Applied Ontologies
Industrial applications of knowledge graphs

Juan Gómez Romero
Research Fellow
http://decsai.ugr.es/~jgomez

Departamento de Ciencias de la
Computación e Inteligencia Artificial
http://decsai.ugr.es
Second oldest university in Spain (1531)
60,000 graduate and post-graduate students
28 teaching centres
75 degrees, 68 MSc degrees, 116 PhD degrees

Computer Science and Artificial Intelligence (1988)
Ranked 42 in ARWU-2015 Computer Science
70 permanent professors and lecturers, 50 research associates
PhD Programme in Data Science

Research Fellow @ DECSAI
1. Motivation
2. Context-aware computing
3. Knowledge-based systems & NLP
4. Current trends and opportunities
“An ontology is a formal, explicit specification of a shared conceptualization.”

**Ontologies** == Knowledge models with special features

**Formal**
Mathematical underpinnings: unambiguous, automatic inference, etc.

**Machine-processable**
Well-defined representation languages: RDF(S), OWL
Information exchange (different serializations), query (SPARQL), storage (*triplestores*), etc.

**Standard**
W3C standardization
Interaction with other property-graph software: TinkerPop (+Gremlin), Neo4j (+Cypher), etc.

**Tools**
Editors: Protégé, TopBraid
APIs: Apache Jena, RDF4J (previously Sesame), OWL API, RDFLib, etc.
Triplestores: Virtuoso, Blazegraph, GraphDB, etc.
Reasoning engines: HermiT, RACER, Stardog, Pellet, ELK, etc.
Knowledge base development

Support knowledge-based systems
From simple (pizza recommender) to complex (galen, umls, gene ontology)
Pizza and (hasTopping some MozzarellaTopping) and ...

Publish open linked data
DBPedia, Wikidata
Geonames
YAGO2
Drugbank
data.gov

Information exchange & annotation format
DCAT (datasets)
RDF Data Cube (statistical data)
1. Motivation
2. Context-aware computing
3. Knowledge-based systems & NLP
4. Current trends and opportunities
An example: **video-surveillance systems**

**Objective**

To achieve a high degree of understanding of the scene from multiple observations to barely require operator attention while cutting component costs

Tracking moving objects with Kalman filter + identification

Issues
**Tracking**

Track 008
- pos ()
- vel()

Track 010
- pos()
- vel()

**Low level**
- Person
- Entry
- > Entering
- Mirror
- > Reflection
- Column

**High level**
- (Entering through Entry 2)
- (Reflected by Mirror 1)

**Interpretation**
- Person 1 is
- (Entering through Entry 2)
- and
- (Reflected by Mirror 1)

**Context**
Context-aware systems

Computational systems that use a massive amount of context knowledge
The interpretation of the available information depends on context knowledge

Ambient Intelligence & Ubiquitous Computing


Axioms:

Fridge
Couch
Person

Tracked

Ambient computing

<owl: Thing rdf: about="fridge">
<rdf:type rdf: resource="#Fridge"/>
<scob: hasObjectSnapshot rdf: resource="#osn_fridge"/>
</owl: Thing>

<owl: Thing rdf: about="#osn_fridge">
<rdf:type rdf: resource="?scobySceneObjectSnapshot"/>
<scob: hasObjectProperties rdf: resource="#fridgeProps"/>
</owl: Thing>

<owl: Thing rdf: about="#fridgeProps">
<rdf:type rdf: resource="?scobyObjectSnapshotProperties"/>
<scob: hasPosition rdf: resource="#fridgePosition"/>
</owl: Thing>

<owl: Thing rdf: about="#fridgePosition">
<rdf:type rdf: resource="?scob1 OPosition"/>
<scob: oPositionValue rdf: resource="#p1"/>
<scob: oPositionValue rdf: resource="#p2"/>
<scob: oPositionValue rdf: resource="#p3"/>
<scob: oPositionValue rdf: resource="#p4"/>
<scob: oPositionValue rdf: resource="#p5"/>
</owl: Thing>

