Case-oriented Semantic Enrichment of Bibliographic Entities
Joffrey Decourselle

To cite this version:
Joffrey Decourselle. Case-oriented Semantic Enrichment of Bibliographic Entities. Theory and Practice of Digital Libraries, Sep 2016, Hannover, Germany. hal-01346830v2

HAL Id: hal-01346830
https://hal.archives-ouvertes.fr/hal-01346830v2
Submitted on 3 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Context: Evolution of Digital Libraries towards Semantic technologies requires to migrate and enrich the existing bibliographic records. Yet building these processes is still a very challenging task.

Semantic Enrichment Process:

1. **FRBRisation**
   - Migration Rules
   - 101 $a fre $c eng
   - 200 $a Da Vinci code
   - $g [trl by Daniel Roche]
   - 454 $t The Da Vinci code
   - 700 $a Brown $b Dan

2. **Expression**
   - Work
   - Agent
   - Name: Dan Brown
   - Translator: Daniel Roche
   - Language: French
   - Expression: Da Vinci Code
   - Language: English

3. **Enrichment**
   - HasTitle: Da Vinci Code
   - Realized Through: Translation
   - Translator: Daniel Roche

4. **Enrichment**
   - Expression
   - Agent
   - Name: Angels and Demons
   - Language: French
   - Spoken word

5. **Enrichment**
   - Place
   - Work
   - Agent
   - Case Publication

Contributions:

- Benchmark for the interpretation of bibliographic records
- Design of a Case-oriented model
  - Based on FRBR and bibliographic patterns
  - Hierarchical dependency between cases
- Case-based enrichment process with Linked Open Data

Challenges:

- Tuning of FRBRisation tools
- Information Extraction
- Entity Matching
- Knowledge Fusion

Focus on the Case-oriented model:

Characteristics of a case:

- Describes a bibliographic pattern
- Contains conditions and mappings
- Managed independently

Benefits of the model:

- Eases the design of migration rules
- Prevents useless computations
- Facilitates the extraction of entities from catalog or external sources

References: