Ontologies for Designing Database-based systems

María del Mar Roldán, Ismael Navas

KHAOS Research

University of Málaga, Spain

Khaos INVESTIGACIÓN

LENGUAJES Y CIENCIAS DE LA COMPUTACIÓN

GISUM

Outline

- Introduction
- Techniques
  - Designing databases from ontologies
  - Generating database schema from ontologies
  - Generating ontologies from database schema
  - Exposing databases as RDF graphs
- Demos

Khaos

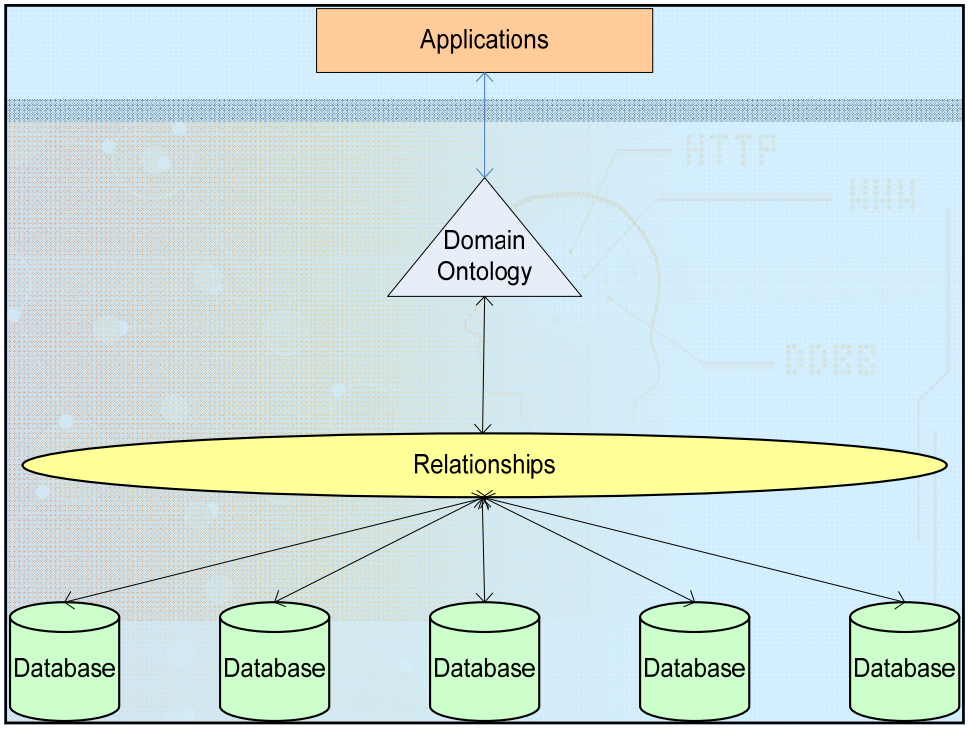
Introduction

- The exponential growth of accessible information on the Web
- Development of distributed systems that make use of common schemas for sharing knowledge.
- This schemas (ONTOLOGIES) represent the conceptualization shared among organizations that work over the same domain
- Many projects are developing tools for creation, maintenance and use of ontologies.
- No tools for importing and using ontologies as starting point in the design of database-based systems, adding semantics to them.

Khaos

Ontologies

- An ontology define concepts and the relationships between them in a formal way.
- Ontologies provide a structured vocabulary as well as a set of valid rules/axioms supported by the set of data sources that commit with it
- The semantics and meaning of every vocabulary term must be well defined to enable communication and the exchange of knowledge across the different organizations.



Khaos

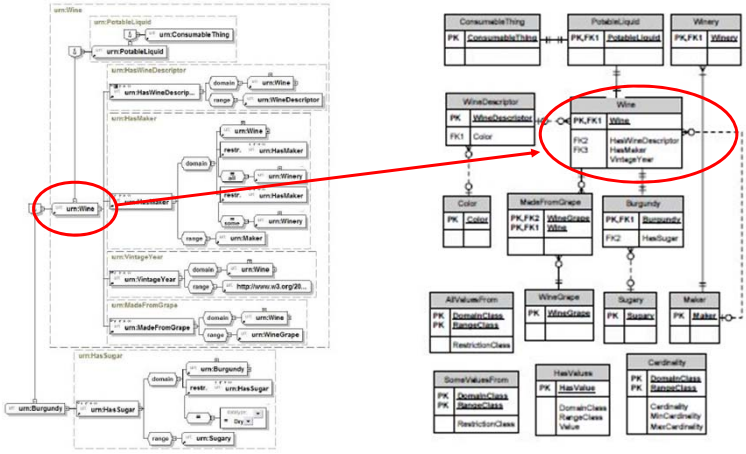
Techniques

- Designing the database from ontologies
- Generating the database schema from an ontology
- Generating an ontology from the database schema
  - Schema quality
- From an ontology and a database generating a SPARQL endpoint
  - Querying relational data using SPARQL
  - Generating RDF graph
  - Reasoning