<owl: Thing rdf: about="#fridgePoint">
<rdf:type rdf: resource="?trenPoint"/>
<tren: x:datatype="xsd: float">687.0</tren: x:datatype=
</owl: Thing>
Communication level information fusion procedures. Smart cameras implement an Tracking data is introduced into the cognitive layer to initiate more comp track multi pipelined structure of several modules, which correspond to the successive stages current images. This process this section, we describe in detail the structure and the functions of the smart C r M sol g!e o t a t a i o t n.\[4 b d c R at e \[45\] l2 I n c r p c e n a.

\[\text{person touches fridge}\]

\[\text{fix track positions}\]
Communication layer, abstract ontologies are used to describe abstract entities. A schema for context information exploitation is introduced into the cognitive layer to initiate more complex tracking data. This process track multiple targets within the local field of view. The tracking layer is arranged in a pipelined structure of several modules, which correspond to the successive stages of the targets within the local field of view. The tracking layer is arranged in a pipelined structure of several modules, which correspond to the successive stages of the targets within the local field of view.

This process involves the ontological model in the update and initialization/deletion steps. In this section, we describe in detail the structure and the functions of the smart cameras and the fusion node.

5.5. Situation fusion procedures and additional tasks may be started as a result of this situation; for example, if a person is using the fridge as the touched object, we could launch a warning to the person or to the remote operator. Besides, the new situation information is encoded with ontologies and the processes to express it in a vocabulary.

UsingFridge

<table>
<thead>
<tr>
<th>Camera 1 simple scene recognition</th>
<th>Camera 2 simple scene recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>%SingleFridge:U singFridge</td>
<td>%SingleFridge:UsingFridge</td>
</tr>
<tr>
<td>?person !tren:Track</td>
<td>?person !tren:Track</td>
</tr>
<tr>
<td>%#!smarthome:Enclosing</td>
<td>%#!smarthome:Enclosing</td>
</tr>
<tr>
<td>A simple rule to fuse</td>
<td>A simple rule to fuse</td>
</tr>
<tr>
<td>%SingleFridge:UsingFridge:Non</td>
<td>%SingleFridge:UsingFridge:Non</td>
</tr>
<tr>
<td>!person !object</td>
<td>!person !object</td>
</tr>
<tr>
<td>%#!smarthome:Fridge</td>
<td>%#!smarthome:Fridge</td>
</tr>
<tr>
<td>&amp;</td>
<td>&amp;</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>#</td>
<td>#</td>
</tr>
<tr>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>person <em>action</em> breakfast</td>
<td></td>
</tr>
</tbody>
</table>
Fuzzy ontologies for situation representation

\[
\langle (a, b) : \text{nearOf} \geq \alpha \rangle, \alpha = \begin{cases} 
1 & \text{dist}(a, b) \leq d_1 \\
0 & \text{dist}(a, b) > d_1 + d_2 \\
\frac{d_1 + d_2 - \text{dist}(a, b)}{d_2} & \text{otherwise (}d_2 \neq 0\text{)}
\end{cases}
\]

Fuzzy / belief-based aggregation for threat assessment
Limitations
Knowledge base must be manually created
Context description
Scene recognition

Solutions
Hybridize with Machine Learning
Automatic feature extraction


7. Grand challenges
C. Flexible models to recognize high-level activities. More complex high-level activities need to be recognized other than only simple daily activities. *It is difficult to determine the hierarchical structure of high-level activities because they contain more semantic and context information*. Existing methods often ignore the correlation between signals, thus they cannot obtain good results.
1. Motivation
2. Context-aware computing
3. Knowledge-based systems & NLP
4. Current trends and opportunities
An example: **building information model**

**BIM**: representation of volumes, materials and equipment in a building

US National Building Information Model Standard Project Committee:
A digital representation of physical and functional characteristics of a facility.
Shared knowledge resource for information about a facility that provides support for decision-making during its life-cycle

