UNIVERSIDAD DE GRANADA

DECSAI

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

http://www.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

UNIVERSIDAD DE GRANADA

http://decsai.ugr.es

Computer Science and Artificial Intelligence (1988)

Ranked 42 in ARWU-2015 Computer Science

70 permanent professors and lecturers, 50 research associates

Soft computing: Fuzzy Logic, Genetic Algorithms, Probabilistic Models

PhD Programme in Data Science

DECSAI

http://decsai.ugr.es/~jgomez

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

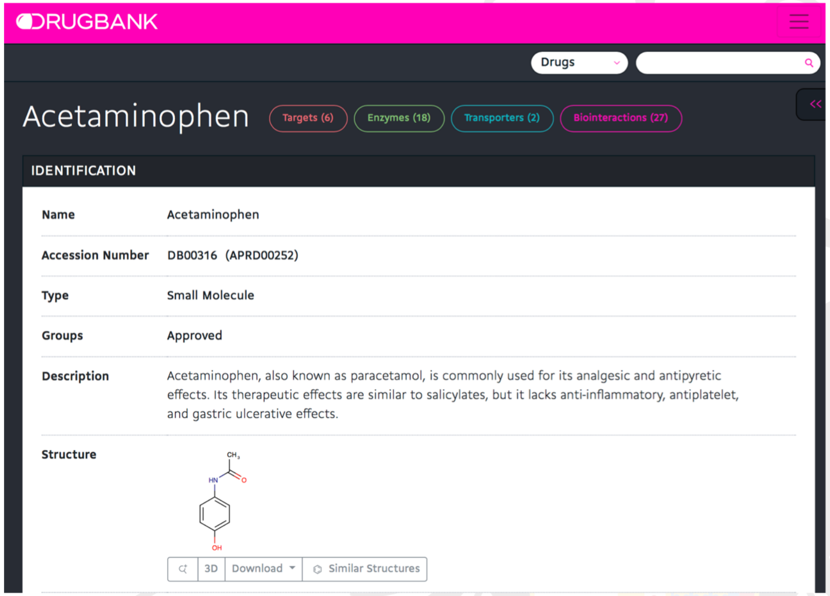
DBPedia, Wikidata

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

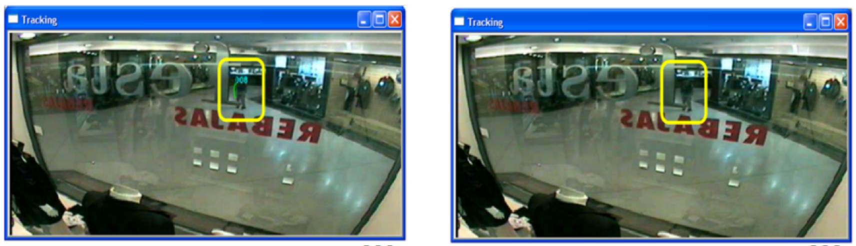
To achieve a high degree of understanding of the scene from multiple observations to barely require operator attention while cutting component costs

From PETS2002 ftp://ftp.pets.rdg.ac.uk/pub/PETS2002

Tracking moving objects with Kalman filter + identification

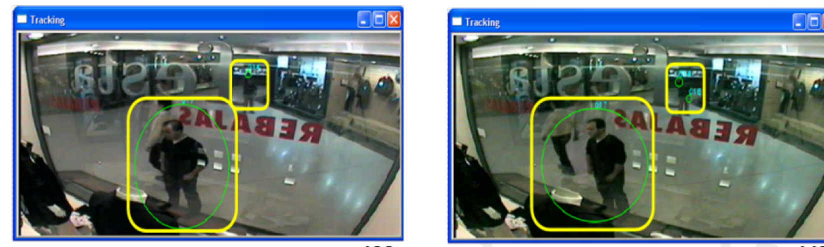
Issues

Track loss



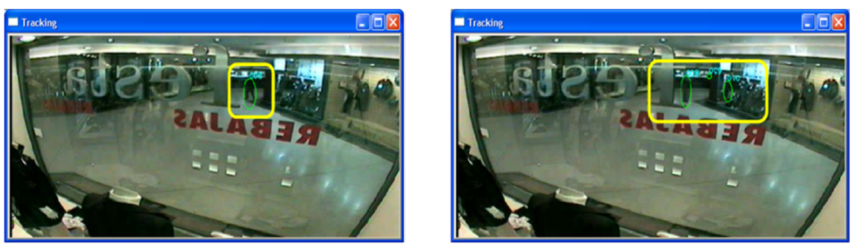
322 323

Bad adaptation to tracked entities (people)



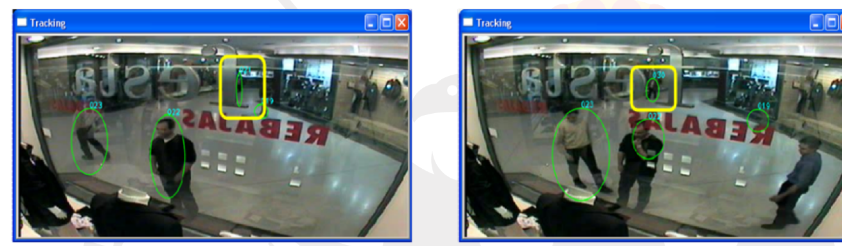
423 442

Reflections



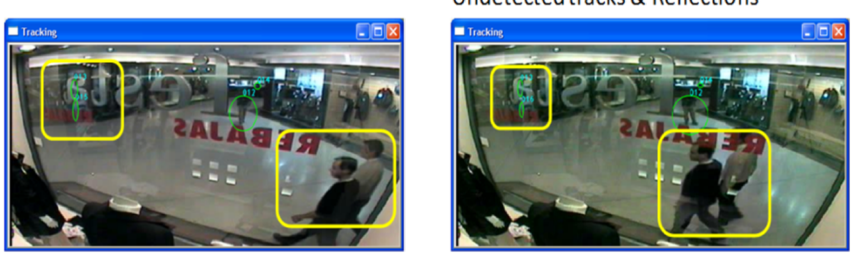
220 225

Occlusions



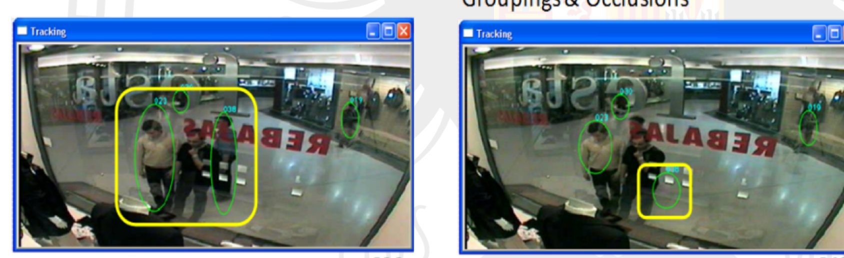
472 571

Undetected tracks & Reflections



350 372

Groupings & Occlusions



629 645

Tracking

```
Track 008
  pos  ()
  vel  ()

Track 010
  pos  ()
  vel()
```

Low level → High level

Person
Entry
> Entering
Mirror
> Reflection
Column

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

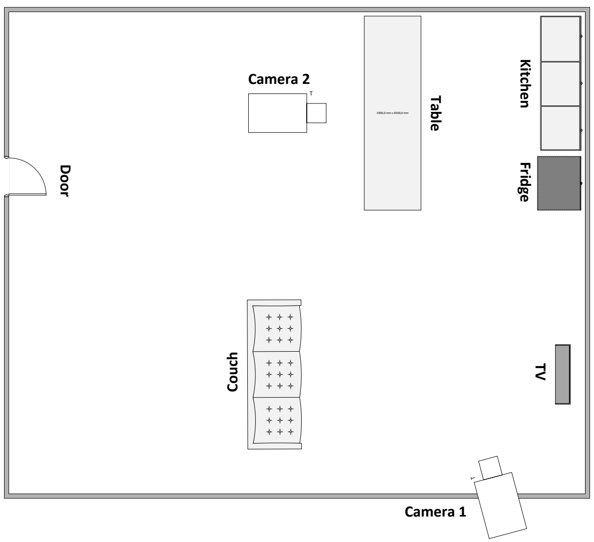
The interpretation of the available information depends on context knowledge

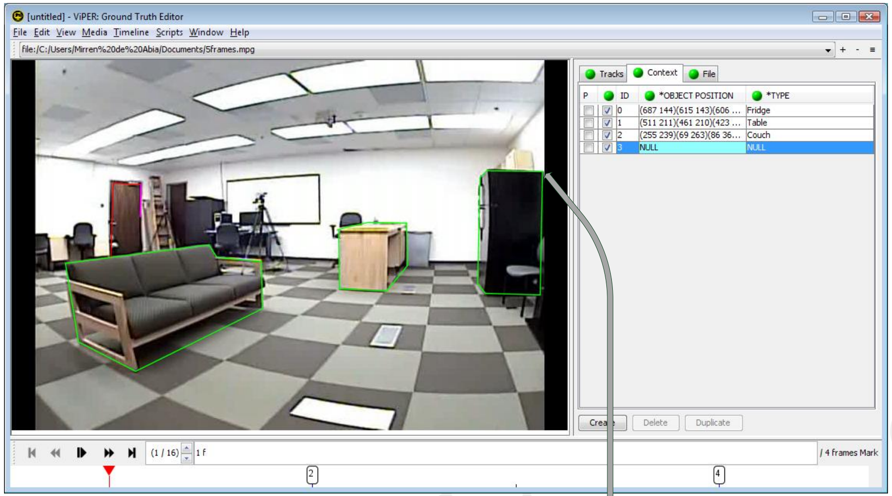
Ambient Intelligence & Ubiquitous Computing

J. Gómez-Romero, M.A. Serrano, M.A. Patricio, J. García & J.M. Molina (2012). *Context-based scene recognition from visual data in smart homes: an Information Fusion approach*. Personal and Ubiquitous Computing 16(7), 835-857.

