

# Structured data interoperability on the Web

François Scharffe  
**INRIA Grenoble Rhone-Alpes, France**

February 25, 2010  
Prague University of Economics

# Introduction

- ▶ The Web of data/Semantic Web is the future of structured data !
- ▶ I'll talk about. . .
  - ▶ structured data
  - ▶ convergence of structured data extraction and structured data publication (SW and linked-data)
  - ▶ Interoperability in the Web of data: issues, research and achievements
- ▶ I'll suppose that. . .
  - ▶ You have a relational database approach to structured data
  - ▶ You know about the Web
  - ▶ You've heard about the semantic Web: RDF, OWL, SPARQL

# Structured data and the Web

- ▶ Structured data is the future of the Web
- ▶ Many ongoing research aiming to extract structured data from the Web
  - ▶ Deep Web, spreadsheets, forms
  - ▶ Fusion tables
- ▶ Effort to provide formalisms, standards and tools to publish structured data and vocabularies: Semantic Web and linked-data
  - ▶ RDF, OWL, RDFa, RIF
  - ▶ Semantic Wiki

# The web of Data and the open linked data project

- ▶ Publish vocabularies
  - ▶ useful Web ontologies
- ▶ Publish data
  - ▶ RDF
  - ▶ dereferenceable URIs
  - ▶ SPARQL endpoints
  - ▶ Link to existing data



# Interoperability issues

## Vocabs

- ▶ Vocabularies should be reused, but anyone is free to provide its own
- ▶ Any published data sources might use its own schema as an ontology
- ▶ Results in many ontologies/vocabularies available and the need to connect them

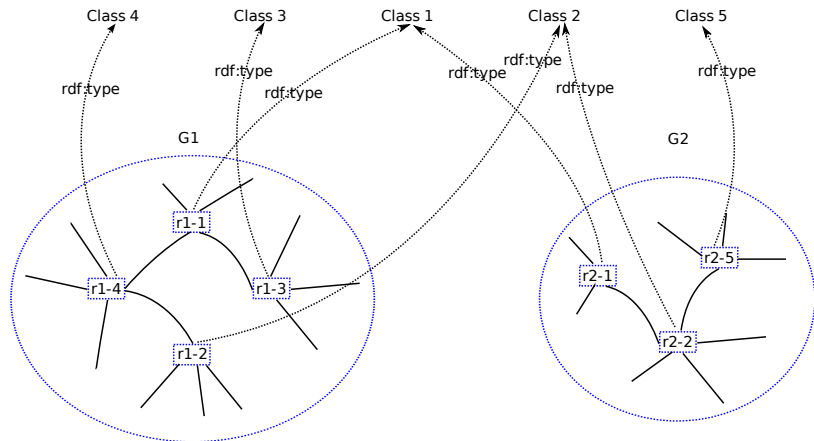
## Data

- ▶ Even when a same ontology is used, if two data-sources are published, they might contain equivalent resources
- ▶ Need to interconnect these resources

# Interconnecting vocabularies

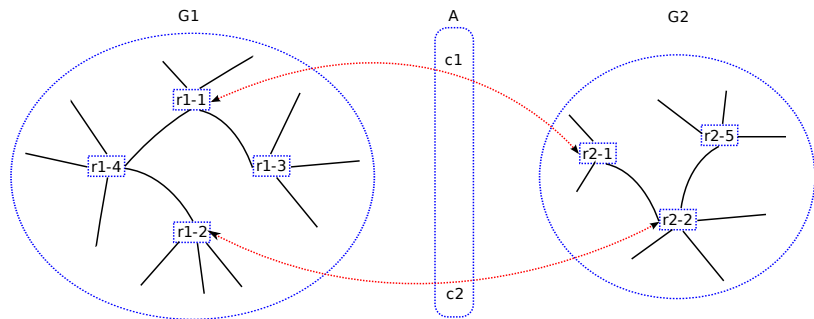
- ▶ Ontology matching and alignment is the key to issue one:
  - ▶ Matching algorithms
  - ▶ GUIs
  - ▶ alignment representation
  - ▶ patterns
  - ▶ mediators
- ▶ Results: Alignment API and server, Ontomap, OAEI, Alignment patterns language
- ▶ Data linkage is the key to issue 2
  - ▶ Matching algorithms
  - ▶ Link specification languages
  - ▶ Linksets
- ▶ Results: Silk, Knofuss, RDF-AI, VOID

