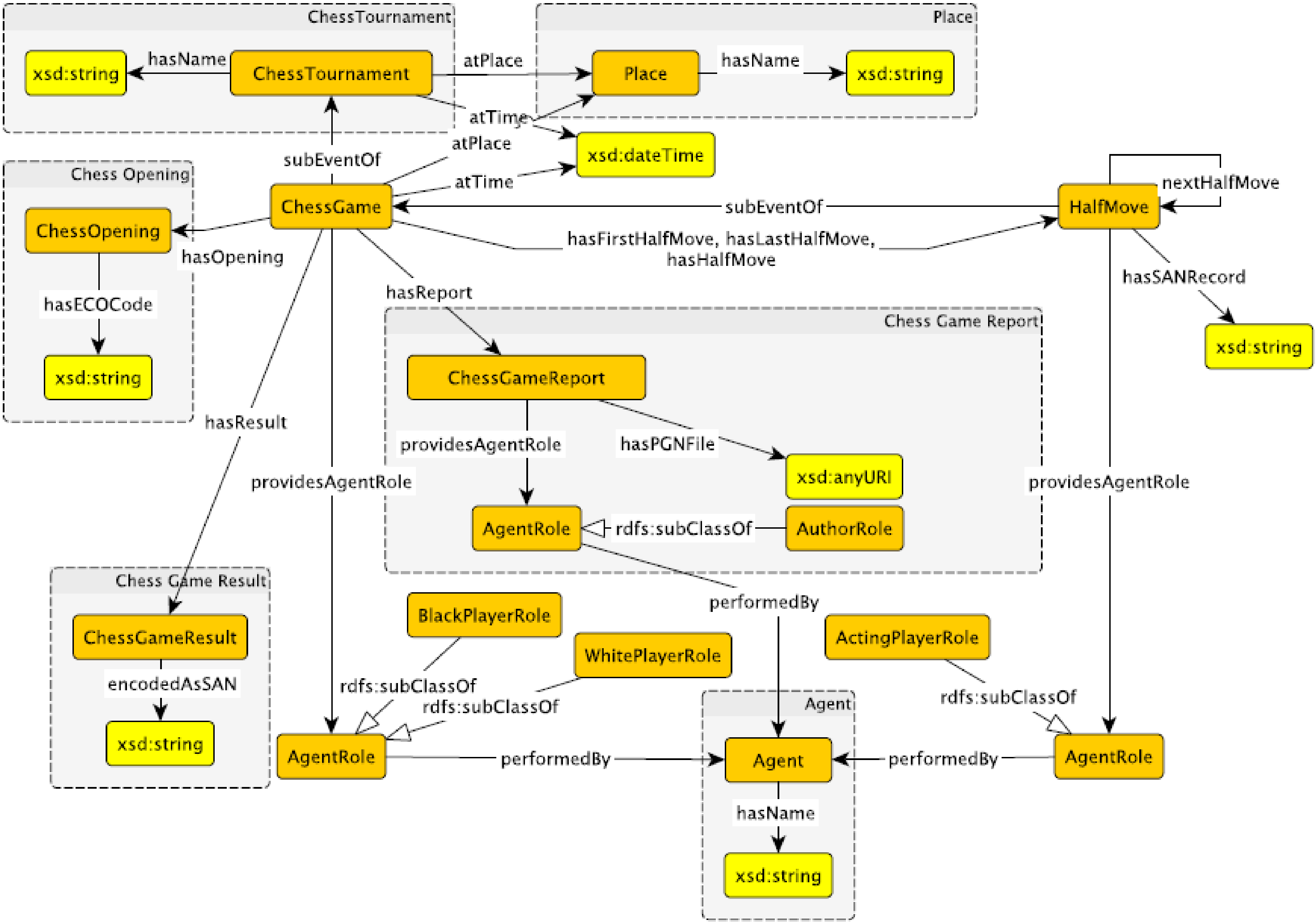
- **Inherited from Event or AgentRole**
- **Scoped domain/range restrictions, possibly with some cardinalities**
- **Basic existentials**
- **Non-cyclicity of half-move sequence**