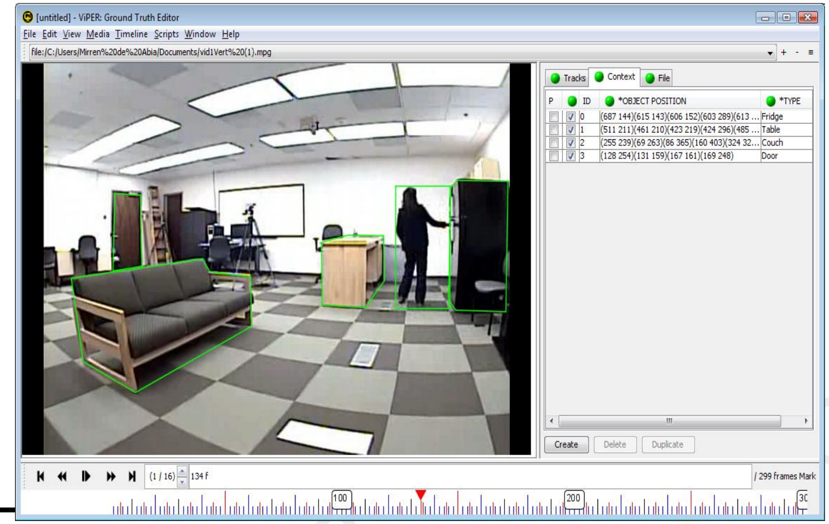
N. Díaz-Rodríguez, M.P. Cuéllar, J. Lilius & M. Delgado (2011). *A fuzzy ontology for semantic modelling and recognition of human behaviour*. Knowledge-Based Systems 66, 46-60.

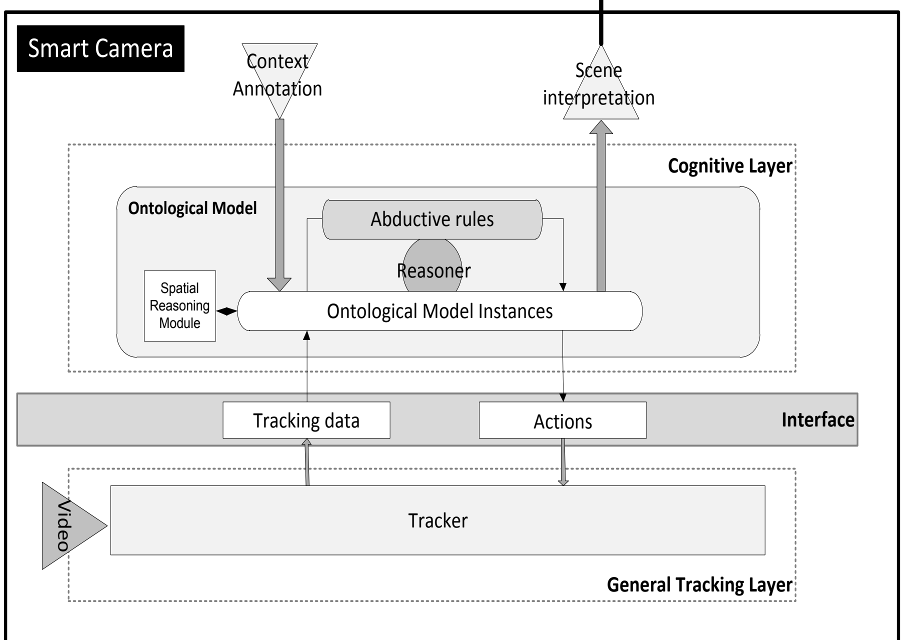
N. Díaz-Rodríguez, M.P. Cuéllar, J. Lilius & M. Delgado (2014). *A survey on ontologies for human behavior recognition*. ACM Computer Surveys 46(4), 43.





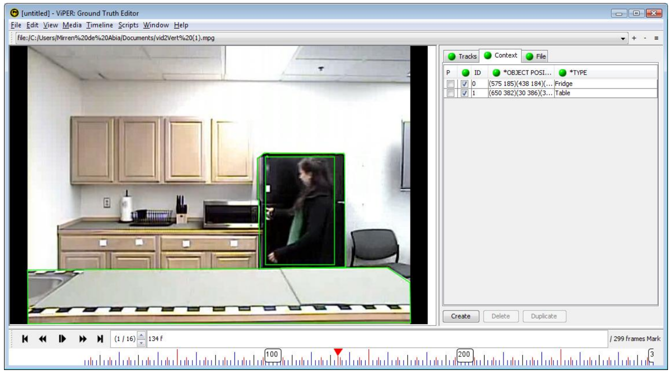

```
<!-- fridge1 instance -->
<owl:Thing rdf:about="#fridge1">
  <rdf:type rdf:resource="#Fridge"/>
  <scob:hasObjectSnapshot rdf:resource="#osn_fridge1"/>
</owl:Thing>

<!-- object snapshot of fridge1 -->
<owl:Thing rdf:about="#osn_fridge1">
  <rdf:type rdf:resource="&scob;SceneObjectSnapshot"/>
  <scob:hasObjectProperties rdf:resource="#fridge1_props"/>
  <tren:isValidInEnd rdf:resource="&tren;unknown_frame"/>
</owl:Thing>

<!-- properties of fridge1 snapshot (position) -->
<owl:Thing rdf:about="#fridge1_props">
  <rdf:type rdf:resource="&scob;ObjectSnapshotProperties"/>
  <scob:OhasPosition rdf:resource="#fridge1_position"/>
</owl:Thing>

<!-- fridge1 position -->
<owl:Thing rdf:about="#fridge1-position">
  <rdf:type rdf:resource="&scob;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"/>
  <scob:OpositionValue rdf:resource="#p6"/>
</owl:Thing>

<!-- fridge1 point1 coordinates -->
<owl:Thing rdf:about="#p1">
  <rdf:type rdf:resource="&tren;2DPoint"/>
  <tren:y rdf:datatype="&xsd;float">687.0</tren:y>
  <tren:x rdf:datatype="&xsd;float">144.0</tren:x>
</owl:Thing>
```

