



SIERA: A Multilingual Knowledge Sharing Project

Pr. Christophe Roche

- Universidade Nova de Lisboa, CLUNL - Portugal
- Université de Savoie, Equipe Condillac – France

roche@univ-savoie.fr



■ **SIERA**



Full title: Integrating **S**ina **I**nstitute into the **E**uropean **R**esearch **A**rea

Duration: 3-years FP7 Project

Partners:

- **Birzeit University, Palestine (Grant holder and Coordinator)**
- *Mustafa Jarrar & Adnan Yahya*
- **New University of Lisbon, Portugal**
- *Christophe Roche (University of Savoie, France) & Rute Costa*
- **Berlin-Brandenburg Academy of Sciences**
- *Christiane Fellbaum (Princeton University, USA)*
- **University of Trento, Italy**
- *Paolo Bouquet*
- **University of Milano-Bicocca, Italy**
- *Carlo Batini & Matteo Palmonari*

Project Associates:

- VU University Amsterdam
- Michael Culture
- Centre for Cultural Heritage Preservation– Bethlehem
- The Ministry of Telecom and Information Technology (Palestine)
- The Egyptian Engineering Company for Developing Digital Systems

Work packages:

WP 1	Research Strategy
WP 2	Networking and Research Setup
WP 3	Capacity Building in FP7
WP 4	Dissemination and Exchange of Knowledge
WP 5	Project Management and Quality Control

■ SIERA



Description:

The rapid growth of cross-border markets and global concerns are creating a huge demand to facilitate **knowledge sharing between societies**. Digitizing information and enabling this information to be accessed and understood by other communities is of fundamental importance. **The diversity of languages, cultures and standards are the main barriers to sharing and consuming knowledge.**

Objectives:

The general objective of the project is to **reinforce closer scientific cooperation between EU and Palestinian scientists** in the field of **multilingual and multicultural knowledge sharing technologies**.

■ SIERA



Key Idea:

Enhancing large-scale knowledge sharing portals with **Arabic language** and **Arabic content**.

Application Domains:

- **MICHAEL**: a cultural heritage portal which provides a multilingual service to explore/search the European cultural heritage
- **KYOTO**: a wiki-portal about environment and ecology.



