

Desafios de P&D e Cenários em LOD

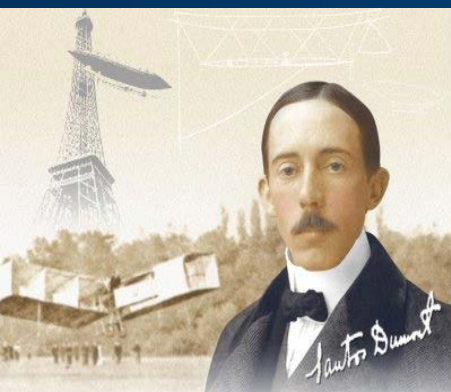
19 a 21 de Novembro de 2014

LOD Brasil, Florianópolis

Prof. José M Parente de Oliveira

Departamento de Ciência da Computação

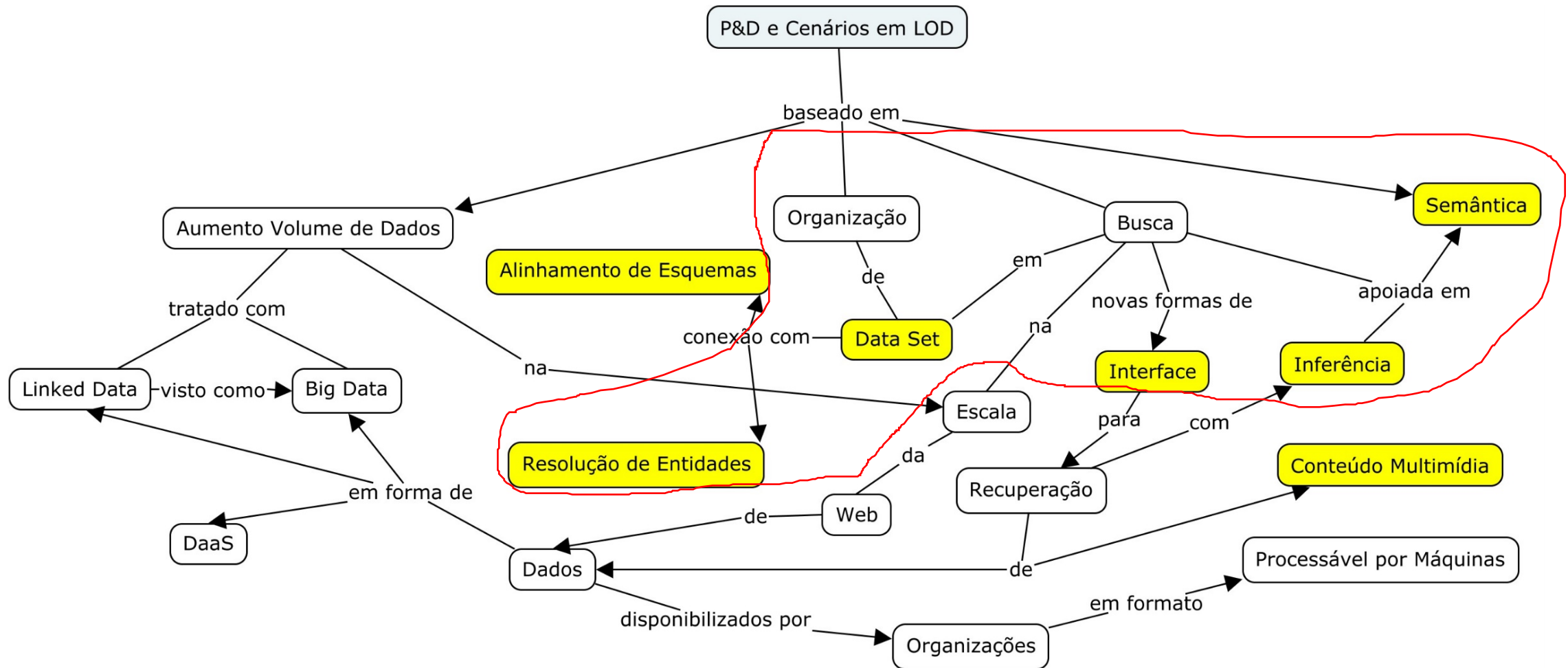
Instituto Tecnológico de Aeronáutica



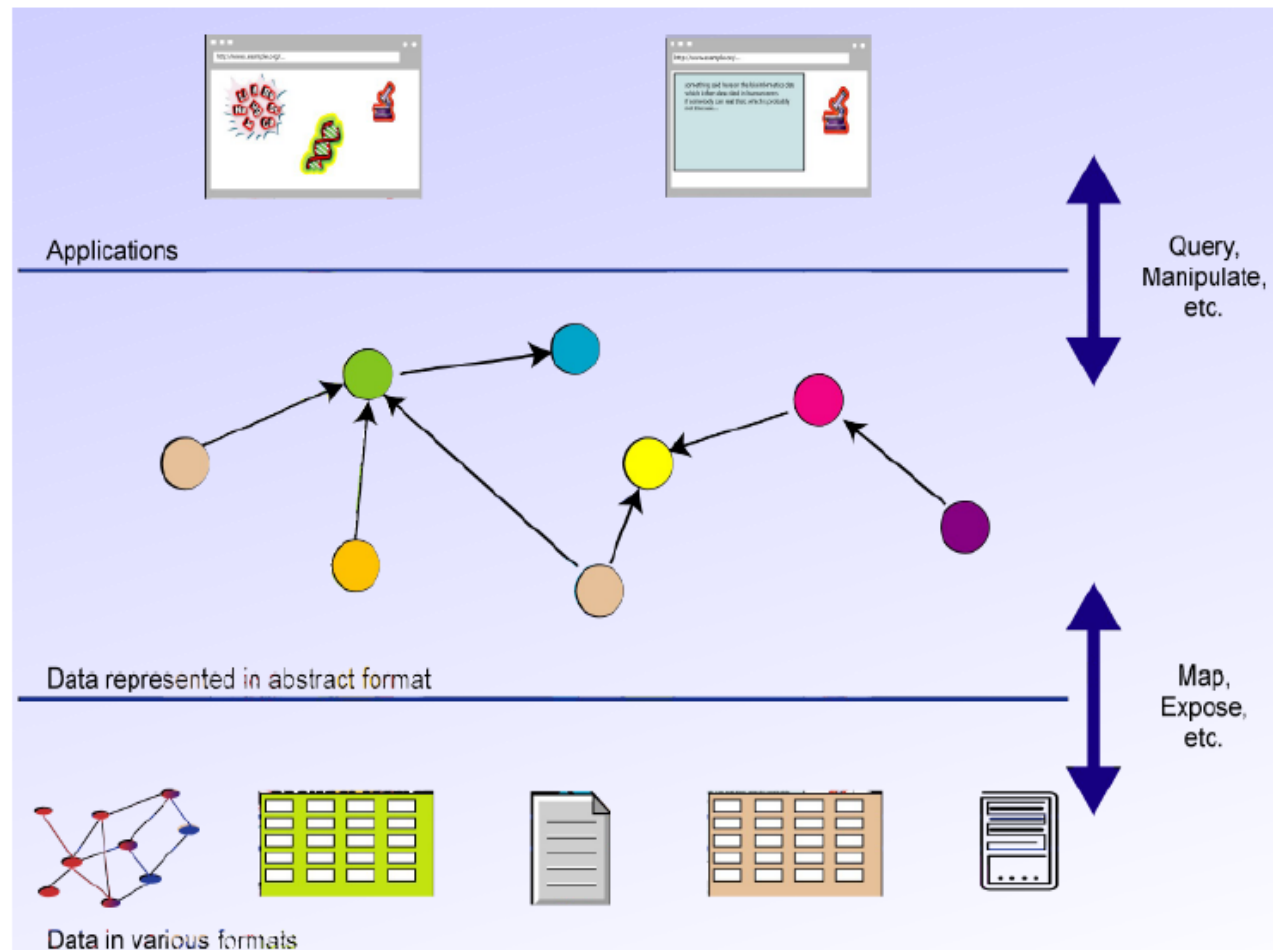