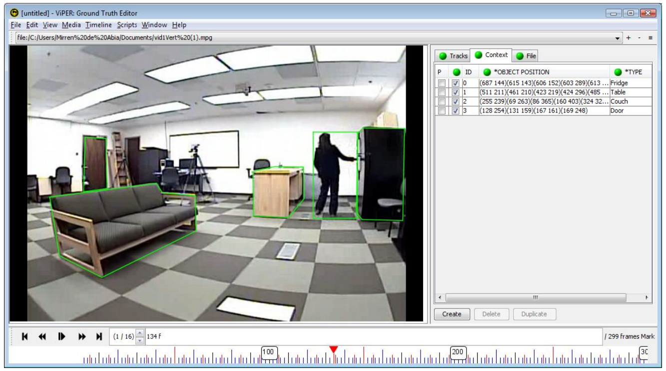


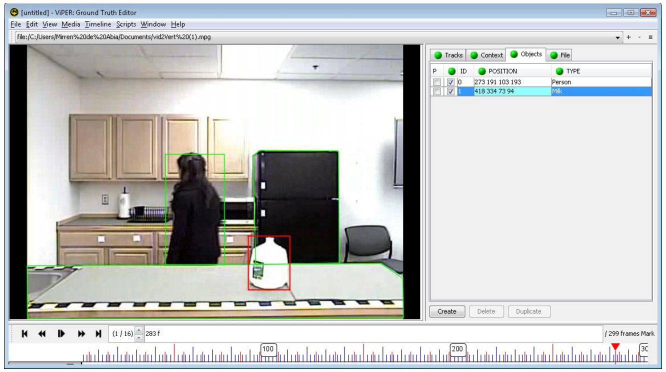


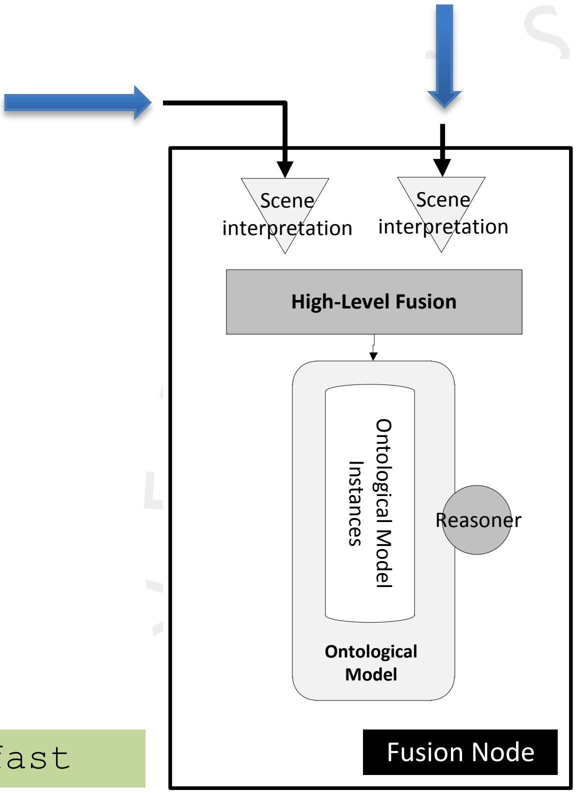
person ***touches*** fridge

fix track positions



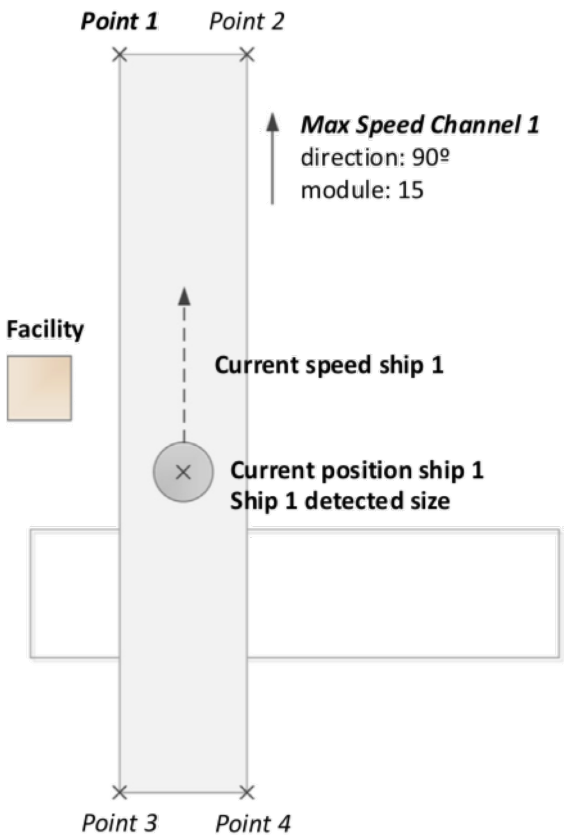






person ***action*** breakfast

J. Gómez-Romero, M.A. Serrano, J. García, J.M. Molina, G. Rogova (2015). *Context-based multi-level information fusion for harbor surveillance*. Information Fusion 21, 173-186.



Representation Channel 1

```
Individual: channel_1
    Types:
        RestrictedArea
    Facts:
     delimitedBy  p_01,
     delimitedBy  p_02,
     delimitedBy  p_03,
     delimitedBy  p_04
     allowedSpeed speed_channel_1

Individual: p_01
    Types:
        Point
    Facts:
     lat  "54.6043",
     lng  "19.1022"

...

Individual: speed_channel_1
    Types:
        Speed
    Facts:
     dir  "90",
     mod  "15"
```

Representation Ship 1

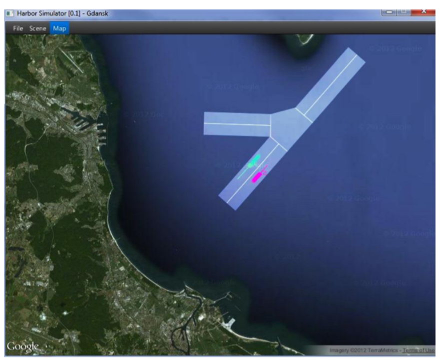
```
Individual: ship1
    Types:
        Vessel
    Facts:
        hasSnapshot sn01

Individual: sn01
    Types:
        Snapshot
    Facts:
        speed speed_1
        length length_1
        position position_1
        insideOf channel_1
        alignedTo channel_1
        closeTo facility_1
...

Individual: length_1
    Types:
        Length
    Facts:
        len "10"
```

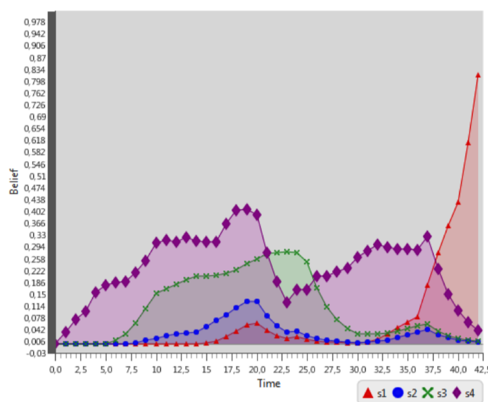
J. Gómez-Romero, M.A. Serrano, J. García, J.M. Molina, G. Rogova (2015). *Context-based multi-level information fusion for harbor surveillance*. Information Fusion 21, 173-186.

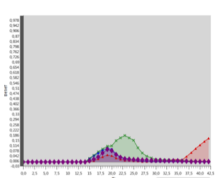
Fuzzy ontologies for situation representation

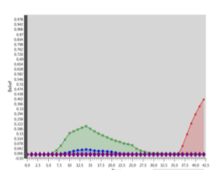
$$\langle (\mathsf{a},\mathsf{b}) : \mathsf{nearOf} \geq \alpha \rangle, \alpha = \begin{cases} 1 & dist(a,b) \leq d_1 \\ 0 & dist(a,b) > d_1 + d_2 \\ \frac{d_1+d_2-dist(a,b)}{d_2} & otherwise\ (d_2 \neq 0) \end{cases}$$

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

J. Wang, Y. Chen, S. Hao, X. Peng, L. Hu (2018). *Deep learning for sensor-based activity recognition: A Survey*. Pattern Recognition Letters, In Press (Corrected Proof).

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

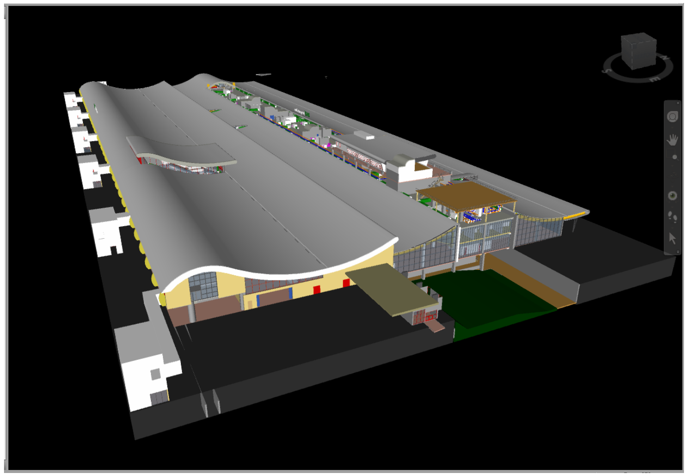
An example: **building information model**

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 Entities according to REVIT MEP elements: HVAC			
Icon	REVIT Element	IFC Type Entity	IFC Occurrence Entity
	Duct	IFCFLOWSEGMENT	IFCDUCTSEGMENTTYPE
	Flexible Duct		
	Duct Fitting	IFCFLOWFITTING	IFCDUCTFITTINGTYPE
	Duct Accessory	IFCBUILDINGELEMENTPROXY	-
	Air Terminal	IFCFLOWTERMINAL	IFCAIRTERMINALTYPE
	Mechanical Equipment	IFCFLOWTERMINAL	IFCAIRTERMINALTYPE

IFC (Industry Foundation Classes) specification

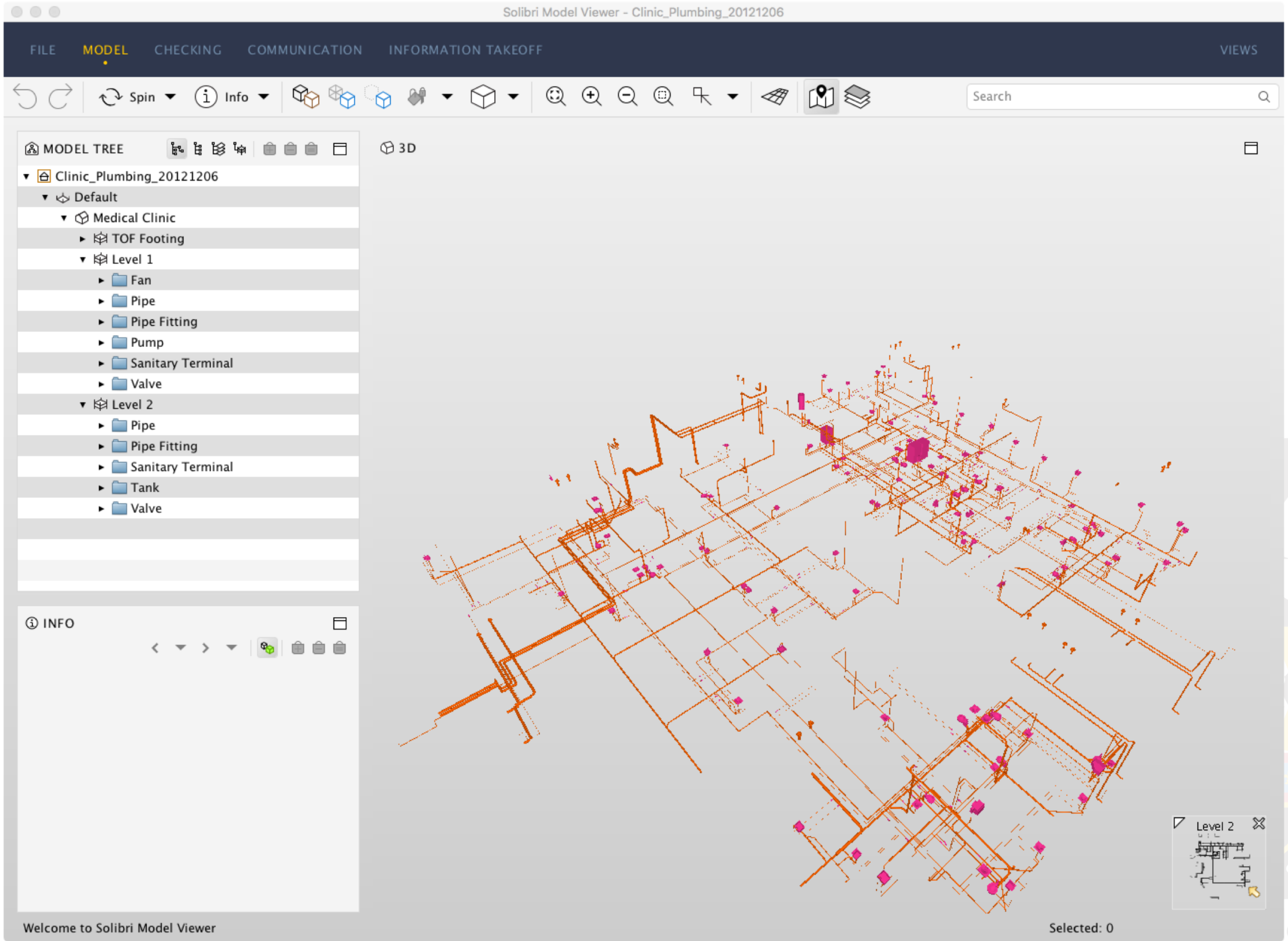
Object-based data model (EXPRESS) + text-based file interchange format (STEP)