Khaos

Generating the DB Schema from an Ontology

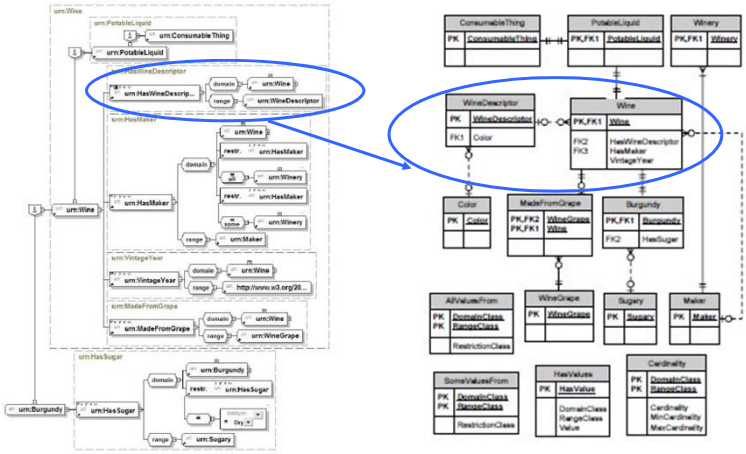
- Vysniauskas, 2006



Khaos

Generating the DB Schema from an Ontology

- Vysniauskas, 2006



Khaos

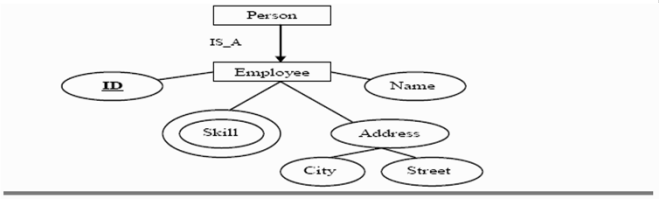
Generating an ontology from the database schema

- Analyzing the relational schema:
  - L.Stojanovic, et al. Migrating data-intensive web sites into the semantic web. In SAC'02.
  - Damir Juric, Dejing Dou. Building owl ontologies by analyzing relational database schema concepts and wordnet semantic relations. In The 9th International Conference on Telecommunications. ConTEL 2007
  - Christian Bizer. D2r map- a database to rdf mapping language. In 12 th Intl World Wide Web Conf, 2003. → **D2R MAP**
  - Lyna Lubyte. Techniques and Tools for the De-sign of Ontologies for Data Access → **DLR-DB**
  - Guoqiang Zhang. Ontology-based knowledge extraction for relational database schema. In: Electronic Commerce and Security, 2009

Khaos

Generating an ontology from the database schema

- Analyzing the relational schema (Tools):
  - **OntoExtract**
  - **ICOM**
  - **ER2OWL**
  - **ERONTO**




```
<owl:Class rdf:ID="Person"/>
 <owl:Class rdf:ID="Employee">
  <rdfs:subClassOf rdf:resource="#Person"/>
</owl:Class>
<owl:InverseFunctionalProperty rdf:ID="ID">
  <rdf:type rdf:resource=" #FunctionalProperty"/>
  <rdfs:range rdf:resource=" #int"/>
  <rdfs:domain rdf:resource="#Employee"/>
  <rdf:type rdf:resource="#DatatypeProperty"/>
</owl:InverseFunctionalProperty>
```

```
<owl:DatatypeProperty rdf:ID="Skill">
  <rdfs:range rdf:resource="#string"/>
  <rdfs:domain rdf:resource="#Employee"/>
</owl:DatatypeProperty>
<owl:FunctionalProperty rdf:ID="Name">
  <rdfs:range rdf:resource="#string"/>
  <rdfs:domain rdf:resource="#Employee"/>
  <rdf:type rdf:resource="#DatatypeProperty"/>
</owl:FunctionalProperty>
```

Khaos

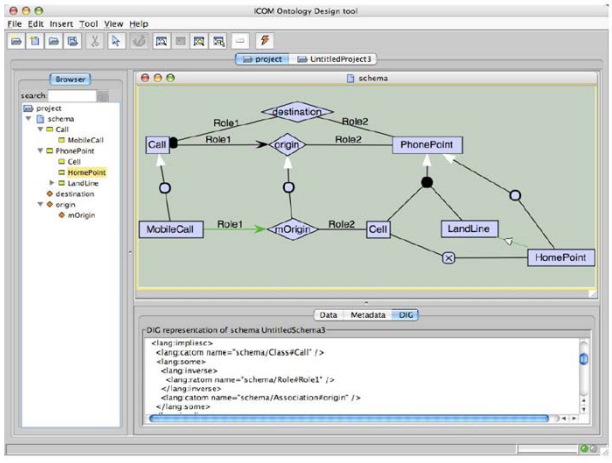
Generating an ontology from the database schema