■ Issues



1. Investigating Arabic support in multilingual knowledge sharing portal

- Arabic APIs
- Extending the Michael portal with Arabic search functionalities

■ Issues



2. Extending multilingual thesauri/ontologies for culture and ecology domains

2.1 Extending the Michael thesaurus

2.1.1 Arabic localization of the Michael thesaurus

2.1.2 Integrating the Bethlehem thesaurus

2.2 Extending the Kyoto thesaurus

Mapping Kyoto and Arabic Ontology

■ Issues

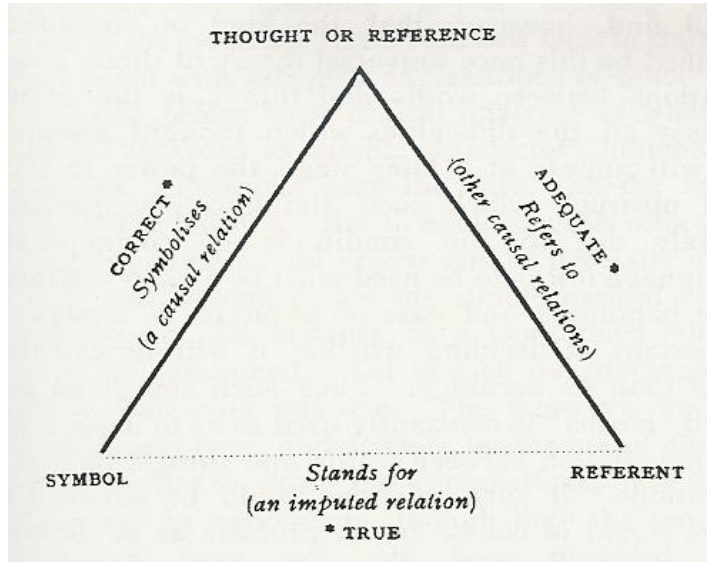


3. Establishing a framework for mapping between Wordnet and Arabic Ontology

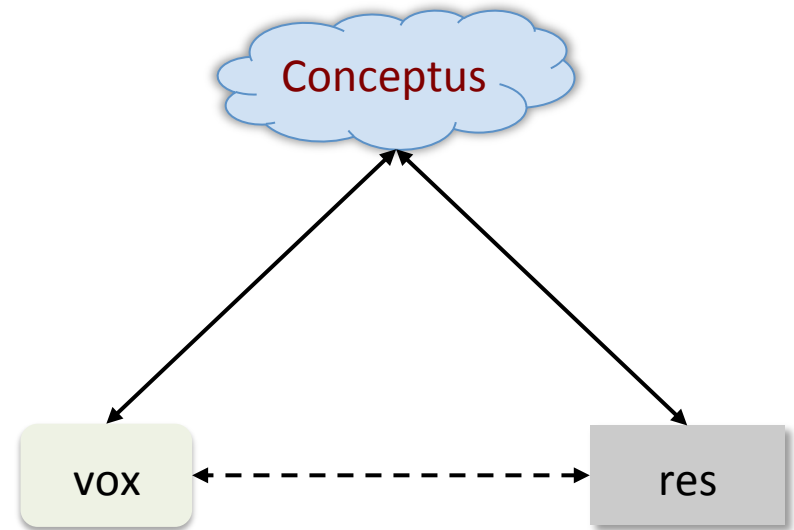
Introduction of the notion of "Semantonym"

4. Resolving and linking entities and identities

■ Key Principle




Ogden & Richards' Semantic Triangle



Scholastic Semantic Triangle

Arabic Ontology



The screenshot shows the website for Arabic Ontology at sites.birzeit.edu/comp/ArabicOntology/. The page features a navigation menu with links for About, People, Publications, News & Events, Downloads, and Online Search. A central diagram illustrates the ontology's structure, with 'جدول' (Table) at the center, branching into 'مصفوفة' (Matrix) and 'أجندة' (Calendar). 'مصفوفة' further branches into 'ترتيب' (Ordering) and 'نهر' (River). 'أجندة' branches into 'جدول دوري' (Periodic Table) and 'جدول بأيام السنة مرتب حسب تعاقبها الزمني' (Table of the year's days ordered by their chronological sequence). 'نهر' branches into 'نهر مائي كبير يتدفق بمجرى' (Large water river flowing in a course) and 'نهر مائي صغير' (Small water river).

About Arabic Ontology

The Arabic Ontology is a formal representation of the concepts that the Arabic terms convey. For each term in the Arabic language, the set of its meanings (i.e. concepts) are identified, and semantic relationships (such as subtype-of and part-of) between all concepts are introduced. For simplicity, the Arabic ontology is a tree of the meanings of the Arabic terms (see the diagram below). The Arabic Ontology can be seen as an Arabic WordNet; however its relationships are well-formalized, and glosses follow strict formulation and ontological rules.

News & Events

- The Arabic Ontology Project receives seed funding from Birzeit University
- First Arabic Ontology Volunteers Meeting
- Dr. Jarrar talk about the Arabic Ontology at Birzeit IT Expo 2010