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%3E#project1

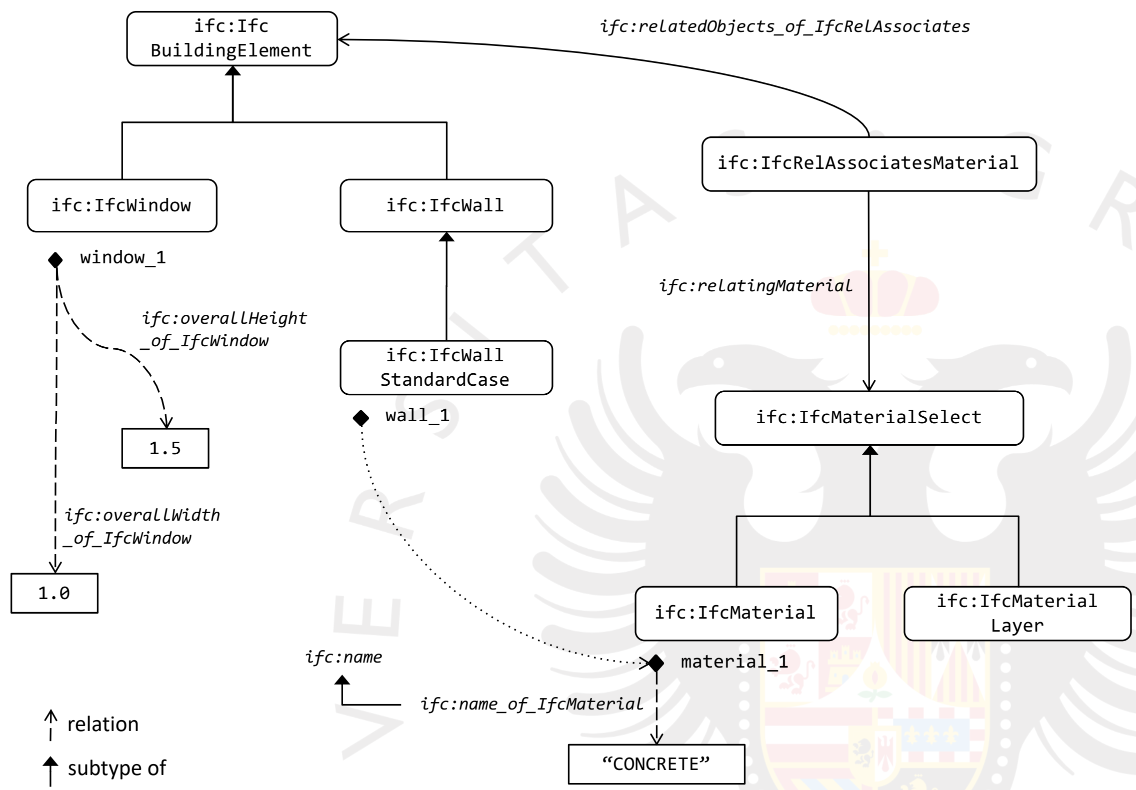
```
ISO-10303-21;
HEADER;
FILE_DESCRIPTION ((''), '2;1');
FILE_NAME ('', '2012-03-26T07:44:57', (''), (''), '', '', '');
FILE_SCHEMA (('IFC2X3'));
ENDSEC;
DATA;
#528817= IFCRELDEFINESBYPROPERTIES('3jRe8Qj014LexP6MAAaocL',#521411,$,$,(#521705),#528819);
#528818= IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(21.422000885009766),$);
#528819= IFCPROPERTYSET('0BTfgrhSzE7A4ylBNY0c08',#521411,'PSet_Revit_Dimensions',$,(#528818,#528772));
#528823= IFCPROPERTYSINGLEVALUE('Volume','Volume',IFCREAL(12.239999771118164),$);
#528825= IFCRELDEFINESBYPROPERTIES('0SxgxlR9HBTv4S8OOy8qKz',#521411,$,$,(#521767),#528827);
#528826= IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(15.319000244140625),$);
#528827= IFCPROPERTYSET('0s2gvnbuHFsPHqUBVmekO5',#521411,'PSet_Revit_Dimensions',$,(#528826,#528815));
#528828= IFCRELDEFINESBYPROPERTIES('1DT1FrbgbAzBJW7JxLFHG0',#521411,$,$,(#521829),#528830);
#528830= IFCPROPERTYSET('3Kej1LMmLFFv1q5cOROun2',#521411,'PSet_Revit_Dimensions',$,(#528831,#528842));
#528831= IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(5.434999942779541),$);
#528800= IFCRELDEFINESBYPROPERTIES('1YTeCslg99wBKwvk5n7MVq',#521411,$,$,(#521668),#528802);
#528803= IFCPROPERTYSINGLEVALUE('Perimeter','Perimeter',IFCREAL(9.840999603271484),$);
#528802= IFCPROPERTYSET('1c9QrLEi51DAOC5wSkN0jT',#521411,'PSet_Revit_Dimensions',$,(#528803,#528823));
```

Mapping from IFC to OWL > ifcOWL ontology

IFC-to-RDF tool

```
Class: ifc:IfcWindow
    SubClassOf:
        ifc:IfcBuildingElement,
        ifc:overallHeight_of_IfcWindow
                only xsd:float,
        ifc:overallHeight_of_IfcWindow
                max 1 xsd:float,
```

```
Individual: window_1
    Types:
        ifc:IfcWindow
    Facts:
        ifc:overallWidth_of_IfcWindow    1.0f,
        ifc:overallHeight_of_IfcWindow   1.5f,
```



Querying IFC RDF

“All the building elements built from concrete”

```
Class: :BuildingElementsMadeOfConcrete
    EquivalentTo:
        ifc:IfcBuildingElement
        and
        (inverse ifc:relatedObjects_of_IfcRelAssociates
            some (ifc:relatingMaterial
                some (ifc:IfcMaterial
                    and (ifc:name value "CONCRETE"))))
```

(solved by a reasoning engine)

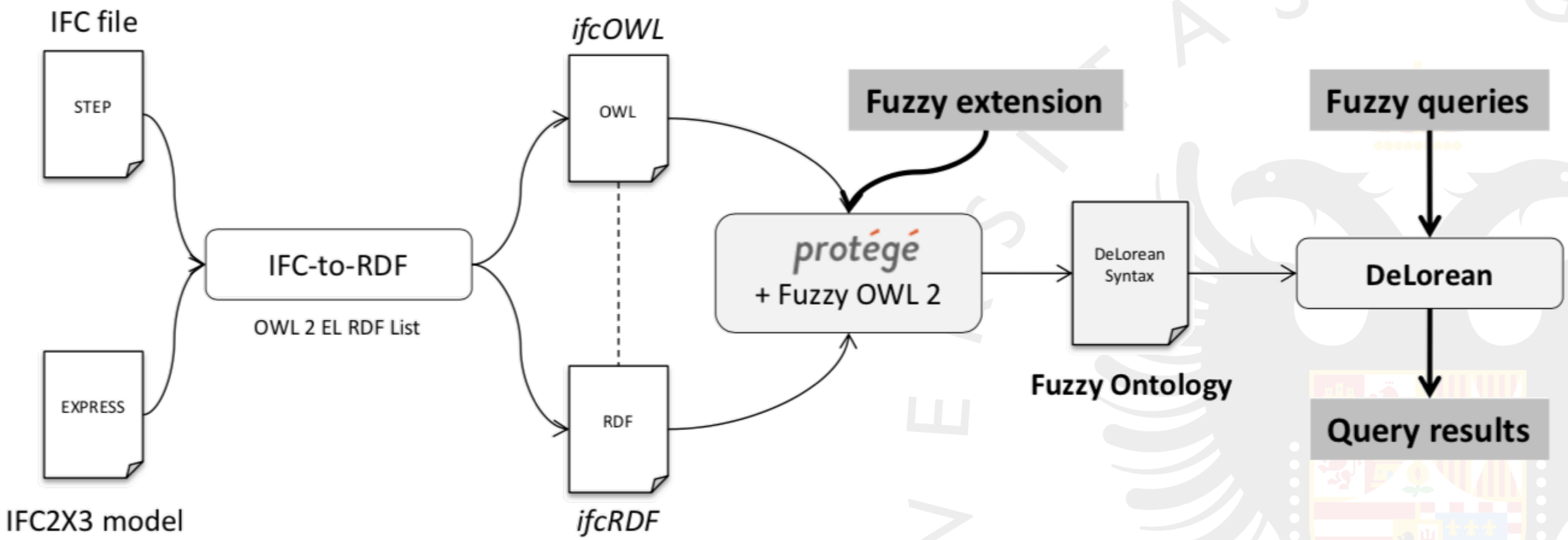
+ More complex expressions

+ User-defined concepts

+ Detection of inconsistencies

Fuzzy IFC

J. Gomez-Romero, F. Bobillo, M. Ros, M. Molina-Solana, M.D. Ruiz, M.J. Martin-Bautista (2015). *A fuzzy extension of the semantic Building Information Model*. Automation in Construction 57, 202-212.