**What about adding, e.g., the following?**

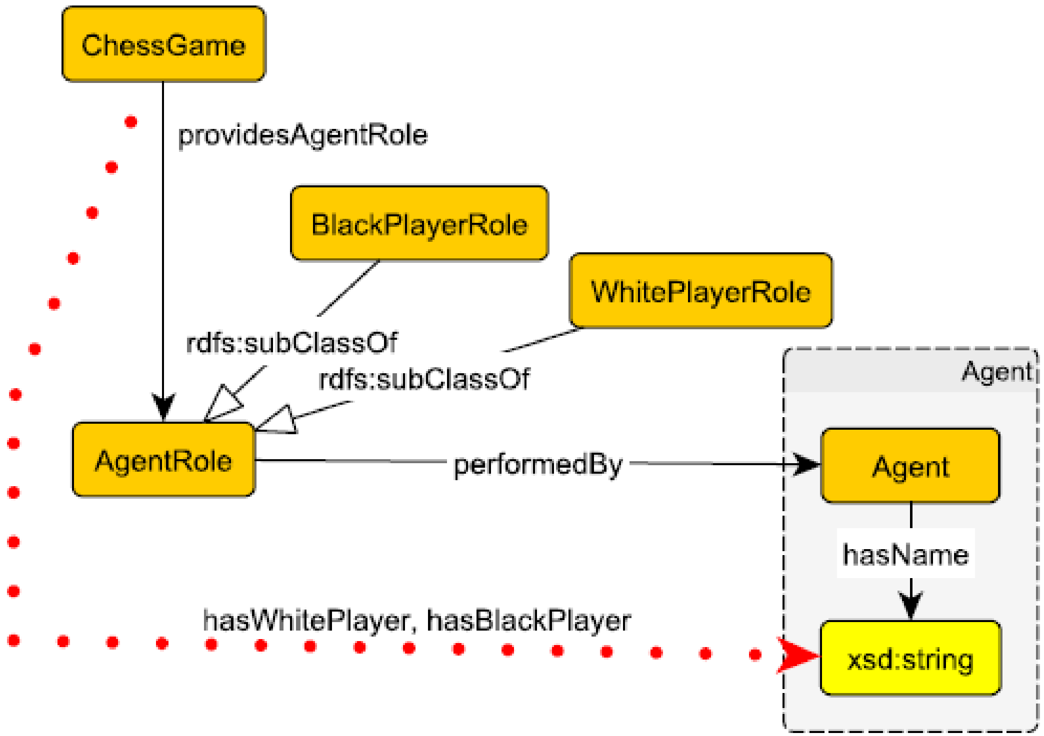
$$\mathsf{ChessGame} \sqsubseteq {\geq}0\mathsf{subEventOf}.\mathsf{ChessTournament}$$

**If one of the roles of axiomatization is to improve human understanding of the ontology, then such axioms are helpful!**

# Shortcuts and Views


ChessTournament
xsd:string
hasName
ChessTournament
atPlace
Place
Place
hasName
xsd:string
atTime
atPlace
xsd:dateTime
subEventOf
Chess Opening
ChessGame
atTime
nextHalfMove
subEventOf
HalfMove
ChessOpening
hasOpening
hasFirstHalfMove, hasLastHalfMove,
hasHalfMove
hasSANRecord
hasECOCode
hasReport
Chess Game Report
xsd:string
xsd:string
ChessGameReport
hasResult
providesAgentRole
hasPGNFile
providesAgentRole
xsd:anyURI
providesAgentRole
AgentRole
rdfs:subClassOf
AuthorRole
Chess Game Result
performedBy
BlackPlayerRole
ActingPlayerRole
ChessGameResult
WhitePlayerRole
encodedAsSAN
rdfs:subClassOf
rdfs:subClassOf
rdfs:subClassOf
Agent
xsd:string
AgentRole
performedBy
Agent
performedBy
AgentRole
hasName
xsd:string

# Shortcuts




$$\mathsf{ChessGame}(x) \land \mathsf{pAR}(x,y) \land \mathsf{WhitePlayerRole}(y) \land \mathsf{performedBy}(y,z) \land \mathsf{Agent}(z) \land \mathsf{hasName}(z,s) \to \mathsf{hasWhitePlayer}(x,s)$$

$$\mathsf{ChessGame}(x) \land \mathsf{pAR}(x,y) \land \mathsf{BlackPlayerRole}(y) \land \mathsf{performedBy}(y,z) \land \mathsf{Agent}(z) \land \mathsf{hasName}(z,s) \to \mathsf{hasBlackPlayer}(x,s)$$