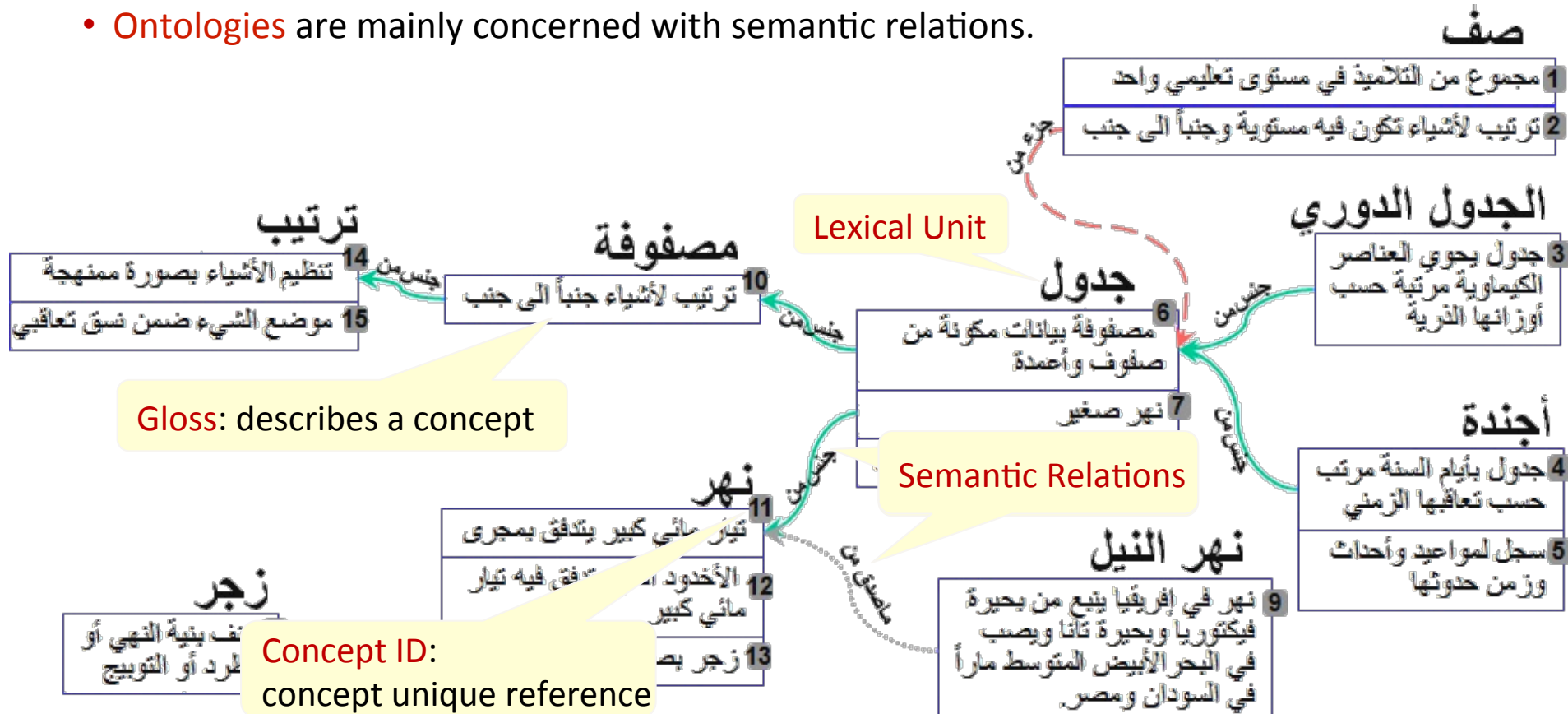
■ Arabic Ontology

- A project started in 2010, at Birzeit University, Palestine
- The Arabic Ontology is more than an Arabic WordNet
- **Unlike WordNet, the Arabic Ontology is logically and philosophically well-founded, as it follows strict ontological principles. Nevertheless it Arabic Ontology can be used as an Arabic WordNet.**

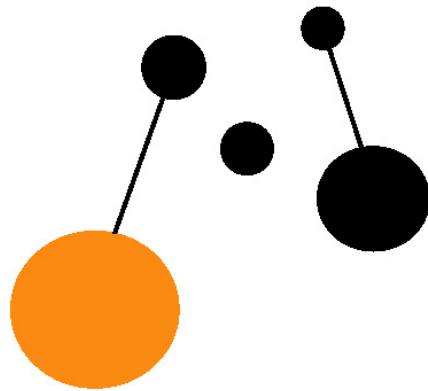
“The Arabic Ontology is a formal representation of the concepts that the Arabic terms convey. For each term in the Arabic language, the set of its meanings (i.e. concepts) are identified, and semantic relationships (such as subtype-of and part-of) between all concepts are introduced”

Arabic Ontology

- Semantic relations** are relationships between concepts (not words), e.g., subtype, part-of, etc.
- Lexical relations** are relationships between words (not concepts), e.g., synonym-of, root-of, abbreviation-of, etc.
- Ontologies** are mainly concerned with semantic relations.