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.)

```
(instance :material_2 ifc:IfcMaterial >= 0.8)
(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"))

(and
    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

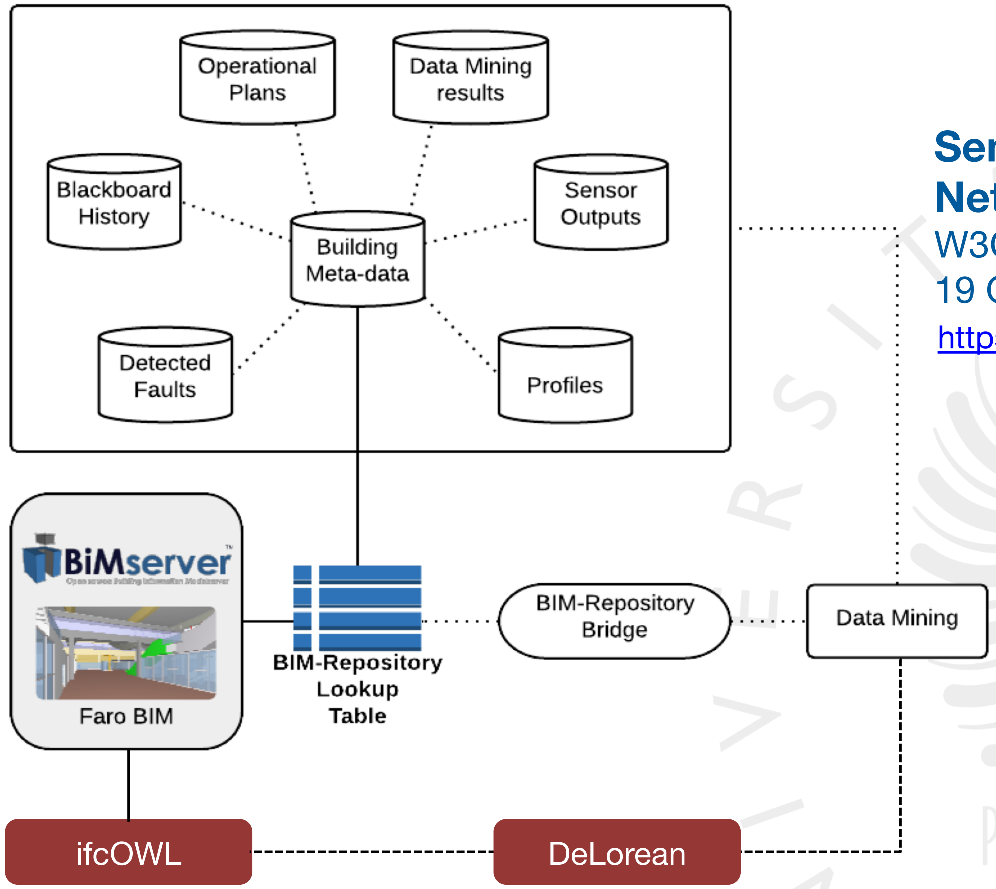
+ Available tools

– Expressiveness is computationally expensive

+/– Ontology modeling knowledge is required

Energy IN TIME

Simulation-based control for energy efficiency building operation and maintenance

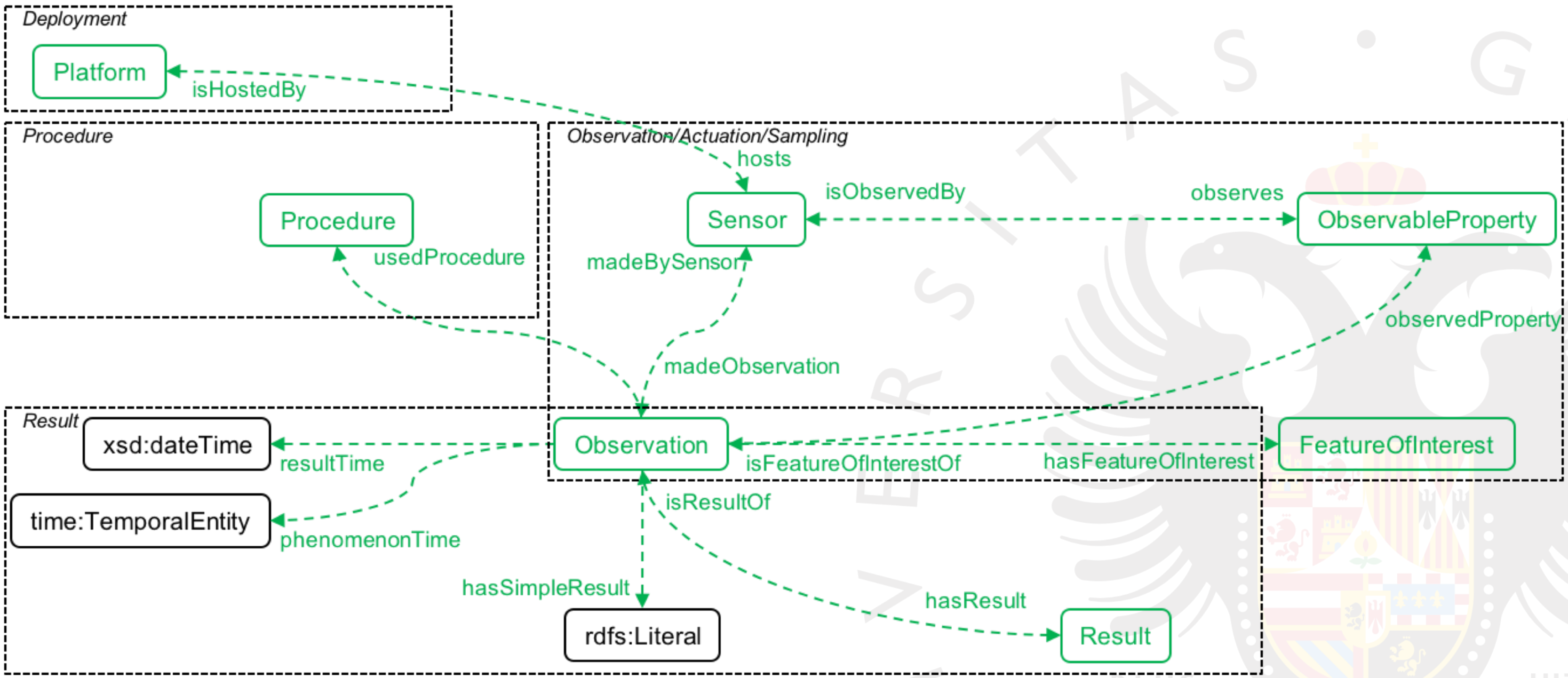


Semantic Sensor Network Ontology
W3C Recommendation
19 October 2017
https://www.w3.org/TR/vocab-ssn/

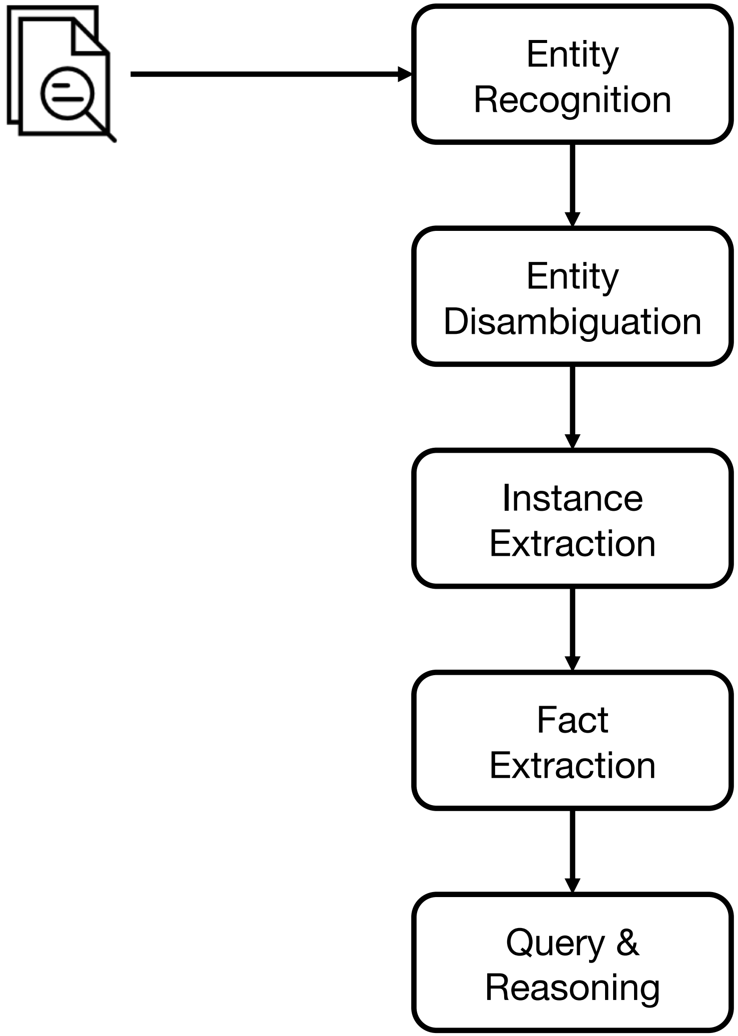
Energy IN TIME

Simulation-based control for energy efficiency building operation and maintenance

Semantic Sensor Network Ontology



An example (II): **Natural Language Processing & Information Retrieval**



Paris

http://dbpedia.org/page/Paris