É tempo, talvez de se instalar uma escola de verdade em um campo adequado. Não é difícil encontrá-lo no Brasil. Nós possuímos, para isso, excelentes regiões planas e extensas, favorecidas de ótimas condições atmosféricas....



Desafios de P&D e Cenários em LOD



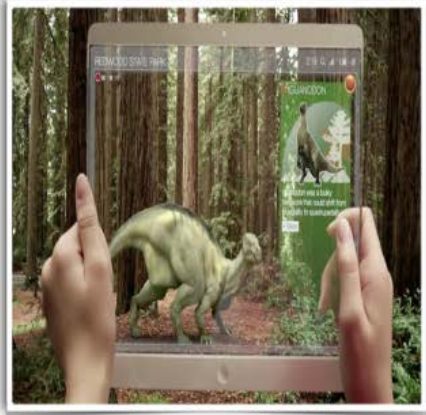
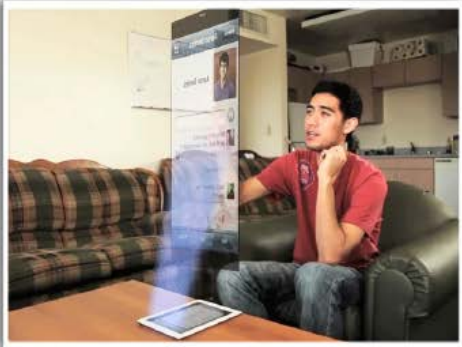
Visão da Web de Dados