**IFC** (Industry Foundation Classes) specification
Object-based data model (EXPRESS) + text-based file interchange format (STEP)
Allows creating readable models and data validation rules
Lacks a mathematical characterization of the semantics of its representation primitives
Clinic_Plumbing_20121206
https://www.nibs.org/page/bsa_commonbimfiles?&hhsearchterms=%22common+and+bim+and+file%22%3E#project3
2012-03-23-Duplex-02-Design-COBIe

https://www.nibs.org/page/bsa_commonbimfiles?&hhsearchterms=%22common+and+bim+and+file%22#project1

ISO-10303-21;
HEADER;
FILE_DESCRIPTION ('', '2;1');
FILE_NAME ('', '2012-03-26T07:45:57', (''), (''), (''), '');
FILE_SCHEMA ('IFC2X3');
ENDSEC;
DATA;
#528817 = IFCRELDEFINESBYPROPERTIES('3jRe8Qj014LexP6MAaoeLC',#521411,$,$,(#521705),#528819);
#528818 = IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(21.422000885009766),$);
#528819 = IFCPROPERTYSET('0BTfgRhSzE7A4ylBNY0c08',#521411,'PSet_Revit_Dimensions',$,(#528818,#528772));
#528823 = IFCPROPERTYSINGLEVALUE('Volume','Volume',IFCREAL(12.239999771118164),$);
#528825 = IFCRELDEFINESBYPROPERTIES('0SxgxlR9HBTv4S800y8qKz',#521411,$,$,(#521767),#528827);
#528826 = IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(15.319000244140625),$);
#528827 = IFCPROPERTYSET('0s2gvnbuHFSPHzUBVmekO5',#521411,'PSet_Revit_Dimensions',$,(#528826,#528827));
#528828 = IFCRELDEFINESBYPROPERTIES('1DT1FrbgAzBJW73xLHG0',#521411,$,$,(#521829),#528830);
#528830 = IFCPROPERTYSET('3kej1LMmLFFv15cORUun2',#521411,'PSet_Revit_Dimensions',$,(#528831,#528842));
#528831 = IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(5.434999942779541),$);
#528800 = IFCRELDEFINESBYPROPERTIES('1YTeCslg99wBkwkn57MVq',#521411,$,$,(#521668),#528802);
#528803 = IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(9.840999603271484),$);
#528802 = IFCPROPERTYSET('1c9QrLEi51DAOC5swSkN0jT',#521411,'PSet_Revit_Dimensions',$,(#528803,#528823));
Mapping from IFC to OWL > ifcOWL ontology
IFC-to-RDF tool
Querying IFC RDF

“All the building elements built from concrete”

```
Class: :BuildingElementsMadeOfConcrete

equivalentTo:
    ifc:IfcBuildingElement
    and
    (inverse ifc:relatedObjects_of_IfcRel Associates
        some (ifc:relatingMaterial
            some (ifc:IfcMaterial
                and (ifc:name value "CONCRETE")))))
```

(solved by a reasoning engine)

+ More complex expressions
+ User-defined concepts
+ Detection of inconsistencies
Fuzzy IFC

Fuzzy IFC

**Example 1.** If we consider the class IfcMaterial as a fuzzy concept, we can add a new instance representing “paper” that can be only partially considered a material. (Note that IfcMaterial already has an individual, material_1, as depicted in Figure 2.)

\[
\text{(instance :material_2 ifc:IfcMaterial } \geq 0.8) \\
\text{(related :material_2 "PAPER" ifc:name_of_IfcMaterial)}
\]
Fuzzy IFC

Example 2. A new fuzzy role has been defined in the ontology to relate the similarity degree between two building materials, namely the similar_to_IfcMaterial object property. This property can be defined as symmetric (R9), because it holds in both directions (with the same degree), and transitive (R8). By extension, it would be possible to define other features of the property with the axioms R3-R14: reflexive, irreflexive, functional, etc. Let us also suppose that we have in the fuzzy ontology additional instances of IfcMaterial representing ‘mortar’ and ‘ecologic mortar’ materials. We can now assert that ‘concrete’ is quite similar to ‘mortar’, but ‘mortar’ is only moderately similar to ‘ecologic mortar’.