# Linking Data

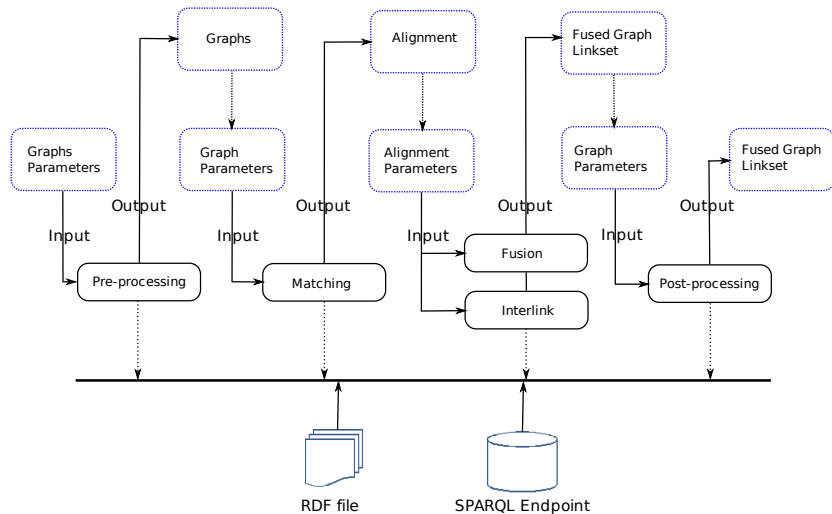




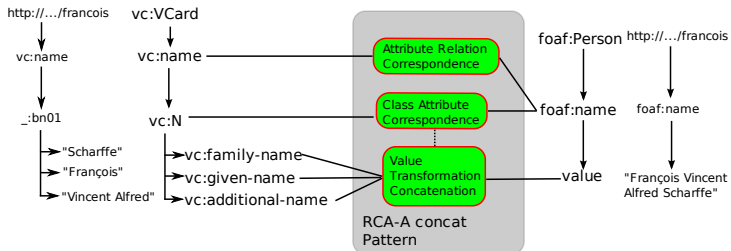
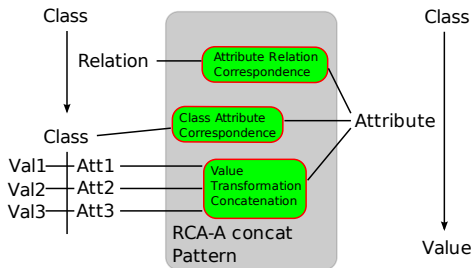
# Linking Data



# Linking Data



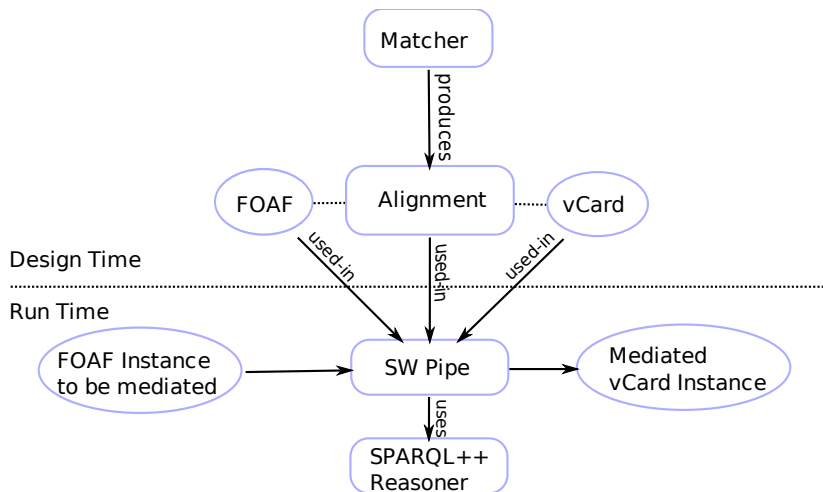
# Aligning Vocabularies



## Alignment representation

```
vcard:family-name a align:Property;
  align:bind-with "?x".
vcard:given-name a align:Property;
  align:bind-with "?y".
vcard:additional-name a align:Property;
  align:bind-with "?z".
foaf:name a align:Property.
:property-group a align:Property;
  or [
    align:item vcard:family-name;
    align:item vcard:given-name;
    align:item vcard:additional-name.
  ].
align:transf [
  align:transf-id transf:concat;
  align:transf-param "?x ?y ?z".
].
:foaf-vcard-names a align:Cell;
  align:entity1 :property-group;
  align:entity2 foaf:name.
```

# Using alignments



# Conclusion

- ▶ The data representation is there
- ▶ Data integration works through owl:sameAs
- ▶ Vocabularies integration works though there is no standard representation
- ▶ Where is the killer app ?