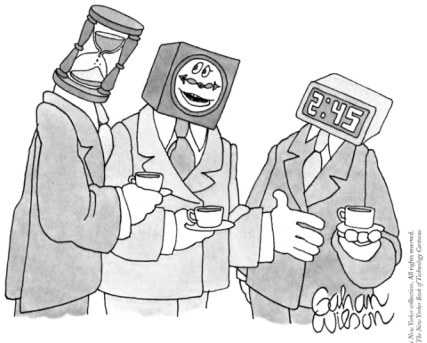
Semantic Interoperability

Jérôme Euzenat
INRIA & LIG
France

Natasha Noy
Stanford University
USA

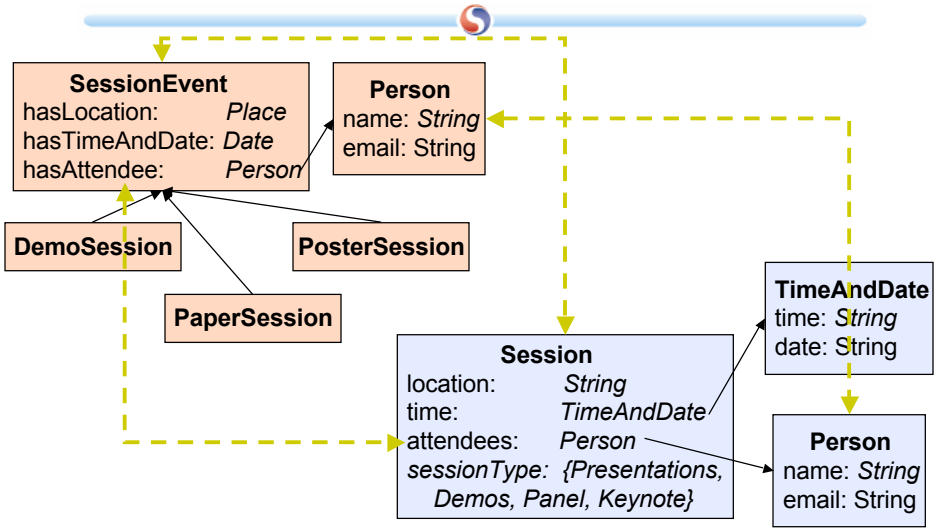
Being serious about the Semantic Web

- 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

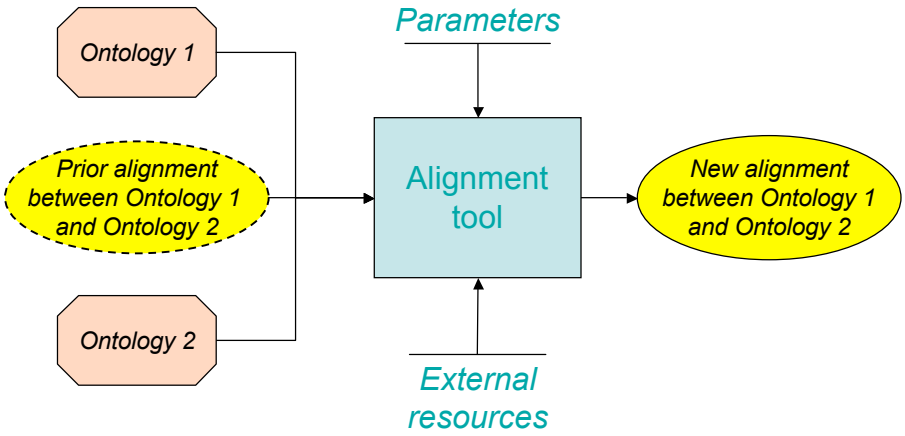


"Basically, we're all trying to say the same thing."

Heterogeneous Ontologies: Example



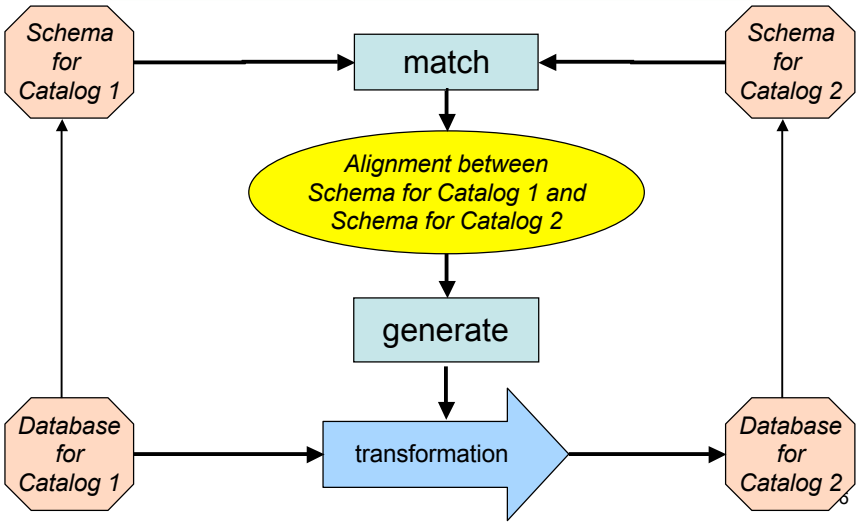
Ontology alignment at a glance



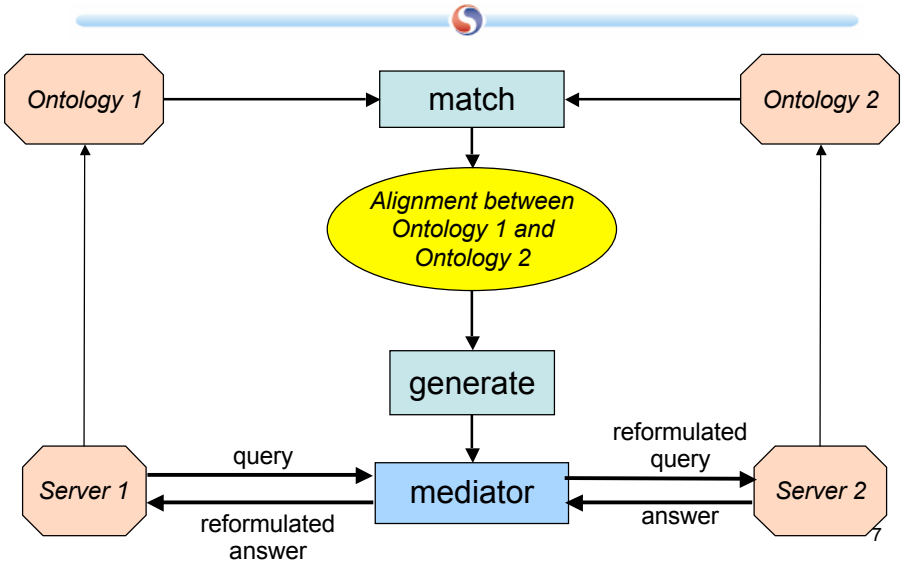
Why should we learn to deal with this?

- 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

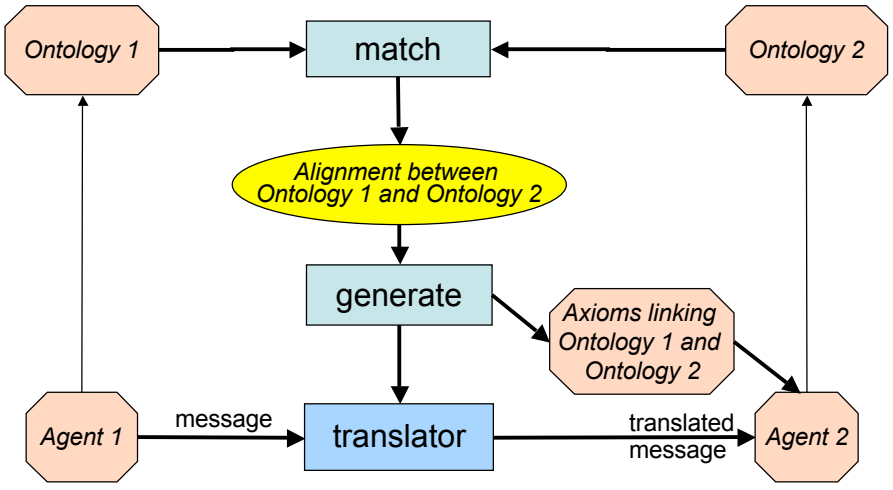
Application: Catalogue integration



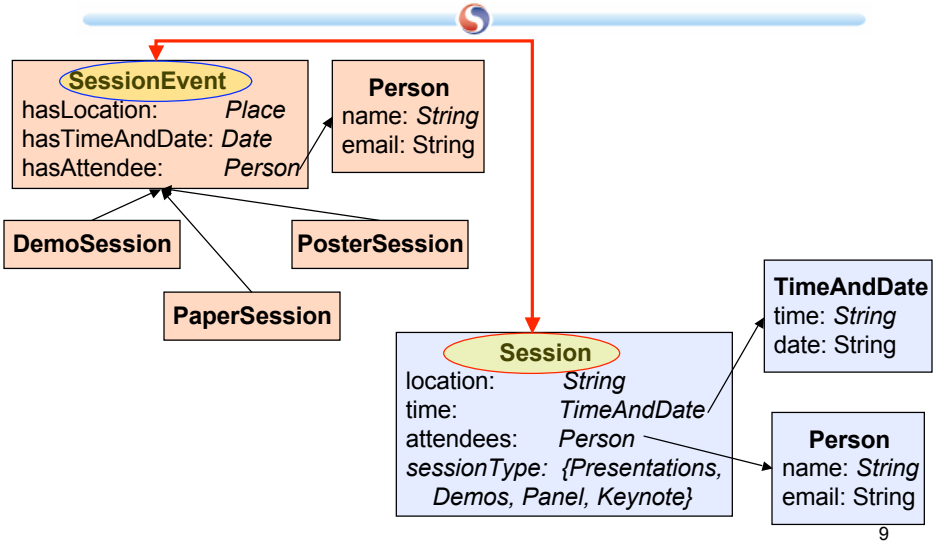
Application: Query answering



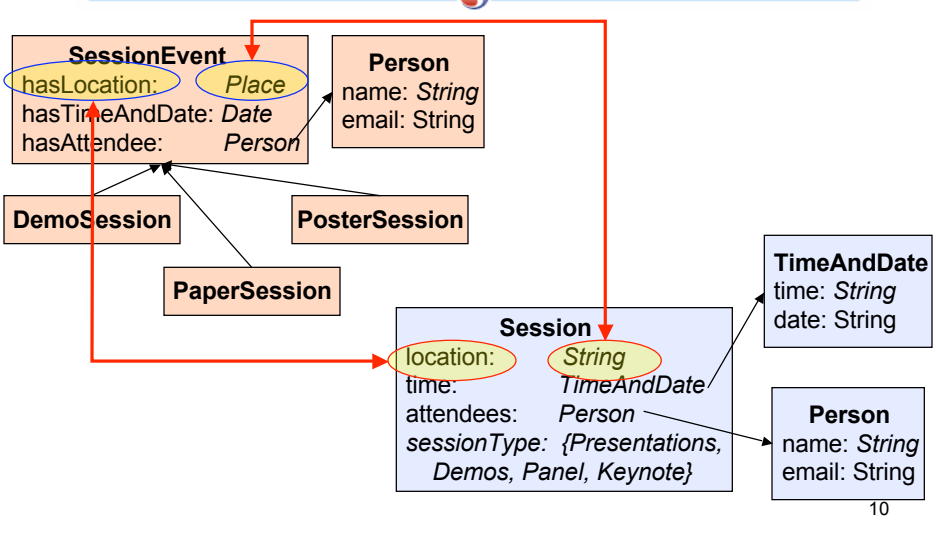
Application: agent communication



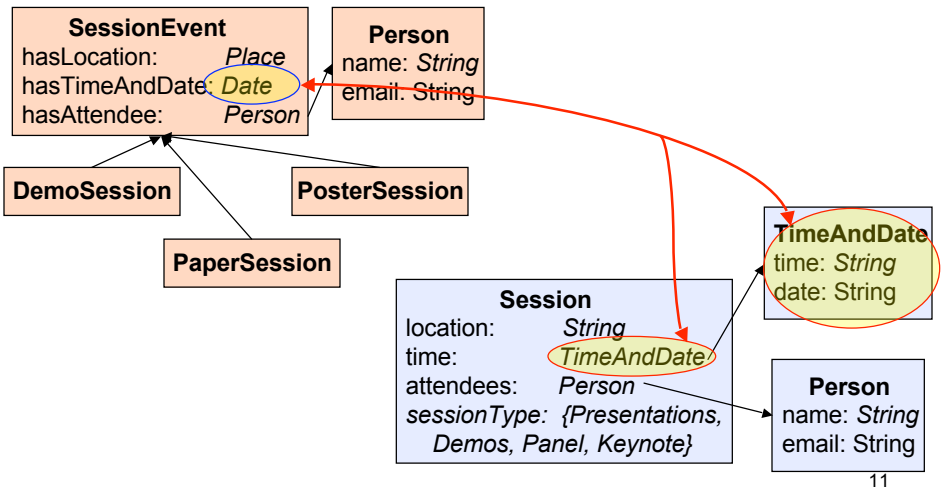
Why is semantic interoperability difficult?



