Semantic Interoperability

Jérôme Euzenat
INRIA & LIG
France

Natasha Noy
Stanford University
USA

Being serious about the Semantic Web

• It is not one person’s ontology
• It is not several people’s common ontology
• It is many people’s ontologies
• So it is a mess, but a meaningful mess
Heterogeneous Ontologies: Example

SessionEvent
  hasLocation: Place
  hasTimeAndDate: Date
  hasAttendee: Person

DemoSession

PosterSession

PaperSession

Session
  location: String
  time: TimeAndDate
  attendees: Person
  sessionType: (Presentations, Demos, Panel, Keynote)

Person
  name: String
  email: String

TimeAndDate
  time: String
  date: String

Person
  name: String
  email: String

Ontology alignment at a glance

Parameters

Alignment tool

External resources

New alignment between Ontology 1 and Ontology 2

Prior alignment between Ontology 1 and Ontology 2

Ontology 1

Ontology 2
Why should we learn to deal with this?

- Applications of semantic integration
  - Catalogue integration
  - Schema and data integration
  - Query answering
  - Peer-to-peer information sharing
  - Web service composition
  - Agent communication
  - Data transformation
  - Ontology evolution

Application: Catalogue integration

[Diagram showing the process of matching schemas and generating transformations between databases for Catalog 1 and Catalog 2]
Application: Query answering

Ontology 1 \rightarrow \text{match} \rightarrow \text{generate} \rightarrow \text{mediator} \rightarrow \text{answer}

Ontology 2

Server 1 \rightarrow \text{query} \rightarrow \text{reformulated query} \rightarrow \text{match} \rightarrow \text{generate} \rightarrow \text{mediator} \rightarrow \text{answer}

Server 2

Alignment between Ontology 1 and Ontology 2

Application: agent communication

Ontology 1 \rightarrow \text{match} \rightarrow \text{generate} \rightarrow \text{translator} \rightarrow \text{translated message}

Ontology 2

Agent 1 \rightarrow \text{message} \rightarrow \text{match} \rightarrow \text{generate} \rightarrow \text{translator} \rightarrow \text{translated message}

Agent 2

Alignment between Ontology 1 and Ontology 2

Axioms linking Ontology 1 and Ontology 2
Why is semantic interoperability difficult?
Why is semantic interoperability difficult?

SessionEvent
hasLocation: Place
hasTimeAndDate: Date
hasAttendee: Person

Person
name: String
e-mail: String

DemoSession
PosterSession

PaperSession

Session
location: String
time: TimeAndDate
attendees: Person

time: TimeAndDate

Session
location: String
time: TimeAndDate
attendees: Person

Session
location: String
time: TimeAndDate
attendees: Person

Session
location: String
time: TimeAndDate
attendees: Person

Session
location: String
time: TimeAndDate
attendees: Person

Why is semantic interoperability difficult?
Possible mismatches

- Different context (databases, ontologies) and different logics
- Same concept, different names
- Same name, different concepts
- Different approaches to conceptualization (e.g., subclasses versus property values)
- Different levels of granularity
- Different, but overlapping, areas

How can we address the problem?

- Names of entities
  - Comments, alternate names, names of related entities
- Structure
  - Internal structure: constraints on relations, types
  - External structure: relations between entities
- Extensions
  - Instances themselves
  - Related resources: annotated documents, exchanged message or queries
- Semantics (models)
- Background knowledge
  - The Web
  - Ontologies
  - Thesauri, e.g. WordNet
Name similarity

SessionEvent
   hasLocation: Place
   hasTimeAndDate: Date
   hasAttendee: Person

DemoSession

PosterSession

PaperSession

Person
   name: String
   email: String

Session
   location: String
   time: TimeAndDate
   attendees: Person
   sessionType: (Presentations, Demos, Panel, Keynote)

TimeAndDate
   time: String
   date: String

Person
   name: String
   email: String

Similar names

Similarity in structure

SessionEvent
   hasLocation: Place
   hasTimeAndDate: Date
   hasAttendee: Person

DemoSession

PosterSession

PaperSession

Person
   name: String
   email: String

Session
   location: String
   time: TimeAndDate
   attendees: Person
   sessionType: (Presentations, Demos, Panel, Keynote)

TimeAndDate
   time: String
   date: String

Person
   name: String
   email: String

Similar property name and range (structure)
Instance similarity

Common set of instances or documents

External sources

• A common reference ontology
• User input
• Lexicons, thesauri, etc.
• Prior matches
• Background knowledge (other ontologies, documents, etc.)
Combining different techniques

- Using several matchers in sequence (composing)
- Using several matchers in parallel (combining)
- Aggregating matcher results
  - aggregating specialised matcher results
  - aggregating competing matcher results
- Filtering results (trimming)
- Extracting alignment (optimizing)
- Iterating
- Learning
How well do these approaches work?

- Ontology Alignment Evaluation Initiative
  - Formal comparative evaluation of different ontology-matching tools
  - Run every year
  - Variety of test cases (in size, in formalism, in content)
  - Results very dependent on the tasks and the data (from under 50% of precision and recall to well over 80% if ontologies are relatively similar)
  - Results consistent across test cases
  - Progress every year!

Compared OAEI Results
Tools you should be aware of

- **Frameworks**
  - PROMPT (a Protégé plug-in): includes a user interface and a plug-in architecture
  - Alignment API: used by many tools in OAEI provides an exchange format and evaluation tools
  - COMA++: oriented toward database integration (many basic algorithms implemented).

- **Matching systems**
  - OAEI best performers (Falcon, RiMOM, etc.)
  - Available systems (FOAM, OLA, Rondo, etc.)
  - …

Current challenges: what to look for in conference papers

- How do we help users perform the alignments interactively?
- How do we explain the alignments that the tools create?
- How do we have system working across all cases? Do we need to?
- Can we use imperfect or inconsistent alignments?
- How do we maintain the alignments when ontologies evolve?
Current challenges (cont’d)

• Design space of alignment approaches
  – Can we create a “toolbox for designing alignment approaches that fit a given problem?
  – We have identified some components, but how can we bring them together?

• Have we discovered a “ceiling” in automatic discovery of alignments?
  – Will it be “lots of work for little gain” from now on?
  – Are there serious untapped resources?

Further reading

• “Ontology Matching” by Euzenat and Shvaiko
• Proceedings of ISWC, ASWC, ESWC, WWW conferences, etc.
• Journal of web semantics, Journal on data semantics, etc.
• http://www.ontologymatching.org