(instance :material_3 ifc:IfcMaterial)  
(related :material_3 "MORTAR"  
ifc:name_of_IfcMaterial)

(instance :material_4 ifc:IfcMaterial)  
(related :material_4 "ECOLOGIC MORTAR"  
ifc:name_of_IfcMaterial)

(symmetric :similar_to_IfcMaterial)  
(transitive :similar_to_IfcMaterial)

(related :material_1 :material_3  
:similar_to_IfcMaterial >= 0.8)  
(related :material_3 :material_4  
:similar_to_IfcMaterial >= 0.6)

Queries

(some :similar_to_IfcMaterial  
(value ifc:name "CONCRETE"))

(ifc:IfcBuildingElement  
(some inv ifc:relatedObjects_of_IfcRelAssociates  
(some ifc:relatingMaterial  
(and  
  ifc:IfcMaterial  
  (some :similar_to_IfcMaterial  
     (value ifc:name "MORTAR"))))))
Fuzzy IFC

more…

Fuzzy taxonomies
A concept is partially included into other concept
GlassMaterial is a MineralMaterial with degree 0.8

Fuzzy datatypes
Imprecise statements over a concrete domain
A HighWindow is a window with height defined by the trapezoid (1.2, 1.7, 10, 10)

Fuzzy modifiers
Change the meaning of a fuzzy concept by modulating its membership function
A VeryHighWindow is a Highwindow modulated by the triangle function (0.4 1 1)
Fuzzy IFC

Applications

Cross-domain knowledge linking
A concept is partially included into other concept; graded relationships

Imprecise BIM query
Retrieve instances of fuzzy concepts; e.g. big room, breezeway

Fuzzy parametric modeling
Define soft constraints & use fuzzy constraint satisfaction

Pros & Cons

+ Inferencing
+ Available tools
– Expressiveness is computationally expensive
+/- Ontology modeling knowledge is required
Energy IN TIME
Simulation-based control for energy efficiency building operation and maintenance

Semantic Sensor Network Ontology
W3C Recommendation
19 October 2017
https://www.w3.org/TR/vocab-ssn/
Energy IN TIME
Simulation-based control for energy efficiency building operation and maintenance

Semantic Sensor Network Ontology
An example (II): **Natural Language Processing & Information Retrieval**

![Diagram](image)

- **Entity Recognition**
- **Entity Disambiguation**
- **Instance Extraction**
- **Fact Extraction**
- **Query & Reasoning**

**Paris**

[Link to DBpedia page](http://dbpedia.org/page/Paris)

```sparql
SELECT ?person
WHERE {
  ?person dbo:birthPlace dbr:Paris ;
  rdf:type dbo:Scientist
} LIMIT 100
```

---

**Example**

- **Natural Language Processing & Information Retrieval**
- **Entity Recognition**
- **Entity Disambiguation**
- **Instance Extraction**
- **Fact Extraction**
- **Query & Reasoning**

**Paris**

[Link to DBpedia page](http://dbpedia.org/page/Paris)

```sparql
SELECT ?person
WHERE {
  ?person dbo:birthPlace dbr:Paris ;
  rdf:type dbo:Scientist
} LIMIT 100
```
ePOOLICE
Early pursuit against organized crime using environmental scanning, the law and intelligence systems

Extracting & processing open data to provide support to strategic analysis by means of an integrated indicator dashboard

**Data acquisition**
Web, External databases, Internal knowledge repository

**Text processing**
Entity recognition, Document categorization and filtering

**Pattern discovery**
Mining of relationships between entities, Discovery of trends correlations

**Situation and threat assessment**
Threat models, Information Fusion, Alarms

**Visualizing, interpreting, discovering**
Map-based dashboard
Knowledge-based systems

Environmental scanning

Available data (World)

Web crawling

Data Acquisition Components

Analysis Goal

Relevant Information Sources

Analysis Product
  > Synthesis/Summary
  > New knowledge / models
  > New Hypothesis

Visualization

New Knowledge (models/sources)

A priori knowledge

Existing models

Enriched data mining

Semantic knowledge representation

Entity recognition

Relation identification

Document classification

Structured data Information

Environmental Knowledge Repository

Relevant raw data

Structured data Information

Structured data Information
Monitoring indicators of Traffic of Human Beings in the UK