```
<dbr:Paris, dbo:country, dbr:France>
<dbr:Paris, dbo:mayor, dbr:Anne_Hidalgo>
<dbr:Paris, dbp:wordnet_type, wn:monument-noun>
```

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

link

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

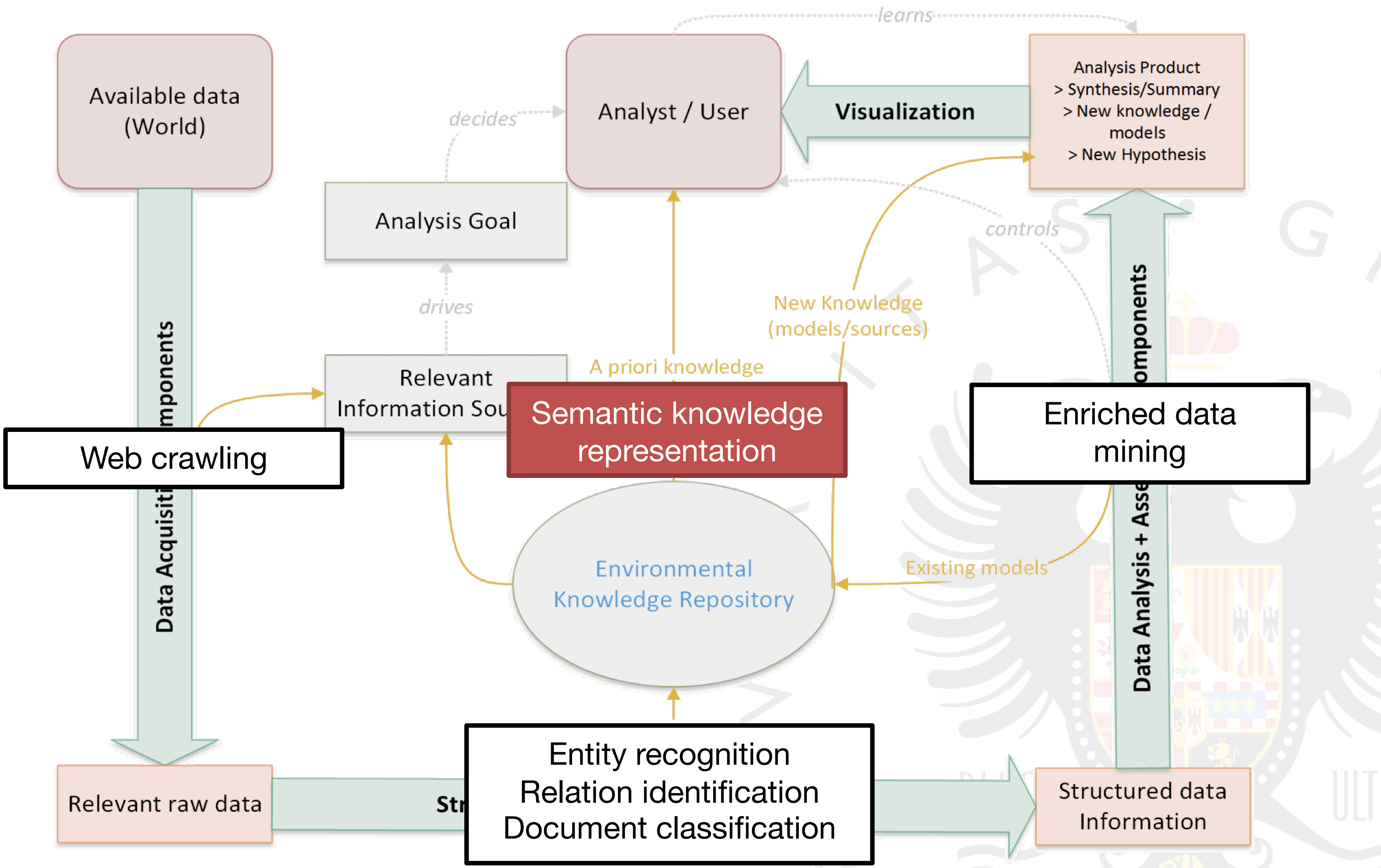
Data acquisition
Web, External databases, Internal knowledge repository

Text processing
Entity recognition, Document categorization and filtering

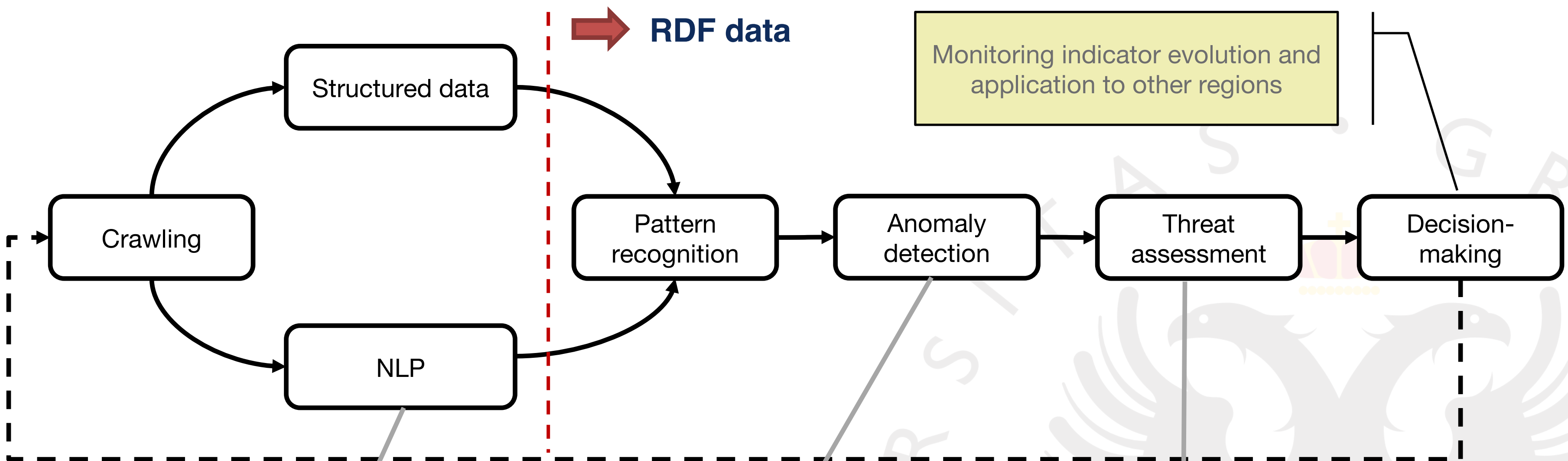
Pattern discovery
Mining of relationships between entities, Discovery of trends correlations

Situation and threat assessment
Threat models, Information Fusion, Alarms

Visualizing, interpreting, discovering
Map-based dashboard


learns
Available data (World)
decides
Analyst / User
Visualization
Analysis Product
> Synthesis/Summary
> New knowledge / models
> New Hypothesis
Analysis Goal
controls
drives
New Knowledge (models/sources)
Relevant Information Sou
A priori knowledge
Semantic knowledge representation
Enriched data mining
Web crawling
Data Acquisiti mponents
Environmental Knowledge Repository
Existing models
Data Analysis + Asse omponents
Relevant raw data
Str
Entity recognition
Relation identification
Document classification