Ivan Herman, W3C

Novas Interfaces

Interface Revolutions: **unprecedented access to information**



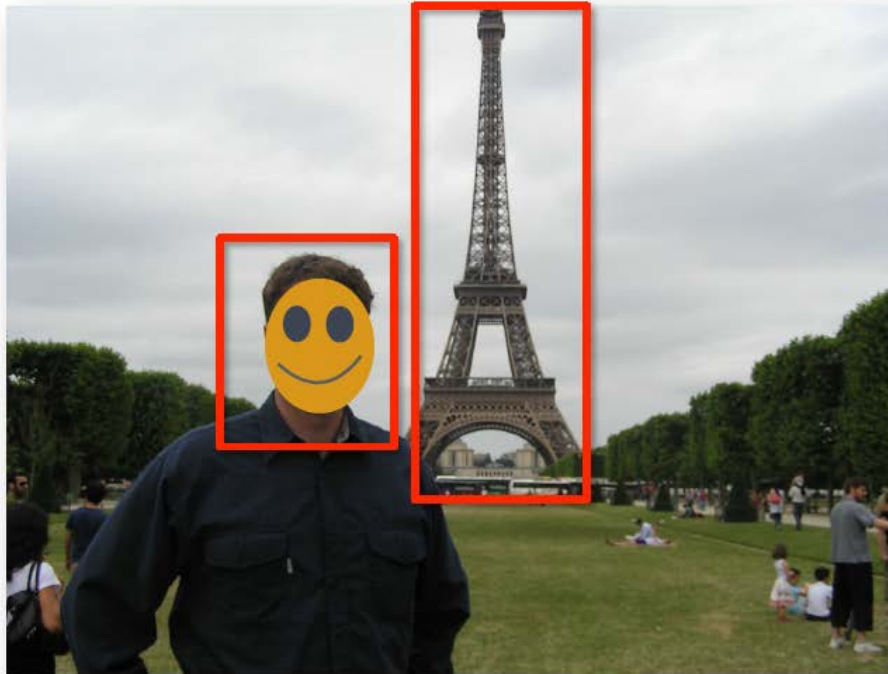
Gerd Leonhard, CEO The Futures Agency

Novas Interfaces



Jer Thorp

Armazenamento e Recuperação de Dados Multimídia



```
:image
:hasFragment
:image#xywh=...

:image#xywh=...
:subject foaf:Person

:image
:hasFragment
:image#xywh=..

:image#xywh=..
:subject
dbpedia:EiffelTower
(50%)
```