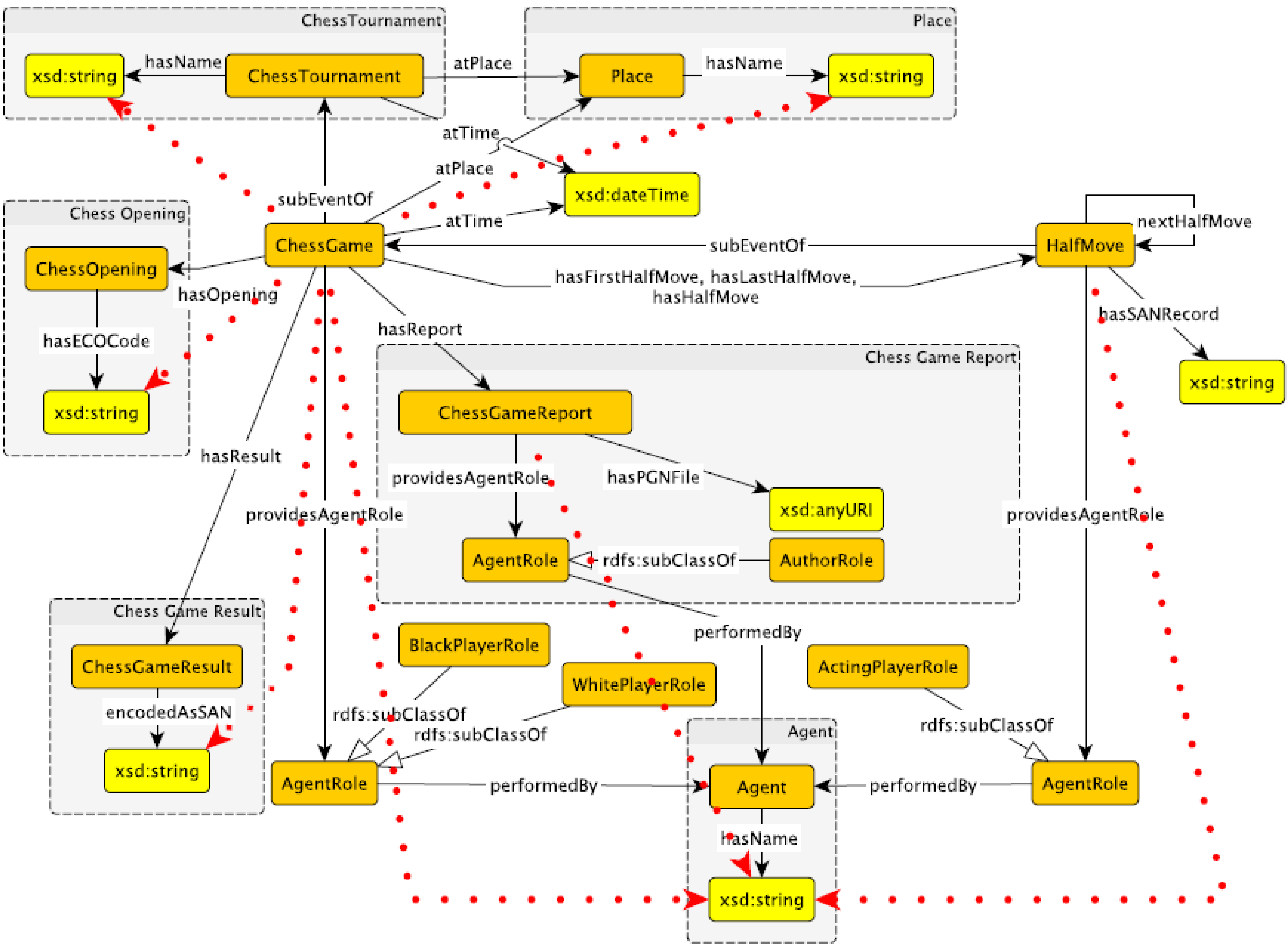
# Translating the rules

DaSe Lab

$$\mathsf{ChessGame}(x) \wedge \mathsf{pAR}(x,y) \wedge \mathsf{WhitePlayerRole}(y) \wedge \mathsf{performedBy}(y,z)$$
$$\wedge\, \mathsf{Agent}(z) \wedge \mathsf{hasName}(z,s) \rightarrow \mathsf{hasWhitePlayer}(x,s)$$
$$\mathsf{ChessGame}(x) \wedge \mathsf{pAR}(x,y) \wedge \mathsf{BlackPlayerRole}(y) \wedge \mathsf{performedBy}(y,z)$$
$$\wedge\, \mathsf{Agent}(z) \wedge \mathsf{hasName}(z,s) \rightarrow \mathsf{hasBlackPlayer}(x,s)$$

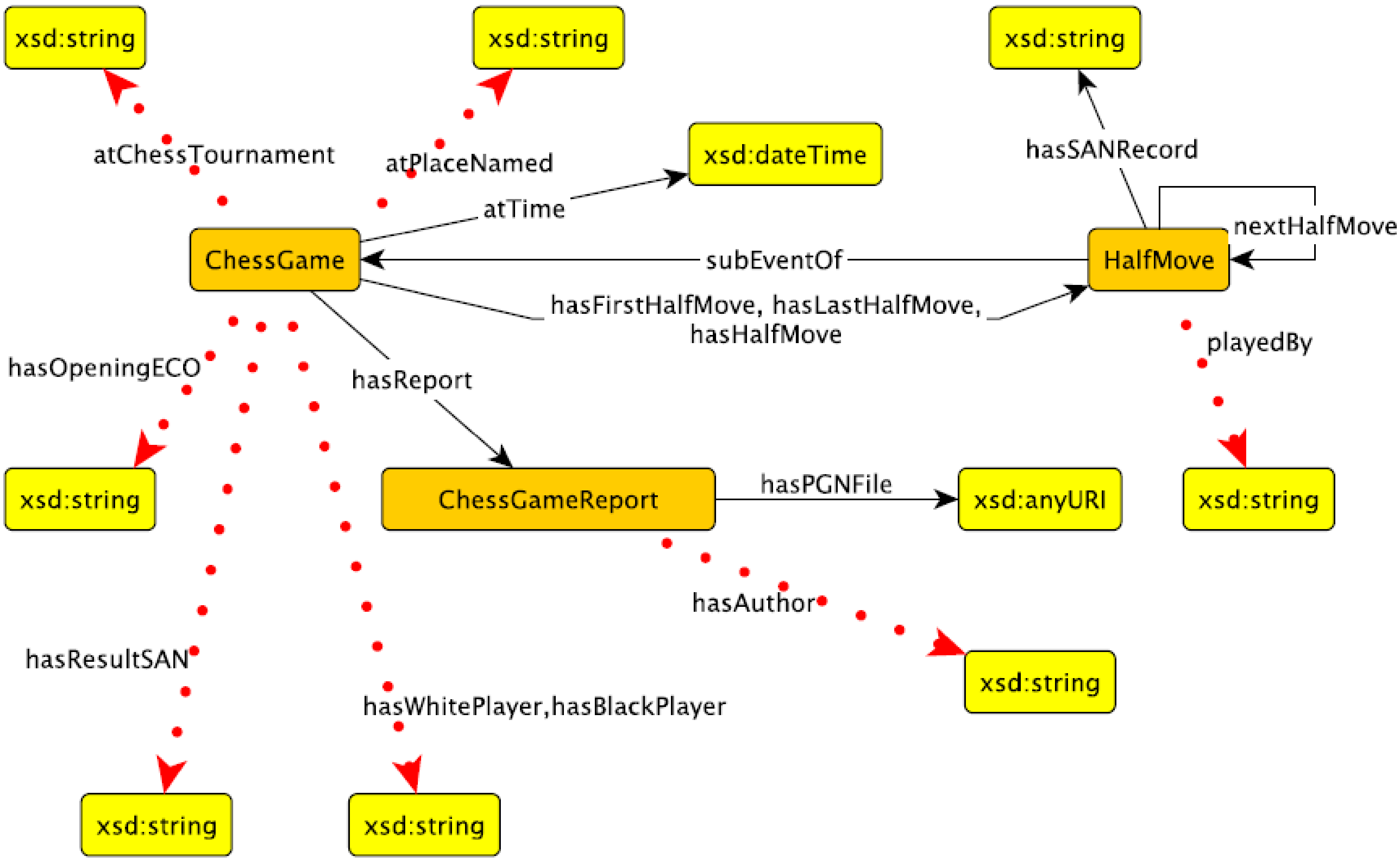
$$\mathsf{ChessGame} \sqsubseteq \exists R_1.\mathrm{Self}$$
$$\mathsf{WhitePlayerRole} \sqsubseteq \exists R_2.\mathrm{Self}$$
$$\mathsf{Agent} \sqsubseteq \exists R_3.\mathrm{Self}$$
$$R_1 \circ \mathsf{pAR} \circ R_2 \circ \mathsf{performedBy} \circ R_3 \circ \mathsf{hasName} \sqsubseteq \mathsf{hasWhitePlayer}$$

**However note that the introduction of additional role chains may cause violations of regularity restrictions.**


ChessTournament
xsd:string
hasName
ChessTournament
atPlace
Place
Place
hasName
xsd:string
atTime
atPlace
xsd:dateTime
subEventOf
Chess Opening
ChessGame
atTime
subEventOf
HalfMove
nextHalfMove
ChessOpening
hasOpening
hasFirstHalfMove, hasLastHalfMove,
hasHalfMove
hasSANRecord
hasECOCode
hasReport
xsd:string
Chess Game Report
xsd:string
ChessGameReport
hasResult
providesAgentRole
hasPGNFile
xsd:anyURI
providesAgentRole
providesAgentRole
AgentRole
rdfs:subClassOf
AuthorRole
Chess Game Result
performedBy
BlackPlayerRole
ActingPlayerRole
ChessGameResult
WhitePlayerRole
encodedAsSAN
rdfs:subClassOf
rdfs:subClassOf
rdfs:subClassOf
Agent
xsd:string
AgentRole
performedBy
Agent
performedBy
AgentRole
hasName
xsd:string