Structured data Information

Monitoring indicators of Traffic of Human Beings in the UK

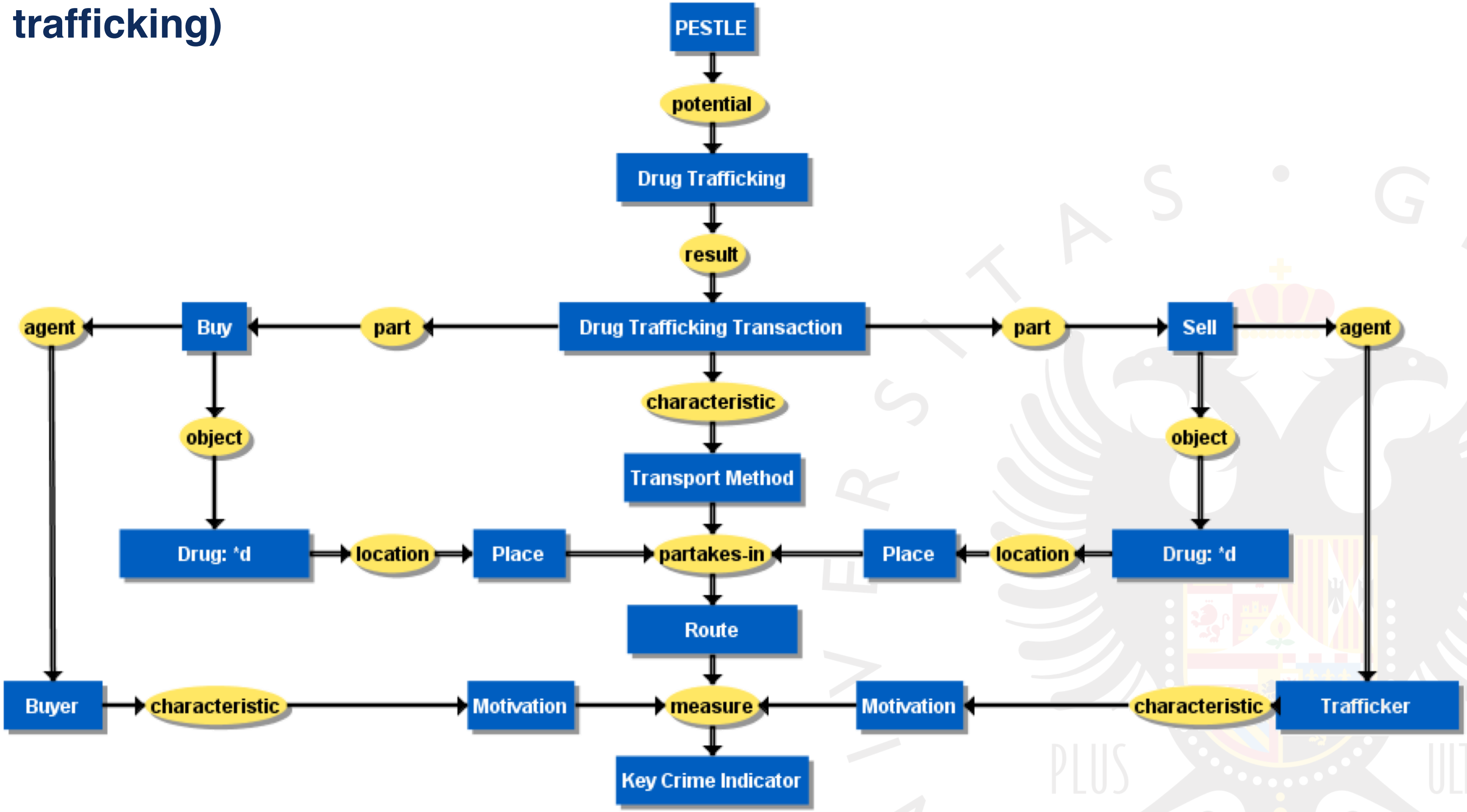


Newspaper data about investigations on fraudulent admissions in private colleges + QAA reports

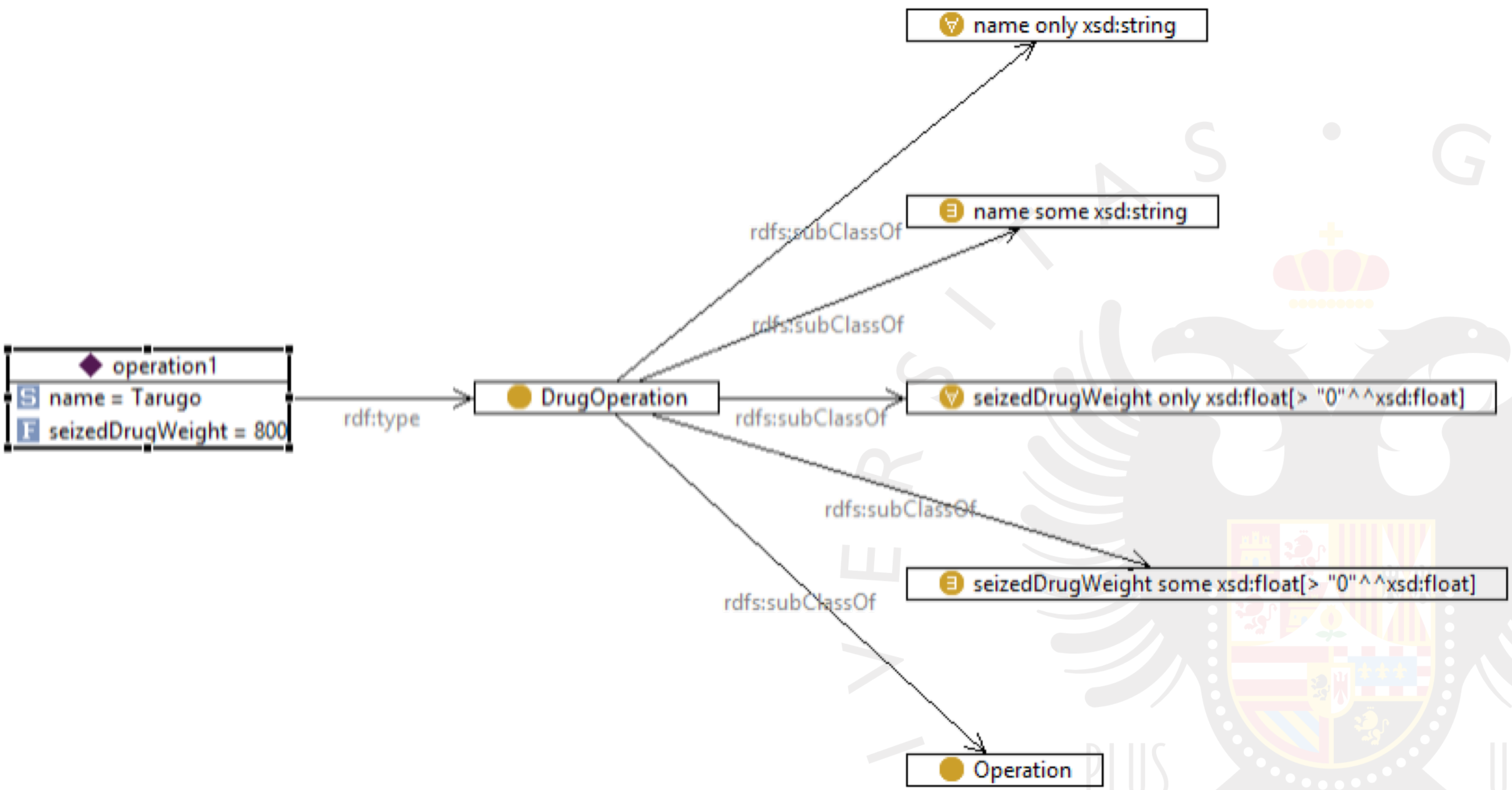
Correlation between presence of non-legit private colleges and organized crime in some regions, but not in other

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

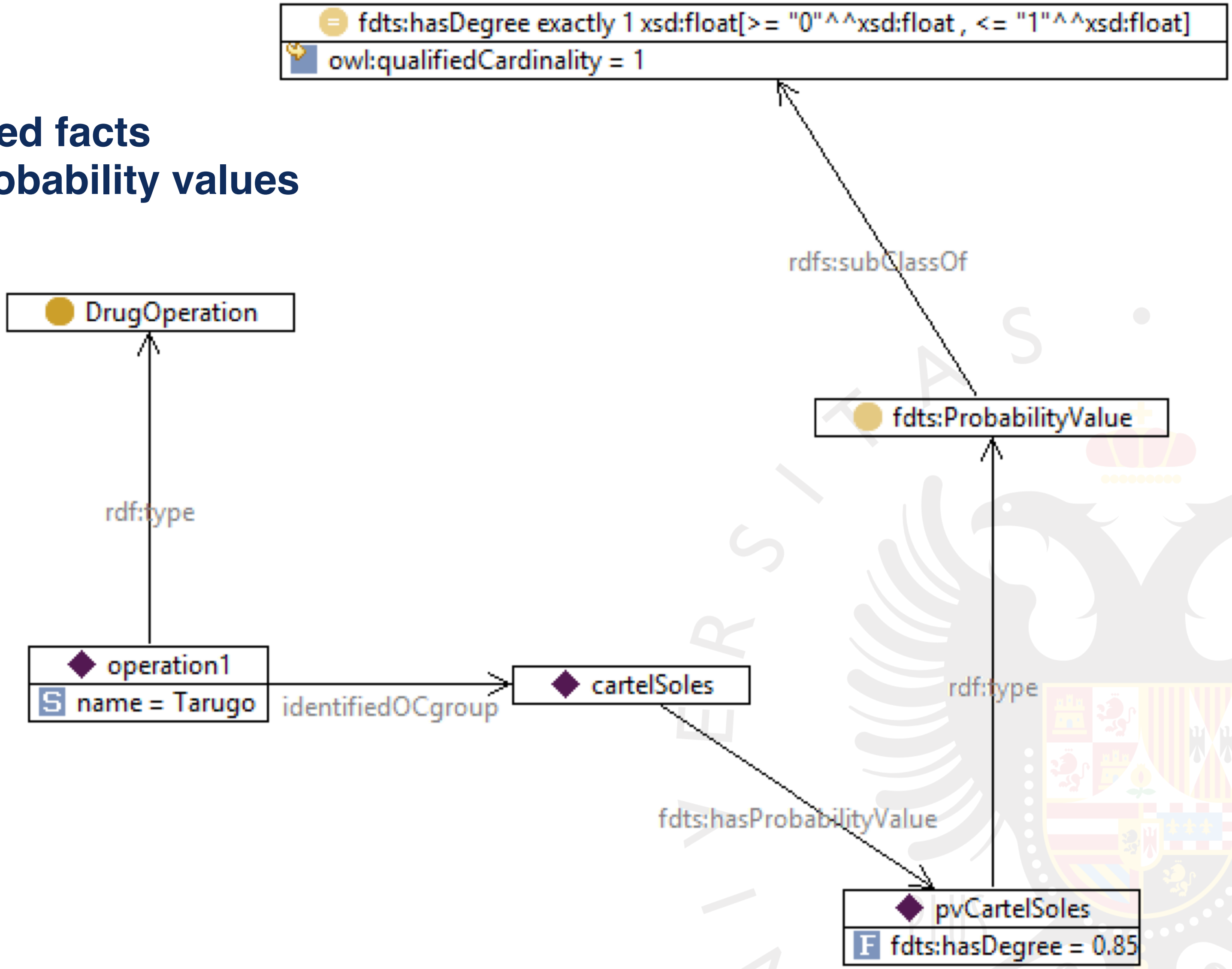
Organized Crime taxonomy used for crawling (drug trafficking)



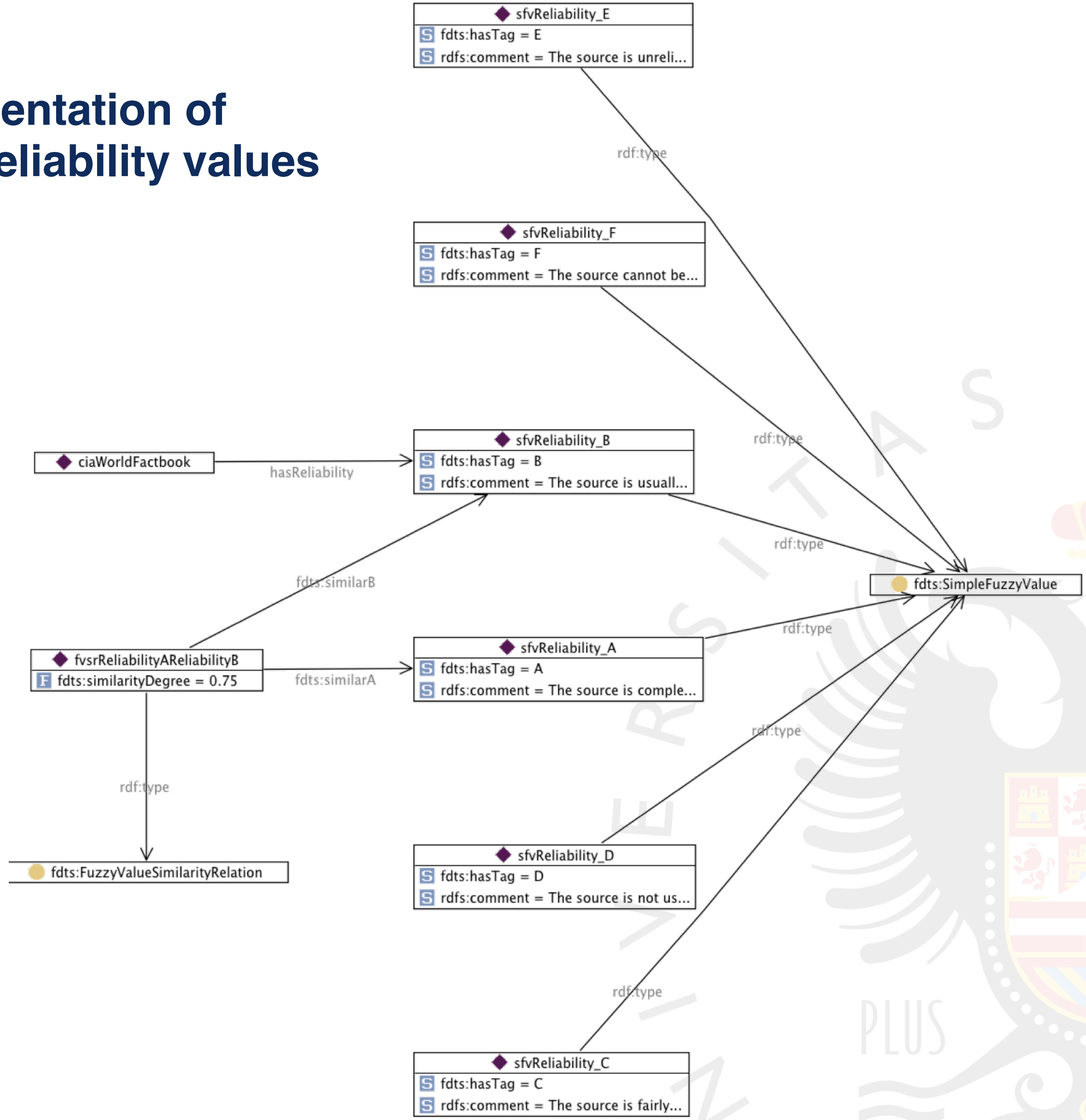
Extracted facts



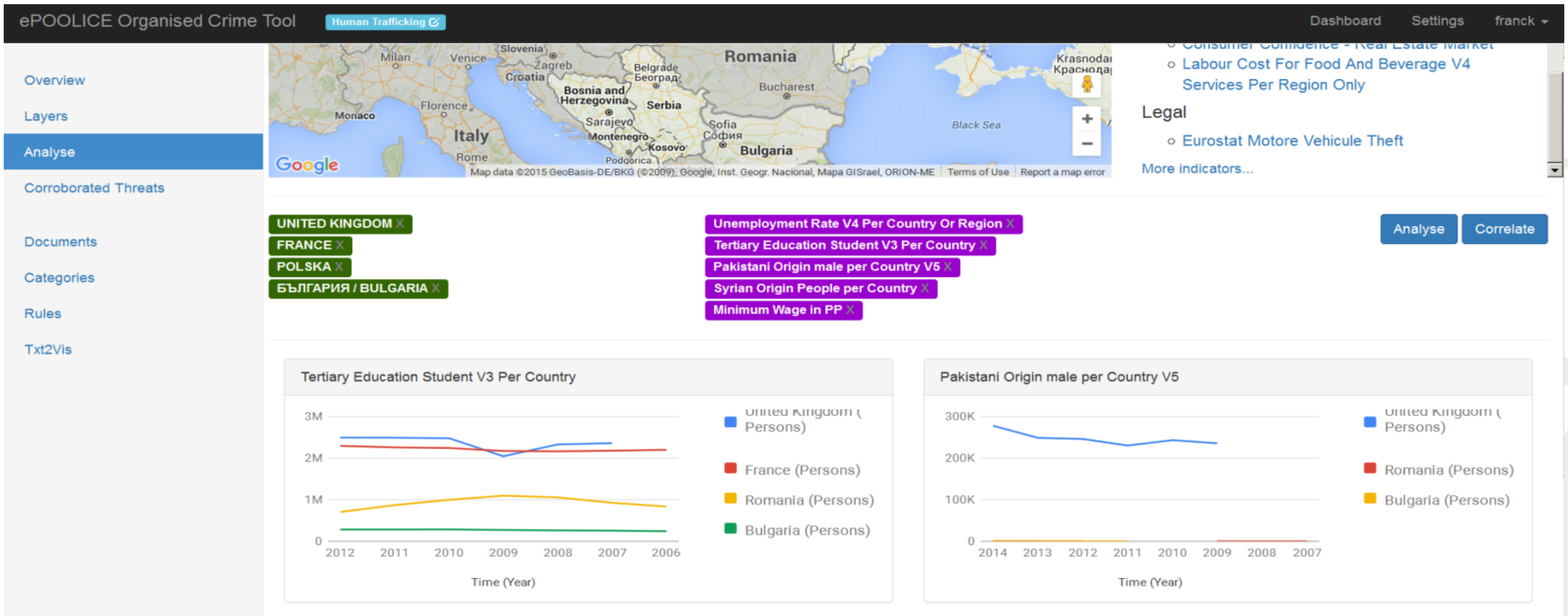
Extracted facts with probability values



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

COPKIT European Commission | Horizon 2020 European Union funding for Research & Innovation

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

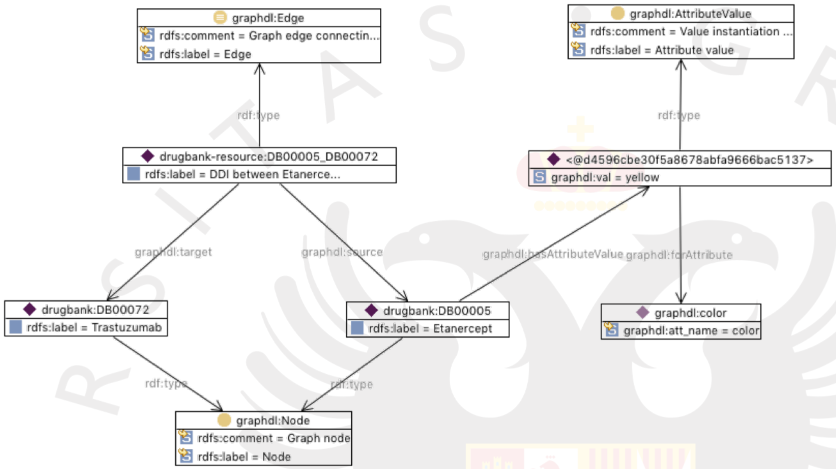
"solution for data access, virtualization, integration and multi-model relational database management (SQL Tables and/or RDF Statement Graphs)"

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)

J. Gómez-Romero, M. Molina-Solana, A. Oehmichen, Y. Guo (2018). *Visualizing large knowledge graphs: A performance analysis.* Future Generation Computer Systems 89, 224-238.