■ Michael



Michael

Multilingual
Inventory
of Cultural Heritage
in Europe

<http://www.michael-culture.org>

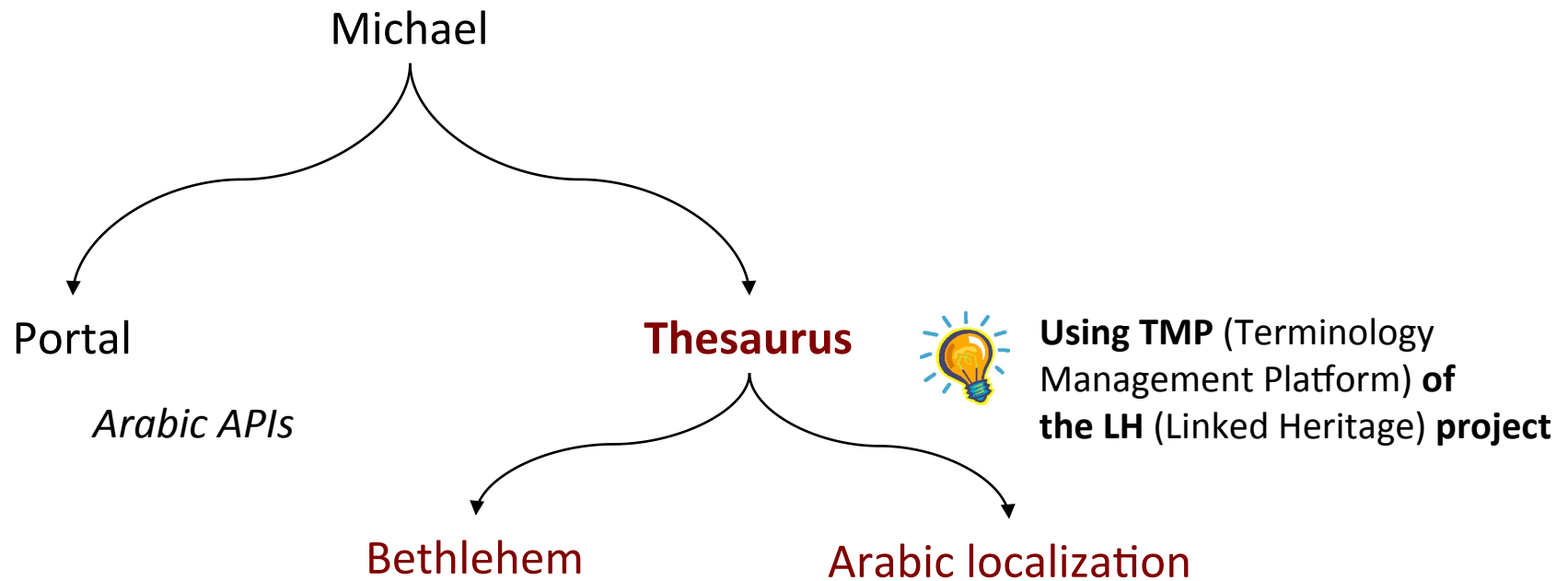


The MICHAEL Project is funded under the European Commission eTEN Programme

