Ontology Design Patterns for Data Repository Integration

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

Krzysztof Janowicz
STKO Lab
UC Santa Barbara
http://stko.geog.ucsb.edu/
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.”
EarthCube requires

- information integration
- interoperability
- conceptual modeling
- intelligent search
- data-model intercomparison
- data publishing support

Semantic Web studies

- information integration
- interoperability
- conceptual modeling
- intelligent search
- data-model intercomparison
- data publishing support

Pascal Hitzler, WSU; Krzysztof Janowicz, UCSB
EarthCube Challenges

The EarthCube “Architecture” must be

- modular
- extensible
- sustainable
- sliceable (i.e. you can adopt part of it without adopting all)
- simple enough for easy adoption
- complex enough to solve real problems
- scalable in terms of breadth of topic coverage
- elastic, in that it allows partners to decide how much they want to share
- respectful of individual modeling choices
EarthCube GeoLink project

We show that our approach meets the mentioned EarthCube challenges.

LDEO: Robert Arko, Suzanne Carbotte, Kerstin Lehnert
WHOI: Cynthia Chandler, Peter Wiebe, Lisa Raymond, Adam Shepherd
UCSB: Mark Schildhauer, Krzysztof Janowicz, Matt Jones, Yingjie Hu
Ocean Leadership: Douglas Fils
Marymount Univ: Thomas Narock
WSU: Pascal Hitzler, Michelle Cheatham, Adila Krisnadhi
UMBC: Tim Finin
GeoLink setup

User Interface

additional application-specific modeling

GeoLink Patterns

mappings

R2R  BCO-DMO  IEDA  LTER  IODP  ...
“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.
Age – very simplistic

Specific, straightforward data modeling is very restrictive regarding data integration and reuse.
Age – still very simplistic

Specific, straightforward data modeling is very restrictive regarding data integration and reuse.
Towards Age

- Thing
  - hasReported Age
    - reportedBy
    - reportedAge
- Referent
- AgeDescription
Towards Age

- Period
  - hasName: xsd:string
  - hasPeriod

AgeDescription
Towards Age

- **Period**: hasName (xsd:string)

  - **AgeDescription**: hasPeriod
  - **hasTimeInterval**: startsAt
  - **endsAt**: hasTemporalValue

  - **Instant**: owltime: Instant
    - hasTemporalValue
Towards Age

AgeDescription

Period

hasPeriod

hasName

xsd:string

Confidence

hasTimeInterval

startsAt

endsAt

hasTemporalValue

owltime:Instant

hasTemporalValue
Oceanographic Cruise

Diagram:
- RepositoryObject
- Event
  - rdfs:subClassOf
  - originatesFrom
- CruiseInformationObject
  - isDescribedBy
- Cruise
  - isUndertakenBy
- Vessel
  - isTraversedBy
- Trajectory
  - hasSegment
  - hasTrajectory

August 2014 – EarthCube C4P – Webinar – Pascal Hitzler
Roles (Cruise as Event)
Cruise Trajectories

- Cruise
  - hasTrajectory
  - isUndertakenBy Vessel
  - isTraversedBy Segment

- Trajectory
  - hasSegment
  - hasFix
  - startsFrom Fix
  - endsAt Fix

- Fix
  - hasLocation
  - hasSpatialFootprint

- Vessel
  - rdf:type port_stop_arrival

- Segment
  - hasAttribute
  - rdf:type port_stop_departure

- Attribute
  - rdf:type
  - time:TemporalEntity

- Place
  - nextFix
  - rdfs:subClassOf Port
Information Objects

Cruise isDescribedBy CruiseInformationObject

hasRelatedCruiseID, hasCanonicalName, hasDescription

hasWebpage isFundedBy isAssociatedWith

FundingAward Program

rdf:PlainLiteral xsd:anyURI

hasCruiseType

CruiseType

rdf:type

operational maintenance transit

other_cruisetype
Patterns as interchange format

- Aggregated data can be “pulled back” along the same mappings, if desired.
- Since the patterns are very generic, there is no loss of information by using them as interchange format.

GeoLink Patterns

mappings

R2R  BCO-DMO  IEDA  LTER  IODP  ...

EarthCube Challenges

The EarthCube “Architecture” must be

- modular
- extensible
- sustainable
- sliceable (i.e. you can adopt part of it without adopting all)
- simple enough for easy adoption
- complex enough to solve real problems
- scalable in terms of breadth of topic coverage
- elastic, in that it allows partners to decide how much they want to share
- respectful of individual modeling choices
Thanks!

www.oceanlink.org
www.geo-link.org
References

References

References