# Simplified View

DaSe Lab

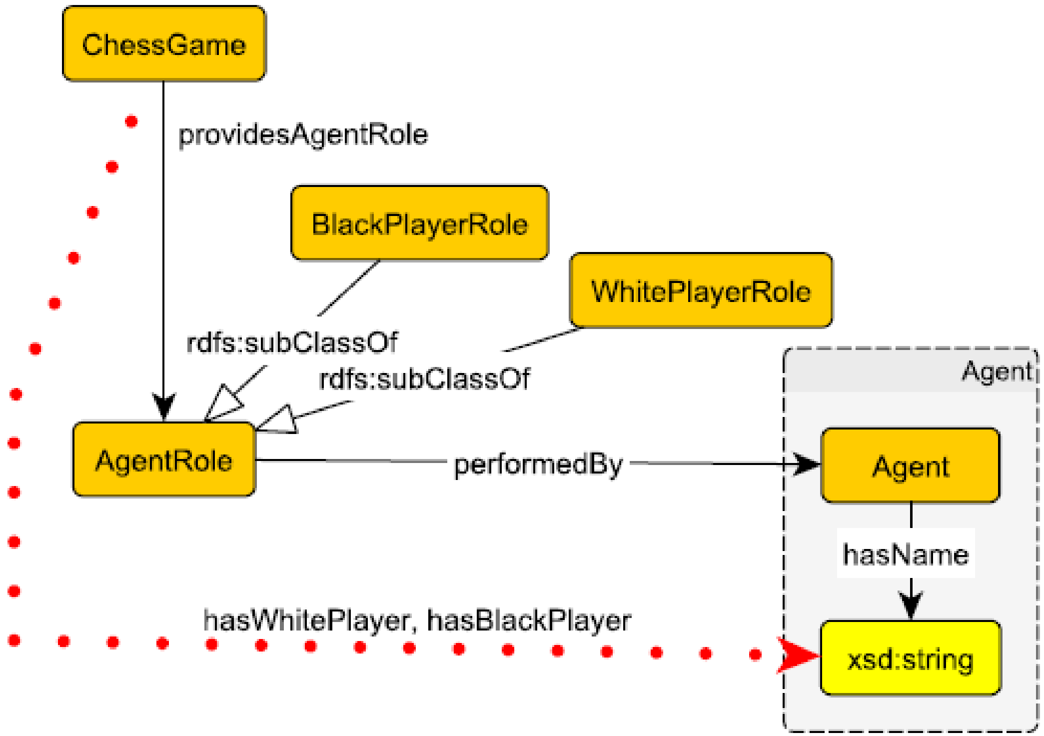



WRIGHT STATE UNIVERSITY

# Mapping from Views

**We used rules (axioms) to express the mapping from the ontology to the view.**

**The reverse direction is much more tricky.**




$$\text{ClassA}(x) \wedge \text{ClassB}(y) \wedge C_1(x_1) \wedge \cdots \wedge C_n(x_n) \wedge R_1(y_1, y_2) \wedge \cdots \wedge R_k(y_k, y_{k+1}) \to \text{shortcut}(x, y).$$

$$\text{shortcut}(x, y) \to \text{ClassA}(x) \wedge \text{ClassB}(y) \wedge \exists x_1 \ldots \exists x_n \exists y_1 \ldots \exists y_n (C_1(x_1) \wedge \ldots \cdots \wedge C_n(x_n) \wedge R_1(y_1, y_2) \wedge \cdots \wedge R_k(y_k, y_{k+1}))$$

# Mapping from views

**Existential rules may be suitable in principle.**

$$\text{shortcut}(x,y) \rightarrow \text{ClassA}(x) \wedge \text{ClassB}(y) \wedge \exists x_1 \dots \exists x_n \exists y_1 \dots \exists y_n (C_1(x_1) \wedge \dots$$
$$\dots \wedge C_n(x_n) \wedge R_1(y_1, y_2) \wedge \dots \wedge R_k(y_k, y_{k+1}))$$

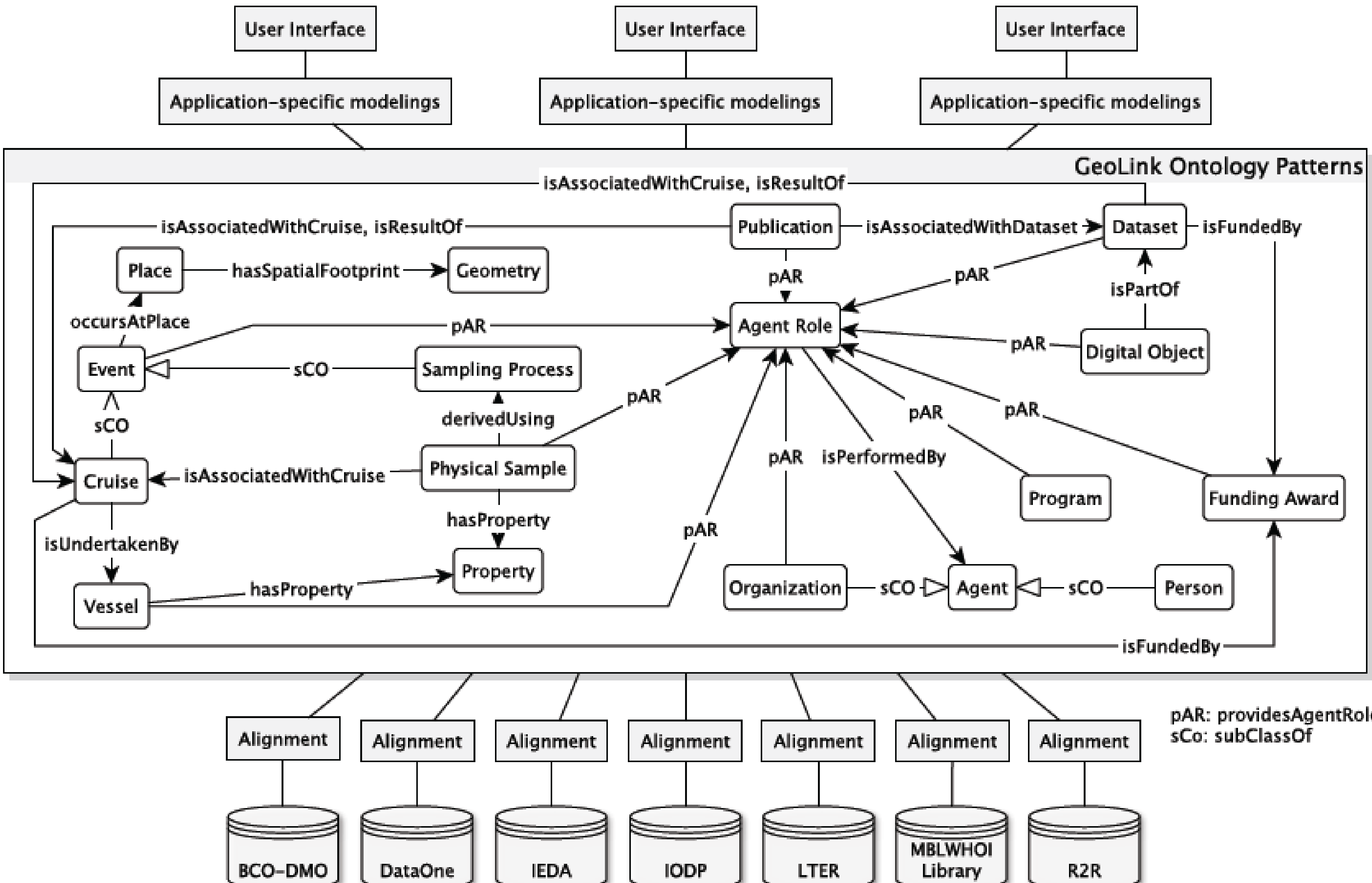
**However automated reasoning with the potentially rather complex rule heads requires investigations, in particular if it is to be integrated with ontology reasoning.**