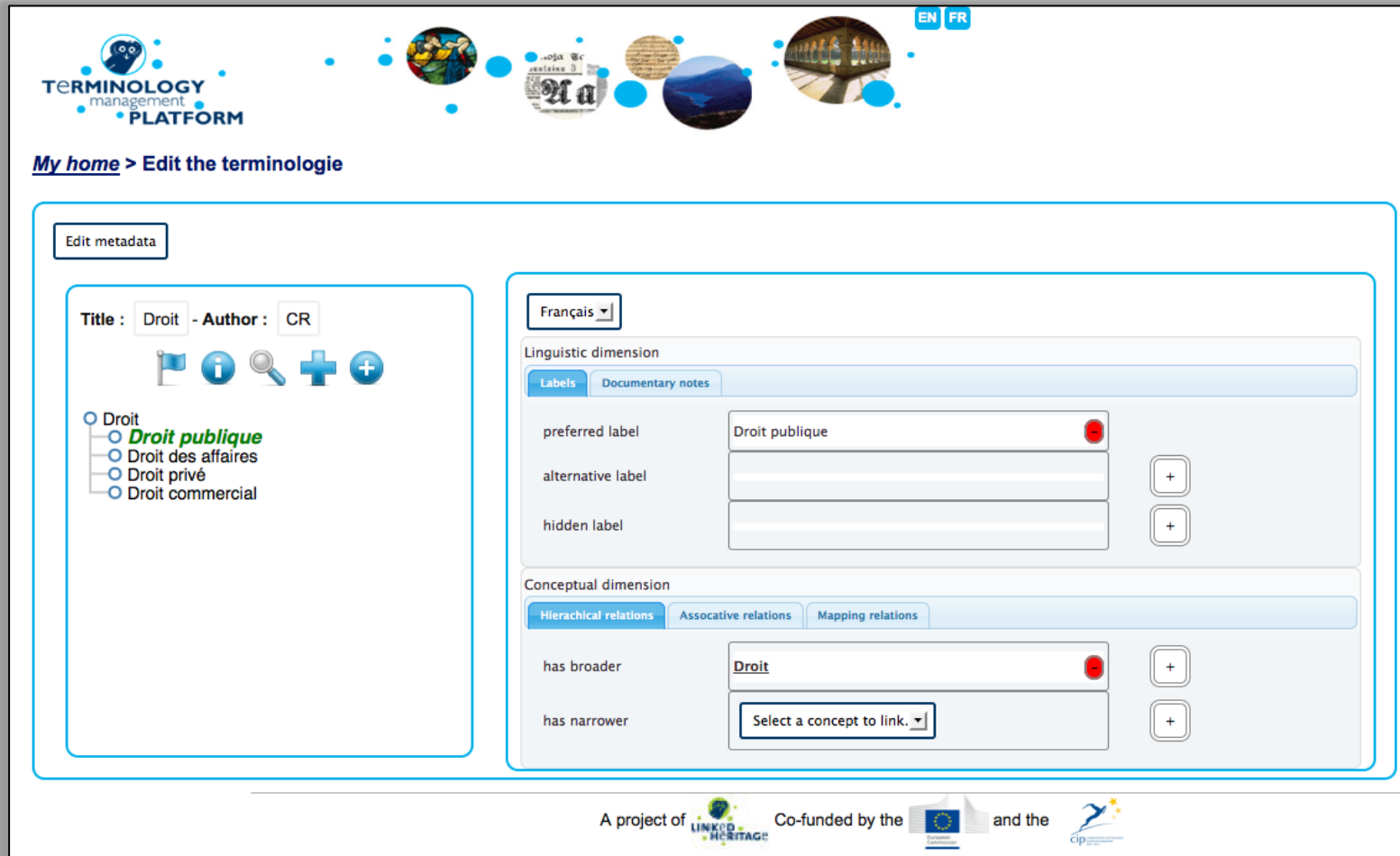


■ Michael

- Extending Michael's Portal functionalities taking into account Arabic language
- Extending Michael's Thesaurus with content from Bethlehem Museum
- Arabic localization of Michael's Thesaurus