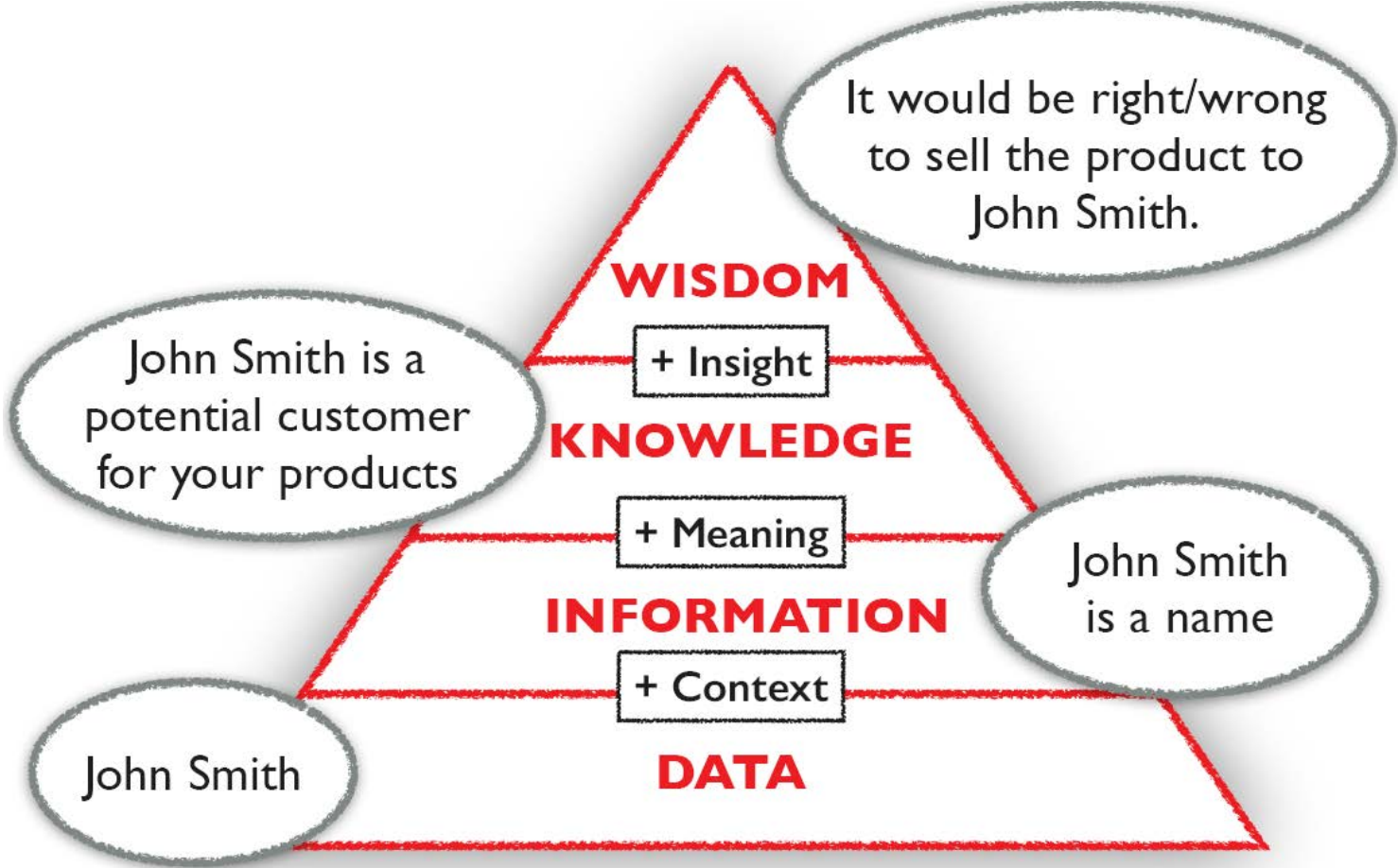
Title: „Me and the big thing“

Album: „A vacation in <http://dbpedia.org/resource/Paris>“

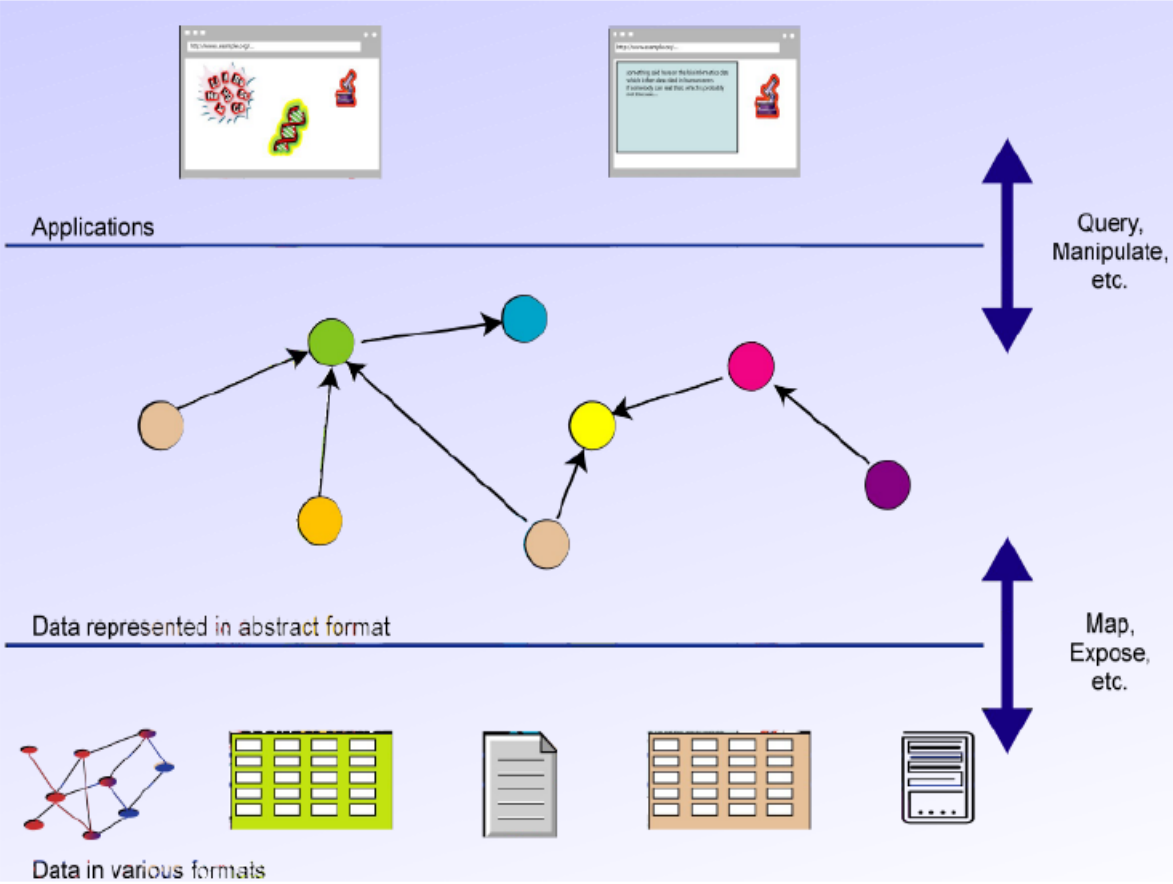
Author: „<http://tom-tester.org/me>“

RedLink

Uso de Semântica para Armazenamento/ Recuperação de Dados



Uso de Semântica para Armazenamento/ Recuperação de Dados



Ivan Herman, W3C

Uso de Semântica para Armazenamento/ Recuperação de Dados

Entity Resolution

Problem of identifying and linking/grouping different manifestations of the same real world object.
(Getoor and Machanavajhala, 2012)

Entity Disambiguation

Record Linkage

Duplicate Detection

Instance Matching



Data Deduplication

Reference Reconciliation

Record Matching

Object Identification

Identity/Coreference/Record Resolution

Data Matching

Uso de Semântica para Armazenamento/ Recuperação de Dados

- Many different vocabularies
- Scale
 - Aproximately 62 billion triples in the LOD Cloud (LODStats - <http://stats.lod2.eu/>)
 - 2.46 billion triples comes from DBPedia 3.9

| DBpedia | dbpedia:Statue_of_Liberty | Freebase | fb:m.072p8 |
|---|---|-----------------------|---|
| rdfs:label | Statue of Liberty, Freiheitsstatue, ... | fb:art_form | fb:m.06msq (Sculpture) |
| dbpprop:location | New York City, New York, U.S., dbpedia:Liberty_Island and | fb:media | fb:m.025rsfk (Copper) |
| dbpprop:sculptor | dbpedia:Frédéric_Auguste_Bartholdi | fb:architect | fb:m.0jph6 (F. Bartholdi), fb:m.036qb (G. Eiffel), fb:m.02wj4z (R. Hunt) |
| dcterms:subject | dbpedia_category:1886_sculptures, ... | fb:height_meters | 93 |
| foaf:isPrimaryTopicOf | http://en.wikipedia.org/wiki/Statue_of_Liberty | fb:opened | 1886-10-28 |
| dbpprop:beginningDate | 1886-10-28 (xsd:date) | YAGO | yago:Statue_of_Liberty |
| dbpprop:restored | 19381984 (xsd:integer) | skos:prefLabel | Statue of Liberty |
| dbpprop:visitationNum | 3200000 (xsd:integer) | rdf:type | yago:History_museums_in_NY, yago:GeoEntity |
| dbpprop:visitationYear | 2009 (xsd:integer) | yago:hasHeight | 46.0248 |
| http://www.w3.org/ns/prov#wasDerivedFrom | http://en.wikipedia.org/wiki/Statue_of_Liberty?oldid=494328330 | yago:wasCreatedOnDate | 1886-##-## |
| | | yago:isLocatedIn | yago:Manhattan, yago:Liberty_Island, |
| | | yago:hasWikipediaUrl | http://en.wikipedia.org/wiki/Statue_of_Liberty |