- Analysing the user queries (Vipul Kashyap. Design and creation of ontologies for environmental information retrieval, 1999)
- Analysing the data (Yuri A. Tijerino et al. Towards ontology generation from table. World Wide Web, 8(3):p 261-285, 2005 → **TANGO, Table Analysis for Generating Ontologies**)

Khaos

Generating an ontology from the database schema

- DEMO ICOM (Franconi)



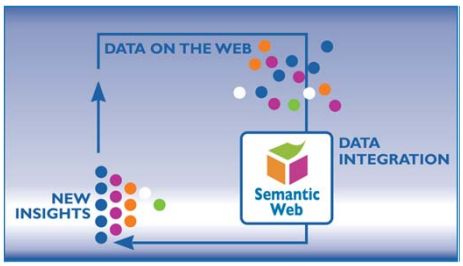
Khaos

From an ontology and a database generating a SPARQL endpoint

- There is a growing number of application patterns referring to the Semantic Web:
  - data integration using RDF, SKOS, OWL, …
  - knowledge engineering with complex ontologies, using, e.g., OWL and/or rule based reasoning
  - better data management, archiving, cataloguing, etc., e.g., digital library applications
  - managing, coordinating, combining Web services
  - intelligent software agents
  - improving search (usually using domain specific vocabularies…)
- It shows that the Semantic Web is a mature technology

Khaos

- But it is good to (re-)emphasize some principles
- The Semantic Web:
  - *extends the principles of the Web from documents to data*; create a *Web of data*



Khaos

- It is the Semantic *Web*, and not only Semantics!
  - data, ontologies, vocabularies, etc, can (and should!) be shared, reused, potentially on Web scale
    - *the "network effect" on data*
- The major importance of RDF is that it provides an abstract integration layer for data on the Web

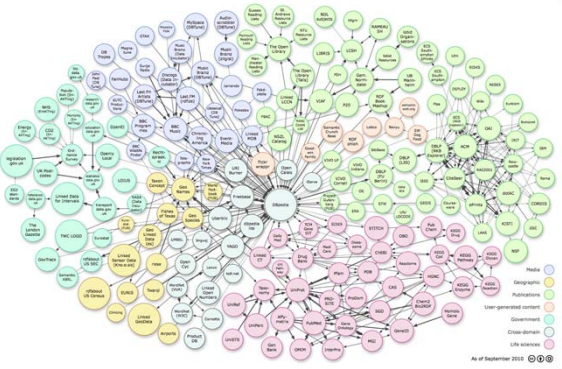
But is there already a Web of Data out there?

Khaos

Linking Open Data Project

- Goal: "expose" open datasets in RDF
- *Set RDF links among the data items* from different datasets
  - Set up query endpoints
  - Billions triples, millions of "links"

The real value is in the links!



Khaos

Example data source: DBpedia

- DBpedia is a community effort to
  - extract structured ("infobox") information from Wikipedia
  - provide a SPARQL endpoint to the dataset
  - interlink the DBpedia dataset with other datasets on the Web

Freie Universität Berlin

UNIVERSITÄT LEIPZIG

OPENLINK SOFTWARE

Khaos

Extracting structured data from Wikipedia

```
@prefix dbpedia <http://dbpedia.org/resource/>.
@prefix dbterm  <http://dbpedia.org/property/>.

                dbpedia:Amsterdam
        dbterm:officialName "Amsterdam" ;
                 dbterm:longd "4" ;
                dbterm:longm "53" ;
                dbterm:longs "32" ;
                        ...
      dbterm:leaderName dbpedia:Job_Cohen ;
                        ...
           dbterm:areaTotalKm "219" ;
                        ...
                 dbpedia:ABN_AMRO
      dbterm:location dbpedia:Amsterdam ;
                        ...
```

Khaos

Accessing the cloud

- Applications can access the data directly (via the URI-s)
- There are several "instantiations" of part of the cloud that user can access
  - these store copies of several "blobs"
    - possibly with some inferred triples based on, eg, OWL
  - often offering a SPARQL endpoint to query to cloud
- Ontologies are necessary to properly integrate data
  - "term used in this dataset relates to the term used there this and this way…"
  - OWL, Rules, RDF vocabularies are vital

Khaos

Publishing Linked Data on the Web

- Publishing a data set as Linked Data on the Web involves the following three basic steps:

  1. Assign URIs to the entities described by the data set and provide for ereferencing these URIs over the HTTP protocol into RDF representations.

  2. Set RDF links to other data sources on the Web, so that clients can navigate the Web of Data as a whole by following RDF links.

  3. Provide metadata about published data, so that clients can assess the quality of published data and choose between different means of access.

Khaos

Tool Demo (Virtuoso Conductor)

Ontologies for Designing Database-based systems

María del Mar Roldán, Ismael Navas

mmar@lcc.uma.es, ismael@lcc.uma.es

Khaos INVESTIGACIÓN