Terminology Management Platform



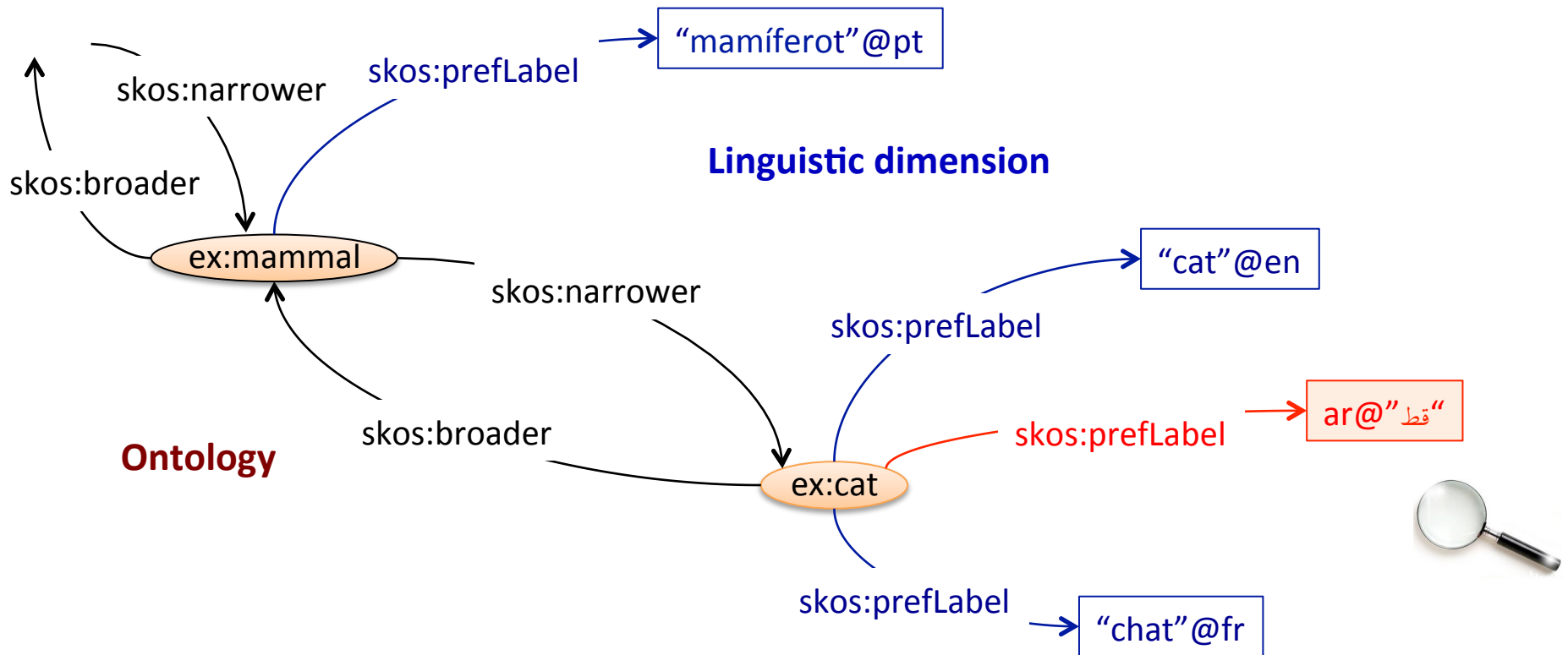
The screenshot shows the 'Edit the terminologie' interface. At the top left is the 'TERMINOLOGY management PLATFORM' logo. Below it is a navigation breadcrumb: 'My home > Edit the terminologie'. The interface is divided into two main sections. The left section, titled 'Edit metadata', contains a 'Title' field with 'Droit' and an 'Author' field with 'CR'. Below these are icons for a flag, information, search, and two plus signs. A tree view shows a hierarchy: 'Droit' (selected) with sub-items 'Droit publique', 'Droit des affaires', 'Droit privé', and 'Droit commercial'. The right section is for editing the concept 'Droit' in 'Français'. It has tabs for 'Linguistic dimension' (with sub-tabs 'Labels' and 'Documentary notes') and 'Conceptual dimension' (with sub-tabs 'Hierarchical relations', 'Associative relations', and 'Mapping relations'). Under 'Linguistic dimension', there are fields for 'preferred label' (containing 'Droit publique'), 'alternative label', and 'hidden label', each with a plus sign button. Under 'Conceptual dimension', there are fields for 'has broader' (containing 'Droit') and 'has narrower' (containing a dropdown 'Select a concept to link...'), each with a plus sign button. At the bottom of the interface, it states 'A project of LINKED HERITAGE Co-funded by the European Union and the CIP' with corresponding logos.