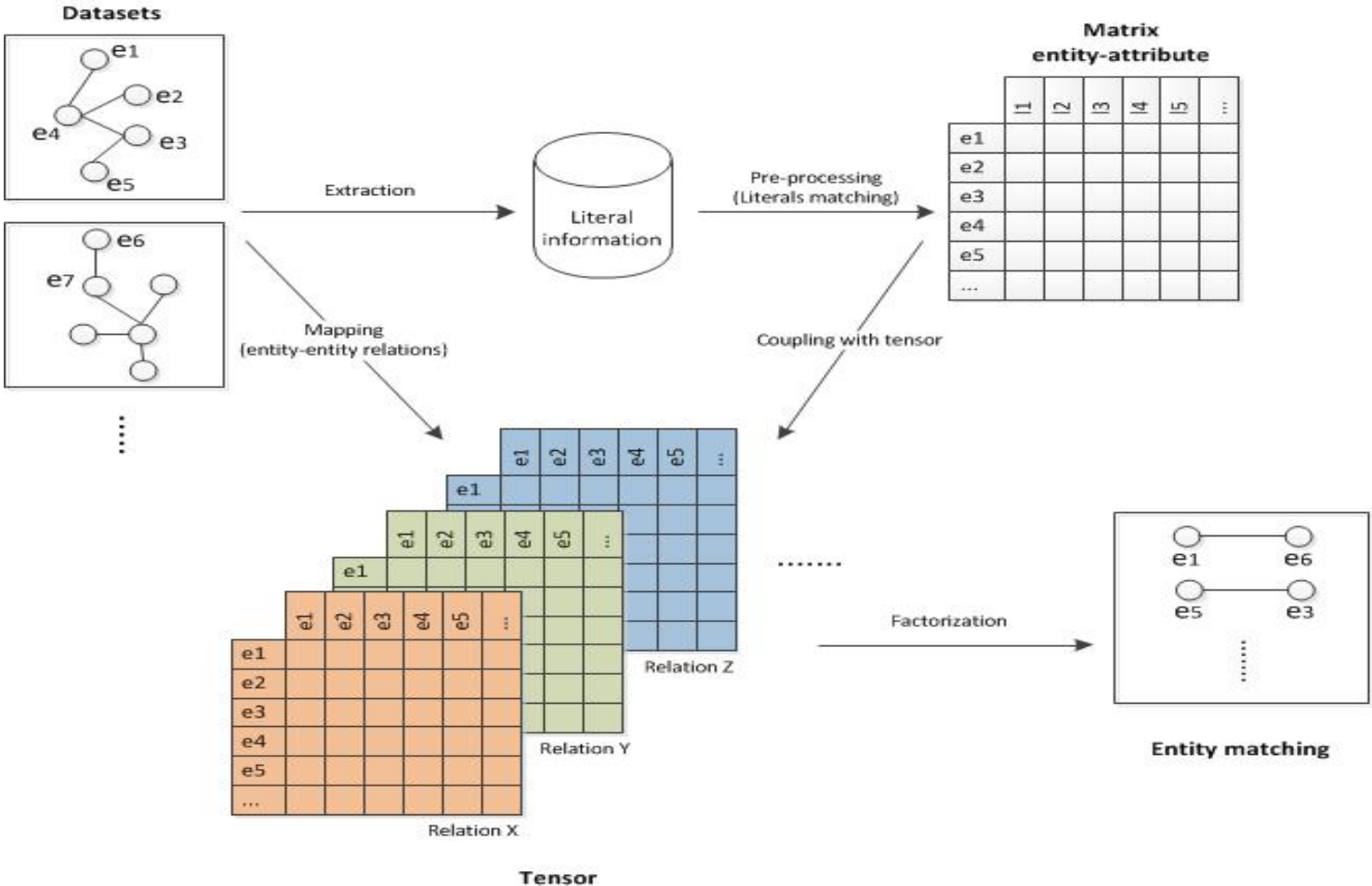
Noise
Quality of data

| DBpedia | dbpedia:Statue_of_Liberty | Freebase | fb:m.072p8 |
|---|---|-----------------------|---|
| rdfs:label | Statue of Liberty, Freiheitsstatue, ... | fb:art_form | fb:m.06msq (Sculpture) |
| dbpprop:location | New York City, New York, U.S., dbpedia:Liberty_Island and | fb:media | fb:m.025rsfk (Copper) |
| dbpprop:sculptor | dbpedia:Frédéric_Auguste_Bartholdi | fb:architect | fb:m.0jph6 (F. Bartholdi), fb:m.036qb (G. Eiffel), fb:m.02wj4z (R. Hunt) |
| dcterms:subject | dbpedia_category:1886_sculptures, ... | fb:height_meters | 93 |
| foaf:isPrimaryTopicOf | http://en.wikipedia.org/wiki/Statue_of_Liberty | fb:opened | 1886-10-28 |
| dbpprop:beginningDate | 1886-10-28 (xsd:date) | YAGO | yago:Statue_of_Liberty |
| dbpprop:restored | 19381984 (xsd:integer) | skos:prefLabel | Statue of Liberty |
| dbpprop:visitationNum | 3200000 (xsd:integer) | rdf:type | yago:History_museums_in_NY, yago:GeoEntity |
| dbpprop:visitationYear | 2009 (xsd:integer) | yago:hasHeight | 46.0248 |
| http://www.w3.org/ns/prov#wasDerivedFrom | http://en.wikipedia.org/wiki/Statue_of_Liberty?oldid=494328330 | yago:wasCreatedOnDate | 1886-##-## |
| | | yago:isLocatedIn | yago:Manhattan, yago:Liberty_Island, |
| | | yago:hasWikipediaUrl | http://en.wikipedia.org/wiki/Statue_of_Liberty |