**A specific case are right-hand-side role chains:**

$$R \sqsubseteq R_1 \circ \dots \circ R_n,$$

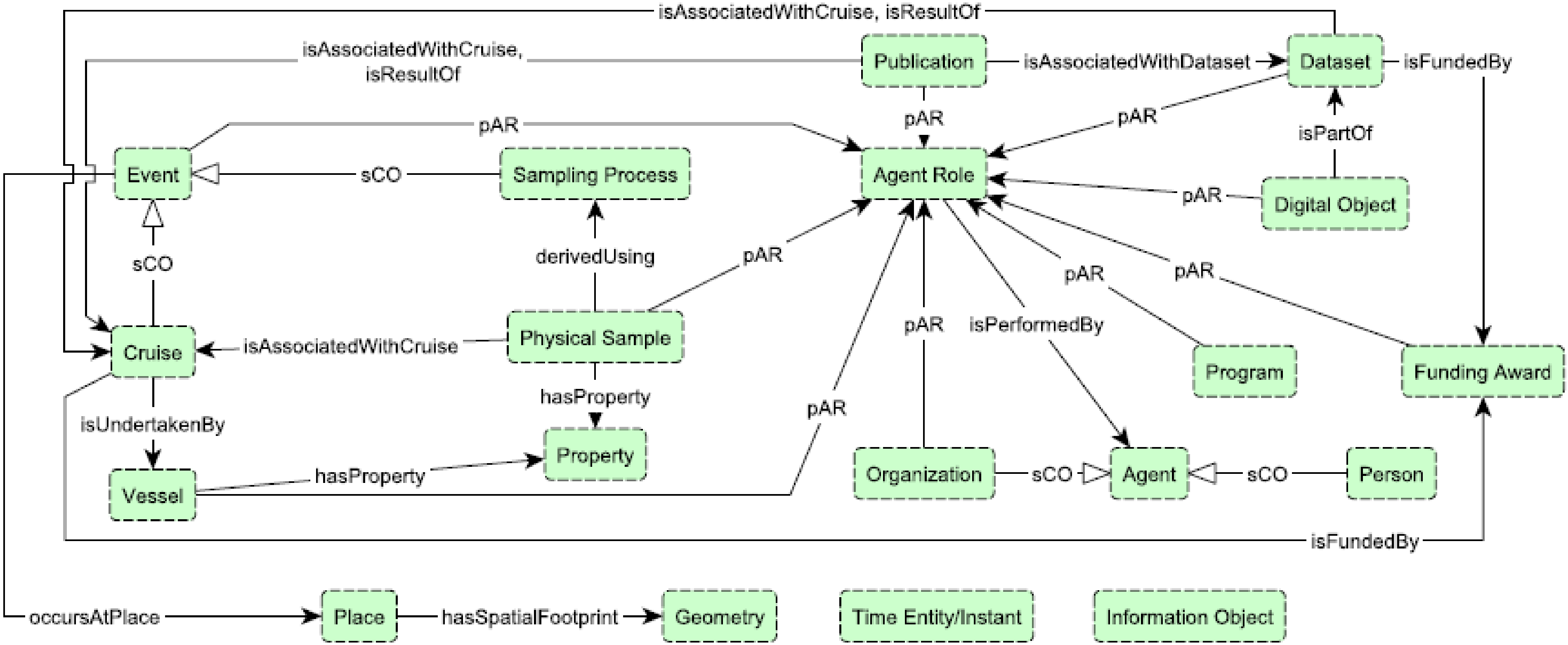
## The GeoLink Modular Ontology (GMO)