➔ A SKOS-oriented Platform



■ Arabic localization of Michael

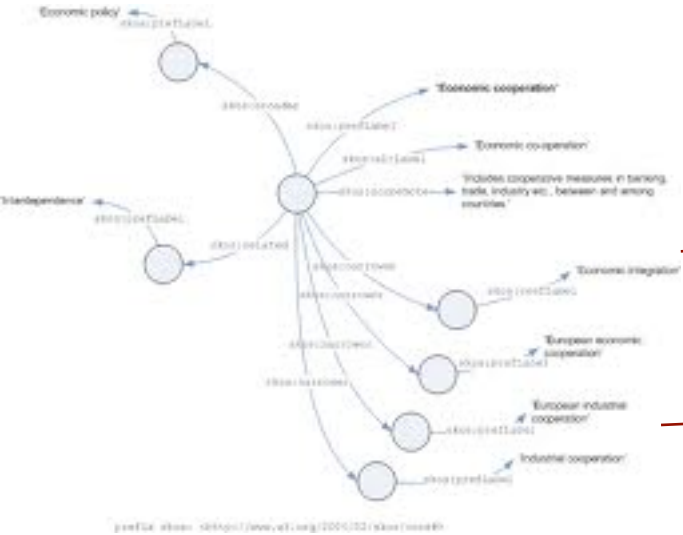
➔ SKOS prefLabel and altLabel



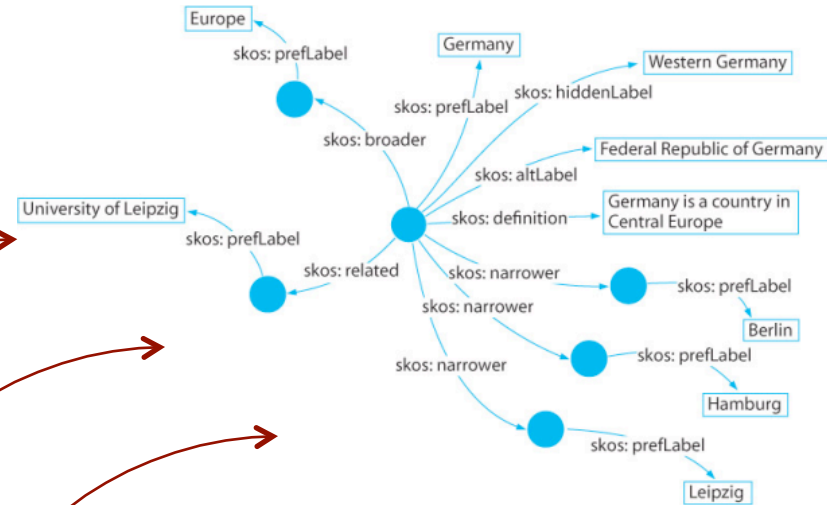
■ Extending Michael Thesaurus with content from Bethlehem

➔ SKOS semantic relationships

Bethlehem thesaurus (in SKOS)



Michael thesaurus (in SKOS)



skos:broader

skos:related

skos:match

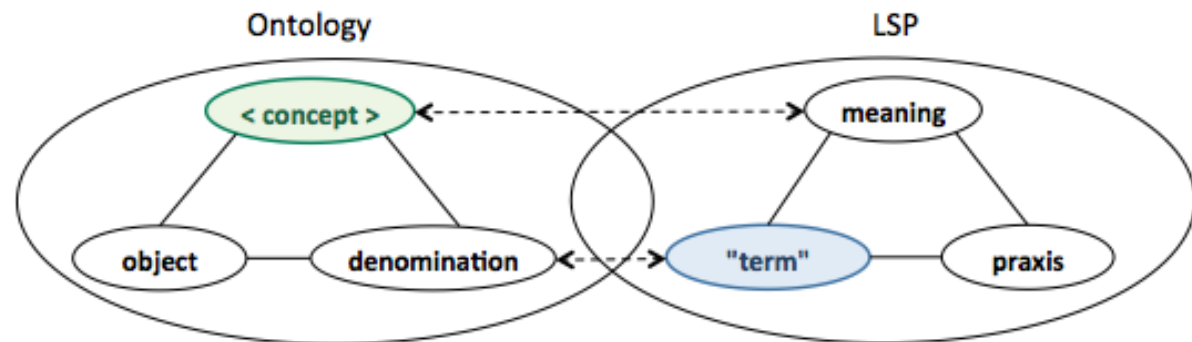
■ Open Issues

➔ Mapping WordNet and Arabic WordNet

➔ SKOS dependent

➔ Too simple principles:

- a thesaurus is neither a terminology nor an ontology
- a terminology is neither an ontology nor a thesaurus
- an ontology is neither a terminology nor a thesaurus



A double semantic triangle



SIERA: A Multilingual Knowledge Sharing Project

Pr. Christophe Roche

- Universidade Nova de Lisboa, CLUNL - Portugal
- Université de Savoie, Equipe Condillac – France

roche@univ-savoie.fr