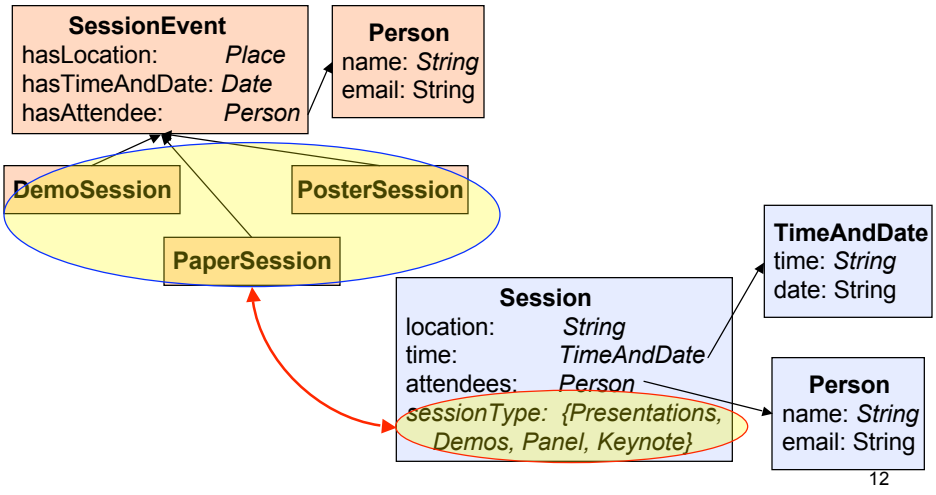
Why is semantic interoperability difficult?



Why is semantic interoperability difficult?



Why is semantic interoperability difficult?



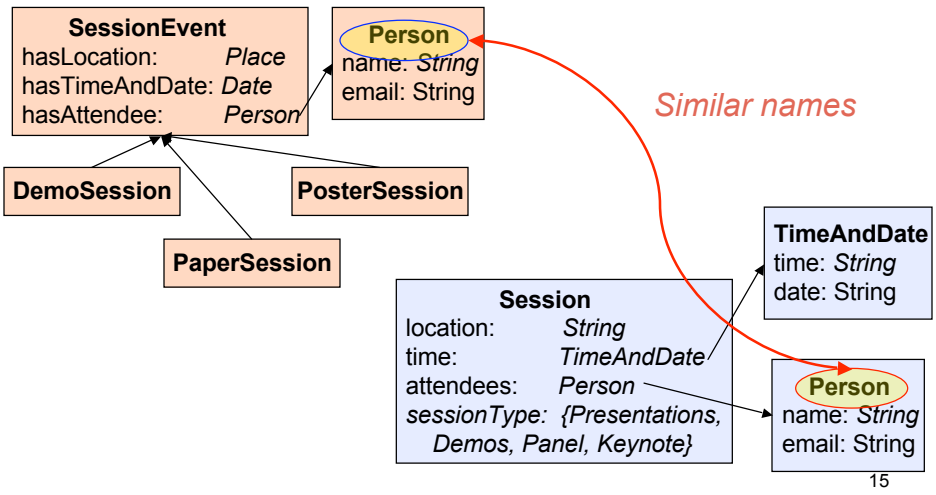
Possible mismatches

- 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

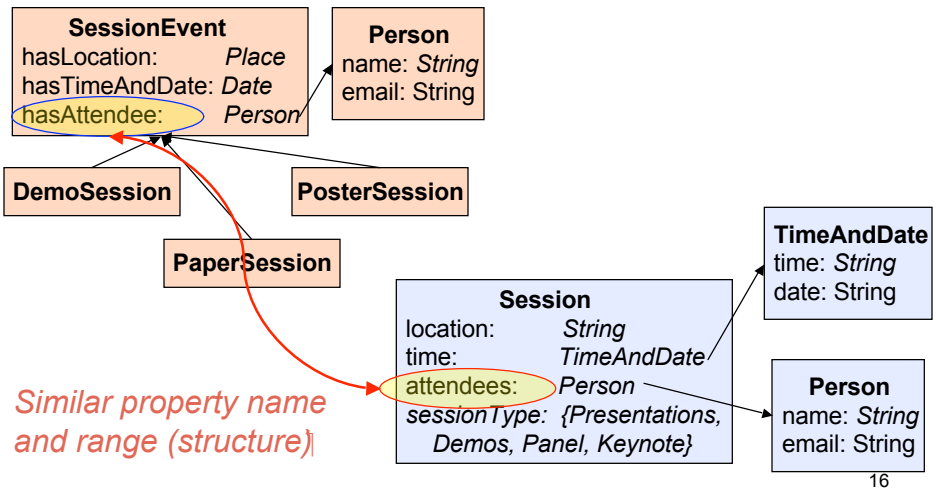
How can we address the problem?