# Back to GeoLink




pAR: providesAgentRole
sCo: subClassOf

# The GeoLink Modular Ontology




**High-level overview of the GeoLink Modular Ontology (GMO).**

**Each box stands for a module, which has been modeled in its own right.**

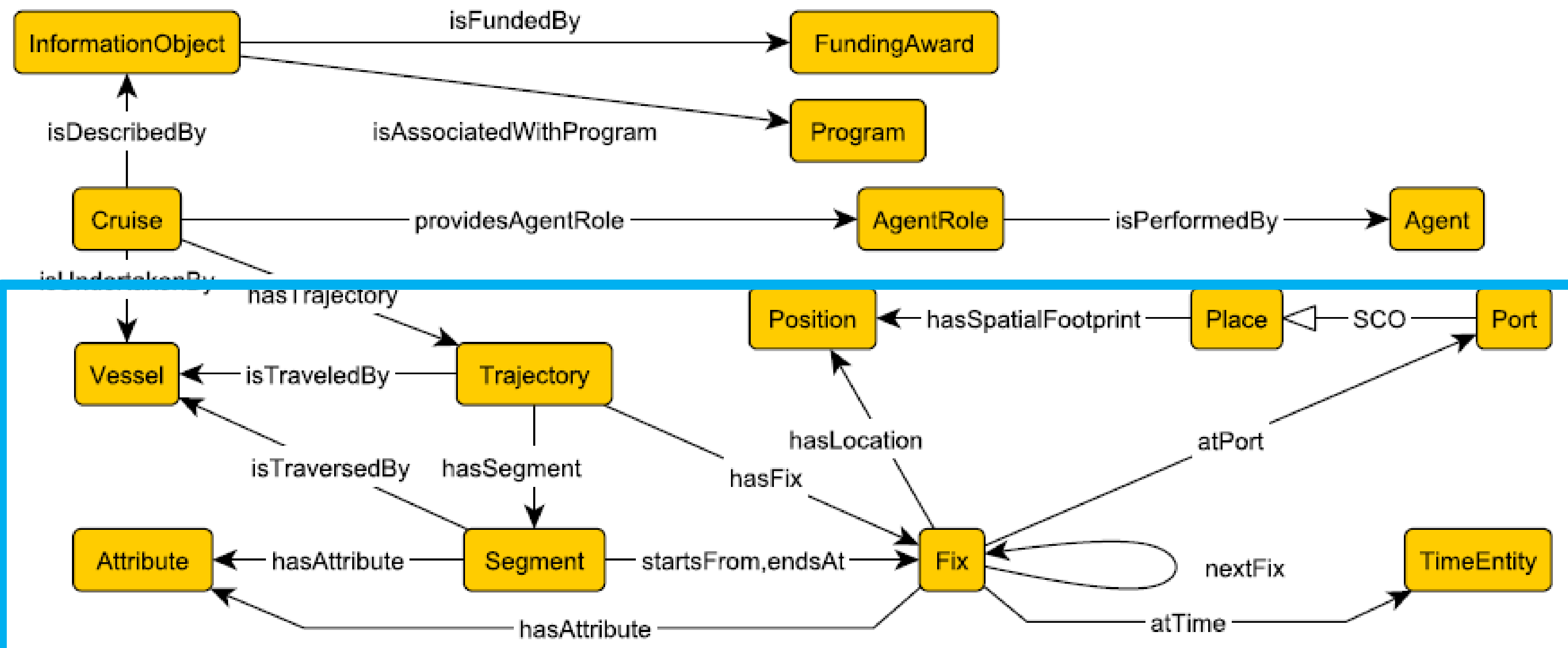
# Example Module: Cruise

**Cruise reused e.g. the generic patterns**

**AgentRole**

**Trajectory**

**and conceptually cruises are understood to be events.**



# GeoLink

An (preliminary) interactive demonstration of the integrated GeoLink data is available at

http://demo.geolink.org

At http://www.geolink.org/ there are links to the complete schema, a SPARQL Endpoint, publications, etc.

# I need to mention

**Several W3C recommendations are relevant in our context:**

- **The Web Ontology Language OWL for expressing ontologies and sharing them on the Web.**
- **The Resource Description Framework RDF for expressing e.g. ABox graphs and aligning them to an ontology.**
- **The SPARQL RDF query language.**

**(See the references for pointers.)**

# Thanks!

# References

Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnathi, Valentina Presutti (eds.), Ontology Engineering with Ontology Design Patterns: Foundations and Applications. Studies on the Semantic Web. IOS Press/AKA Verlag, 2016/2017. To appear.