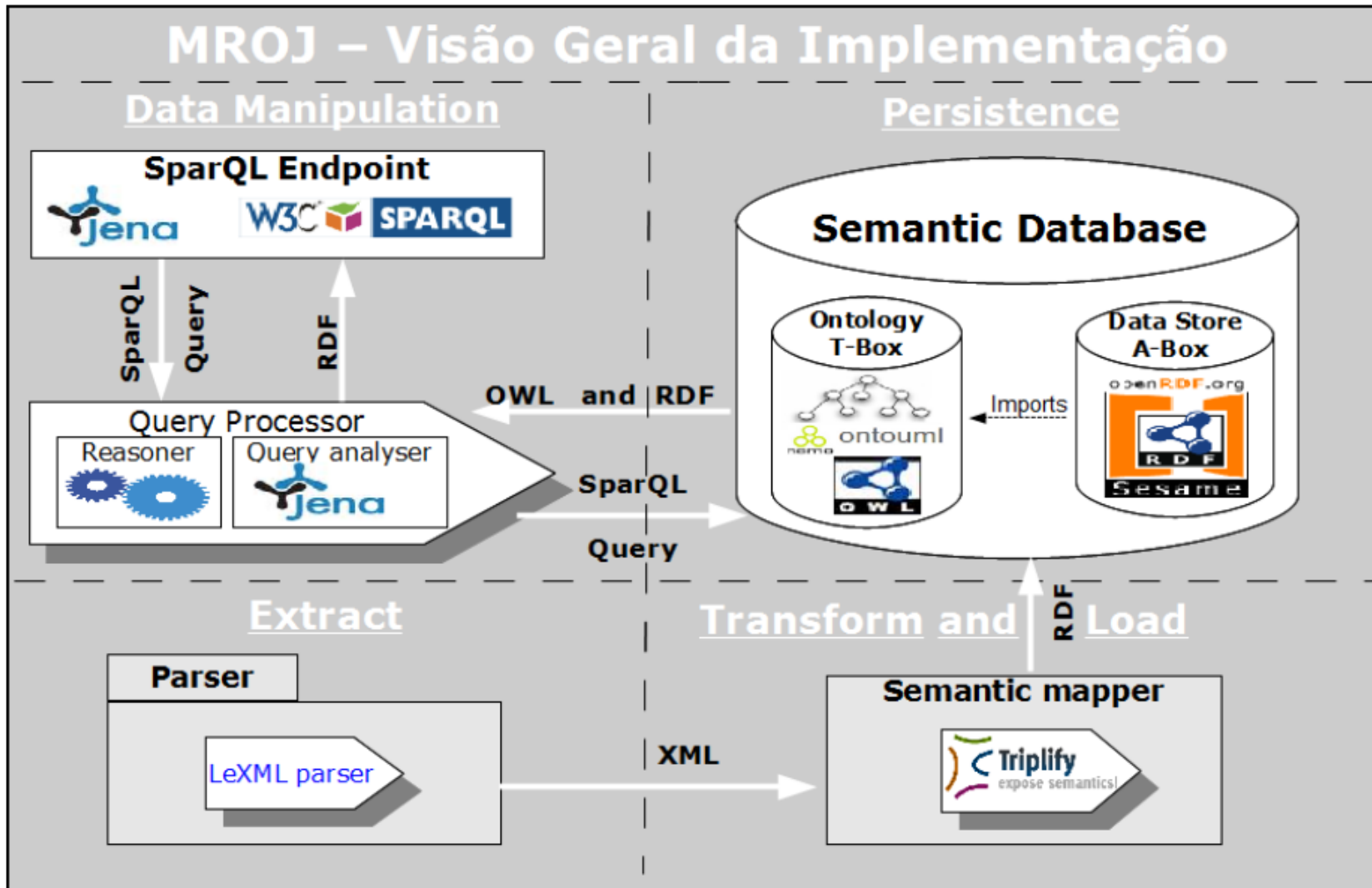
Inconsistency
Same entity description

Uso de Semântica para Armazenamento/ Recuperação de Dados

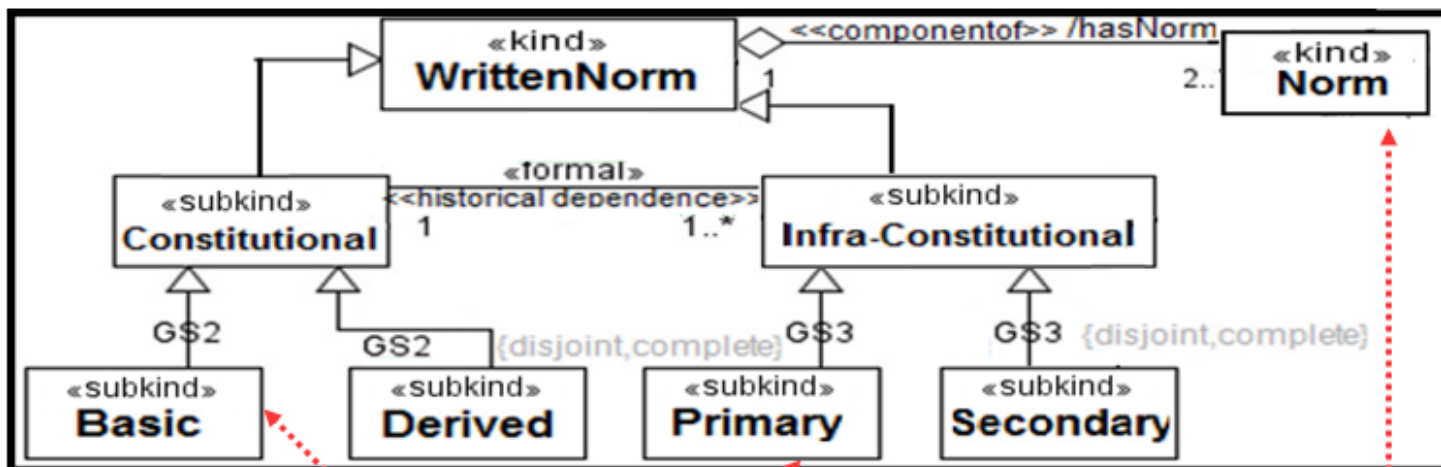
Relational Learning Approach



Semântica x Consulta SparQL



Semântica x Consulta SparQL



```

(text = "CONSTITUIÇÃO DA REPÚBLICA FEDERATIVA DO BRASIL DE 1988" )
...
(text = "Art. 37º" )
(text = "A administração pública direta e indireta de qualquer dos Poderes da União, dos Estados, do Distrito Federal e dos Municípios obedecerá aos princípios de legalidade, impessoalidade, moralidade, publicidade e eficiência e, também, ao seguinte:" )
(text = "XVI - é vedada a acumulação remunerada de cargos públicos, exceto, quando houver compatibilidade de horários, observado em qualquer caso o disposto no inciso XI:" )
(text = "a) a de dois cargos de professor;" )
(text = "b) a de um cargo de professor com outro, técnico ou científico;" )
(text = "c) a de dois cargos ou empregos privativos de profissionais de saúde, com profissões regulamentadas;" )
(text = "Art. 40." )
(text = "Aos servidores titulares de cargos efetivos da União, dos Estados, do Distrito Federal e dos Municípios, incluídas suas autarquias e fundações, é assegurado regime de previdência de caráter contributivo e solidário, mediante contribuição do respectivo ente público, dos servidores ativos e inativos e dos pensionistas, observados critérios que preservem o equilíbrio financeiro e atuarial e o disposto neste artigo." )
(text = "§ 11 - Aplica-se o limite fixado no art. 37, XI, à soma total dos proventos de inatividade, inclusive quando decorrentes da acumulação de cargos ou empregos públicos, bem como de outras atividades sujeitas a contribuição para o regime geral de previdência social, e ao montante resultante da adição de proventos de inatividade com remuneração de cargo acumulável na forma desta Constituição, cargo em comissão declarado em lei de livre nomeação e exoneração, e de cargo eletivo." )