- 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

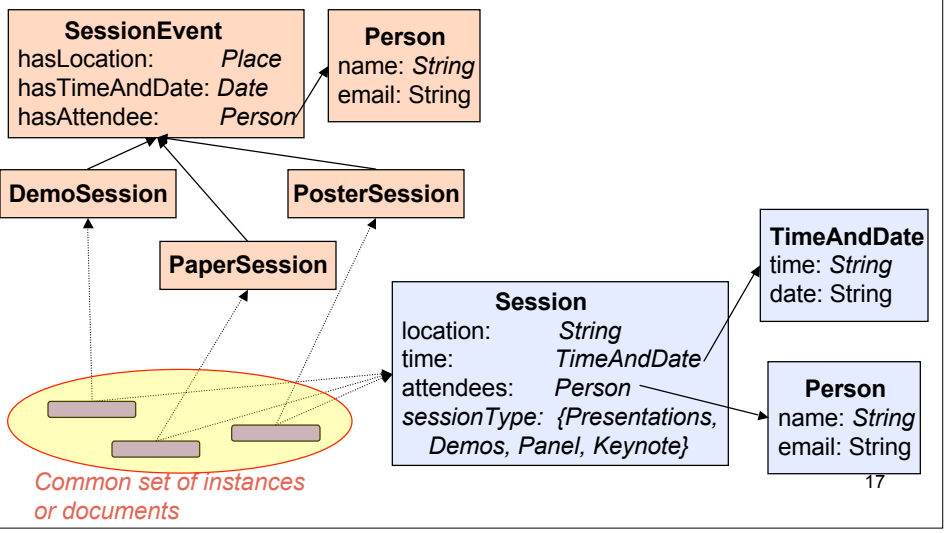
Name similarity



Similarity in structure



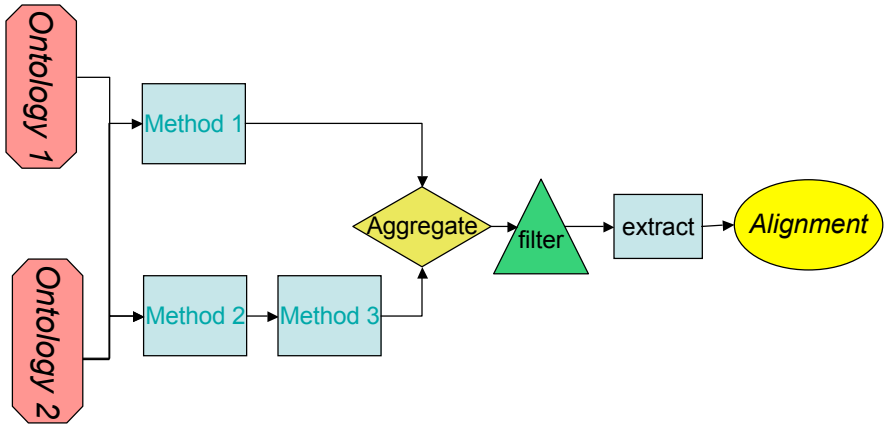
Instance similarity



External sources

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

Combining different techniques



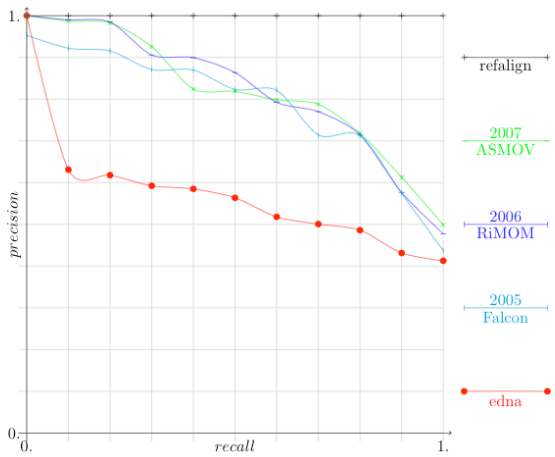
Combining different techniques

- 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