Hitzler, Krötzsch, Rudolph, Foundations of Semantic Web Technologies, CRC/Chapman & Hall, 2010

Adila A. Krisnadhi, Yingjie Hu, Krzysztof Janowicz, Pascal Hitzler, Robert Arko, Suzanne Carbotte, Cynthia Chandler, Michelle Cheatham, Douglas Fils, Tim Finin, Peng Ji, Matthew Jones, Nazifa Karima, Kerstin Lehnert, Audrey Mickle, Tom Narock, Margaret O'Brien, Lisa Raymond, Adam Shepherd, Mark Schildhauer, Peter Wiebe, The GeoLink Framework for Pattern-based Linked Data Integration. In: Proceedings of the Posters and Demos session at the 14th International Semantic Web Conference (ISWC 2015), Bethlehem, Pensylvania, USA, October 2015.

# References

Adila A. Krisnadhi, Yingjie Hu, Krzysztof Janowicz, Pascal Hitzler, Robert Arko, Suzanne Carbotte, Cynthia Chandler, Michelle Cheatham, Douglas Fils, Tim Finin, Peng Ji, Matthew Jones, Nazifa Karima, Audrey Mickle, Tom Narock, Margaret O'Brien, Lisa Raymond, Adam Shepherd, Mark Schildhauer, Peter Wiebe, The GeoLink Modular Oceanography Ontology. In: Marcelo Arenas, Óscar Corcho, Elena Simperl, Markus Strohmaier, Mathieu d'Aquin, Kavitha Srinivas, Paul T. Groth, Michel Dumontier, Jeff Heflin, Krishnaprasad Thirunarayan, Steffen Staab (eds.), The Semantic Web - ISWC 2015 - 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part II. Lecture Notes in Computer Science 9367, Springer, Heidelberg, 2015, 301-309.

Adila Krisnadhi, Ontology Pattern-Based Data Integration. Dissertation, Department of Computer Science and Engineering, Wright State University, 2015.

# References


A. Gangemi. Ontology design patterns for semantic web content. In Y. Gil et al. (eds), The Semantic Web - ISWC 2005 – 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6-10, 2005, Proceedings, volume 3729 of Lecture Notes in Computer Science, pages 262-276. Springer, 2005.

Adila Krisnadhi, The Role Patterns. In: Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnathi, Valentina Presutti (eds.), Ontology Engineering with Ontology Design Patterns: Foundations and Applications. Studies on the Semantic Web. IOS Press/AKA Verlag, 2016/2017. To appear.

Víctor Rodríguez-Doncel, Adila A. Krisnadhi, Pascal Hitzler, Michelle Cheatham, Nazifa Karima, Reihaneh Amini, Pattern-Based Linked Data Publication: The Linked Chess Dataset Case. In: Olaf Hartig, Juan Sequeda, Aidan Hogan (eds.), Proceedings of the 6th International Workshop on Consuming Linked Data co-located with 14th International Semantic Web Conference (ISWC 2105), Bethlehem, Pennsylvania, US, October 12th, 2015. CEUR Workshop Proceedings 1426, CEUR-WS.org, 2015.

# References

Adila Krisnadhi, Pascal Hitzler, Modeling With Ontology Design Patterns: Chess Games As a Worked Example. In: Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnathi, Valentina Presutti (eds.), Ontology Engineering with Ontology Design Patterns: Foundations and Applications. Studies on the Semantic Web. IOS Press/AKA Verlag, 2016/2017. To appear.

Adila Krisnadhi, Nazifa Karima, Pascal Hitzler, Reihaneh Amini, Victor Rodriguez-Doncel, Krzysztof Janowicz, Ontology Design Patterns for Linked Data Publishing. In: Pascal Hitzler, Aldo Gangemi, Krzysztof Janowicz, Adila Krisnathi, Valentina Presutti (eds.), Ontology Engineering with Ontology Design Patterns: Foundations and Applications. Studies on the Semantic Web. IOS Press/AKA Verlag, 2016/2017. To appear.

# References

**Adila A. Krisnadhi, Pascal Hitzler, Krzysztof Janowicz, On capabilities and limitations of OWL regarding typecasting and ontology design pattern views. In: V. Tamma, M. Dragoni, R. Goncalves, A. Lawrynowicz (eds.), Ontology Engineering. 12th International Experiences and Directions Workshop on OWL, OWLED 2015, co-located with ISWC 2015, Bethlehem, PA, USA, October 9-10, 2015, Revised Selected Papers. Lecture Notes in Computer Science 9557, Springer, Heidelberg, 2016, pp. 105-116.**

**Richard Cyganiak, David Wood, Markus Lanthaler, RDF 1.1 Concepts and Abstract syntax. W3C Recommendation 25 February 2014.**

**Pascal Hitzler, Markus Krötzsch, Bijan Parsia, Peter F. Patel-Schneider, Sebastian Rudolph, OWL 2 Web Ontology Language: Primer (Second Edition). W3C Recommendation, 11 December 2012.**

**Steve Harris, Andy Seaborne, SPARQL 1.1 Query Language. W3C Recommendation 21 March 2013.**