Crawling

Structured data

Pattern recognition

Anomaly detection

Threat assessment

Decision-making

RDF data

Monitoring indicator evolution and application to other regions

Structured data

NLP

Correlation between presence of non-legit private colleges and organized crime in some regions, but not in other

Newspaper data about investigations on fraudulent admissions in private colleges + QAA reports

Threat assessment model: Fraud + Conditions favoring traffic of human beings (e.g. low wages) + events of interest (e.g. changes in visa laws) => Opportunities for organized crime groups
Organized Crime taxonomy used for crawling (drug trafficking)
Extracted facts
Extracted facts with probability values
Fuzzy representation of credibility / reliability values
Dashboard indicator
COPKIT

https://copkit.eu/copkit-project-presentation-video/

Analyzing, investigating, mitigating and preventing the use of new information and communication technologies by organized crime and terrorist groups. For this purpose, COPKIT proposes an intelligence-led Early Warning (EW) / Early Action (EA) system for both strategic and operational levels.

Improvements
Federated knowledge base
API for read/write knowledge base
Crowdsourced expert knowledge
Enhanced support for NLP
Blazegraph

https://www.blazegraph.com

“ultra-scalable, high-performance graph database with support for the Blueprints and RDF/SPARQL APIs”

1. High Performance Native graph database
2. Apache TinkerPop™ API or RDF/SPARQL
3. Single machine data storage to ~50B triples/quads
4. REST API with embedded and/or webapp deployment

Virtuoso

https://virtuoso.openlinksw.com

“solution for data access, virtualization, integration and multi-model relational database management (SQL Tables and/or RDF Statement Graphs)”

1. Not-Only-SQL (NoSQL) data management
2. Web application deployment
3. Data privacy & security
4. Maximizing investments in legacy system
GraphDL

https://github.com/jgromero/graphdl

“OWL ontology that allows describing graphs with a simple vocabulary denoting nodes, edges, and properties that can be easily translated into other formats”

J. Gomez-Romero, M. Molina-Solana (2018). GraphDL: An Ontology for Linked Data Visualization. 18th Conference of the Spanish Association for Artificial Intelligence (CAEPIA 2018)

Topbraid Composer


TopBraid Composer™ Maestro Edition (TBC–ME) combines world's leading semantic web modeling capabilities with the most comprehensive data conversion options and a powerful Integrated Development Environment (IDE) for building semantic web and Linked Data applications.
**SHACL**

[https://www.w3.org/TR/shacl/](https://www.w3.org/TR/shacl/)

Shapes Constraint Language: Language for validating RDF graphs against a set of conditions (*shapes*), which are as well expressed in RDF.

```plaintext
Example shapes graph

ex:PersonShape
    a sh:NodeShape ;
    sh:targetClass ex:Person ; # Applies to all persons
    sh:property [ # _:_b1
        sh:path ex:ssn ; # constrains the values of ex:ssn
        sh:maxCount 1 ;
        sh:datatype xsd:string ;
        sh:pattern "^\d{3}-\d{2}-\d{4}" ;
    ] ;
    sh:property [ # _:_b2
        sh:path ex:worksFor ;
        sh:class ex:Company ;
        sh:nodeKind sh:IRI ;
    ] ;
    sh:closed true ;
    sh:ignoredProperties ( rdf:type ) .
```
**SHACL to GraphQL**

[https://www.topquadrant.com/graphql/shacl-graphql.html](https://www.topquadrant.com/graphql/shacl-graphql.html)

"GraphQL schemas are automatically generated using data shape definitions in the Shapes Constraint Language (SHACL)"

**GraphQL**

[https://graphql.org](https://graphql.org)

"Query language for APIs and a runtime for fulfilling those queries"
“A little semantics goes a long way”
James Hendler, co-creator of the Semantic Web

https://www.cs.rpi.edu/~hendler/LittleSemanticsWeb.html
Thanks!

Juan Gómez Romero
Research Fellow
http://decsai.ugr.es/~igomez