...
(text = "LEI Nº 8.112, DE 11 DE DEZEMBRO DE 1990 " )
...
(text = "Capítulo III Da Acumulação" )
(text = "Art. 118." )
(text = "Ressalvados os casos previstos na Constituição, é vedada a acumulação remunerada de cargos públicos," )
(text = "§ 1º A proibição de acumular estende-se a cargos, empregos e funções em autarquias, fundações públicas, empresas públicas, sociedades de economia mista da União, do Distrito Federal, dos Estados, dos Territórios e dos Municípios." )
(text = "§ 2º A acumulação de cargos, ainda que lícita, fica condicionada à comprovação da compatibilidade de horários." )
(text = "§ 3º Considera-se acumulação proibida a percepção de vencimento de cargo ou emprego público efetivo com proventos da inatividade, salvo quando os cargos de que decorram essas remunerações foram acumuláveis na atividade." )

```

Perguntas Norteadoras do Painel

Quais os principais desafios de pesquisa na área de LOD no Brasil e no mundo?

Busca e inferência em grande número de repositórios de dados de diferentes escalas

How to construct a Semantic Web browser that effectively visualizes and navigates the huge connected RDF graph

Quais os principais desafios de desenvolvimento de repositórios e aplicações LOD no Brasil?

Aumentar o uso de semântica, resolução de entidades

How to establish trust and provenance of the content

Quais os cenários possíveis para utilização de LOD nos próximos anos?

Muitas aplicações empresariais e governamentais, como no domínio legal, dados científicos, etc.

Obrigado!

José M Parente de Oliveira
parente@ita.br