Language Technologies for Europe:

Public Sector, new approaches for old problems

Alan Mas Soro (everis)
Enric Staromiejski (everis)
Stefanos Vrochidis (CERTH)
I. Kompatsiaris (CERTH)
Public Sector, new approaches for old problems

A "stand-still" problem in Public Sector

⇒ Most platform's document still 100% narrative

⇒ "Structured" documents contain narrative/non-structured information elements

⇒ Current controlled vocabularies are insufficient

⇒ New "modes/channels" starting to be used
Public Sector, new approaches for old problems

Real cases of application in Public Sector: Using Language technologies to break the e-Procurement “stand-still inertia”

• Objectives:

1. For existing narrative multimodal information sources: extract data from them and structure it based on the current e-Procurement standard vocabularies (e.g. OASIS UBL)

2. For new document drafting: Facilitate the user the generation of a kind of narrative information that can be easily (i.e. automatically) transformed into structured data (e.g. special Procurement document drafting assistants)

3. e-Proc. Use Cases: Use the structured data in combination with knowledge/intelligent artefacts to generate added value services (e.g. matching offer and demand. Need for new artefacts: Terminologies, Ontologies, Landscapes, other).
Public Sector, new approaches for old problems

Public Innovation e-Procurement platform (pilot proposal, based on MS technologies)

Evolve language technologies for Public Sector procurement purposes.
Public Sector, new approaches for old problems

Public Innovation e-Procurement platform (pilot proposal, based on MS technologies)

Captures data from different linked sources containing both structured and narrative information.

Analyses

Helps the user to build new contents in a very user-friendly manner. Supports the user on decision-making operations.

Assists

Independently of the incoming format it transforms it into structured queryable data. Produces logical inference axioms.

Structures

Integrates the inputs and feeds-back from its own outcomes, the transactional operations and the user behavioural analysis (data access and interoperability as source of knowledge).

Learns

Automated operations are launched to provide added-value services such as the discovery of good opportunities for Economic Operators and Contracting Authorities, and many more.

Automates

The PIeP essentials are (1) capturing, (2) structuring and (3) analysing data, most of them generated during the public procurement activity, published in different formats and EU languages.
Public Sector, new approaches for old problems

Public Innovation e-Procurement platform (pilot proposal, based on MS technologies)

**Possible services:** The amount of information generated in the public procurement allows to design a platform with an ambitious functional scope. Its implementation could be limited to a pilot to guarantee automatic and intelligent intersection of supply and demand, as higher value-added service for system agents and basis for the development of downstream services.

- Discover innovative companies and innovative products and services
- Compare products and services
- Knowing best practices from other CA

**Intelligent demand**

**Contracting Authorities (CA)**

**Market analysis (offer)**

**Market analysis (demand)**

**Economic Operators (EO)**

**Intelligent offer**

- Meet demands of CA that can satisfy with my products and services
- Study the CA
- Studying the range of products and services from other companies, and similar experiences

- Post my needs
- Discover CA with my same needs
- Facilitate the purchase grouped with other CA
- Receive proposed solutions that help configure my final demand

**Process automation**

- Introduce my products and services
- Discover "the good chance"
- Get in touch with potential "partners"
- Receiving "feedback" on my offer
Public Sector, new approaches for old problems

Elements of the Platform

**USER ASSIST.**
- A1: CA Profile Tool
- A2: EO Profile Tool

**BASES**
- SQL: Operational
- RDF: Triple Graphs
- CMS: Documents & Artefacts

**MODULES**
- M1: Contextual Analysis
- M2: Automatic Translation
- M3: Conceptual Analysis
- M4: Integration & Enrichment
- M5: Structuring & Persisting
- M6: Data matching + Inferences
- M7: Weighting & Summarizing

**INTEROP**
- I1: Between Platforms
- I2: Between Layers
- I3: Between Modules
- I4: With 3rd Parties

**PRESENTATION**
- P1: CA web site
- P2: EO web site
- P3: Sindication (pull)
- P4: Notification (push)
- P5: SPARQL End-point
- P6: UE Federation
- P7: Linked Open Data
- P8: RESTFUL Services
Use Case: discovering good opportunities for the EO

Public Sector, new approaches for old problems

- CA Profiles
- EO Profile
- Tenders
- Crawlers
- CMS
- M1
- M2
- M3
- M4
- M5
- M6
- M7

Inputs:
- CA Profiles
- EO Profile
- Tenders
- Crawlers

Outputs:
- RDF
- CMS

Process:
1. Context Analysis
2. SQL
3. Data matching & inferences
4. Data structuring & persisting
5. RDF
6. New axioms
7. Integration & enrichment
8. RDF
9. New axioms
10. Concept extraction
11. Yes
12. Translation
13. Yes
14. New axioms
15. Yes
16. RDF
17. New axioms
18. Concept extraction in narrative fields (e.g. EO activities)
19. Yes
20. Automatic translation
21. Yes
22. Weighting and summarising (english)
23. Found Opportunities (summarised and weighted)
Public Sector, new approaches for old problems

Summary of LT-related needs in e-Procurement

- **MULTILINGUALISM**: Huge need for translating user interfaces, structured transactional documents, and large narrative documentation into all the EU official languages. Translations are currently done manually. Automatic translation is perceived as a powerful tool for a sound pan-European e-Procurement.

- **CONCEPT EXTRACTION**: Most information that is narrative could be structured into existing standard e-Documents. Concept extraction could facilitate this task (Need for an Ontology of the Public Procurement).

- **MATCHING OFFER AND DEMAND**: Both CA and EO have to identify the good business opportunities manually. A Matching engine could be built based on the automatic extraction and comparison of essential concepts.

- **SUMMARIZATION**: Presenting accurate summaries of the essential content of a Call for Tender (and vice-versa of the essence of a Tender) is critical for swift and balanced man-made decisions.
MULTISENSOR

- FP7 research project
  - Develops user-driven LT technologies and solutions
  - Supports e-procurement
    - Multilingual analysis and machine translation tools
    - Semantic analysis and concept extraction tools
    - Semantic reasoning oriented decision support and risk assessment
    - Decision support and risk assessment for business opportunities
    - Summarisation of multilingual market information

- http://www.multisensorproject.eu/
Thank you!

Supported by MULTISENSOR project
http://www.multisensorproject.eu/