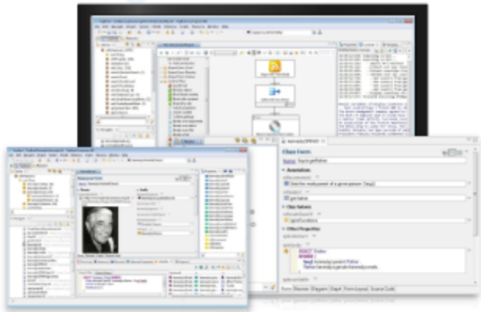
Topbraid Composer

https://www.topquadrant.com/tools/ide-topbraid-composer-maestro-edition/

TopQuadrant™

Solutions | Products | Services | Resources | Company | Technology

TopBraid Composer™ Maestro Edition



TopBraid Composer Maestro Edition™

Ready to build powerful semantic web and Linked Data applications for your enterprise? Let's get to work.

Buy Now »

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.

Download 30-Day FREE Trial »

Related Links

- TopBraid EDG — **Repository**
- TopBraid EVN — **Collaborative editor**
- TopBraid Live — **Web server**
- TopBraid Composer Standard Edition

Application Development Tools

- SHACL — **Data shapes**
- SPIN — **Rules**
- SPARQLMotion — **ETL**
- SPARQL Web Pages — **Info pages**
- SPARQL Web Application — **Endpoint**

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.

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

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

GraphQL

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

UNIVERSIDAD DE GRANADA

Thanks!

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

DECSAI