Ismahane KOURTIN

ELLIADD Laboratory, Bourgogne-Franche-Comté University Besançon, France

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

Kourtin_ismahane.math@yahoo.fr

Samir MBARKI

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

mbarkisamir@hotmail.com

Max SILBERTZTEIN

ELLIADD Laboratory, Bourgogne-Franche-Comté University Besançon, France

max.silberztein@gmail.com

Abdelaaziz MOULOUDI

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

mouloudi_aziz@hotmail.com

NOOJ 2020 INTERNATIONAL CONFERENCE
Zagreb, 05 - 07 June 2020

Legal Question-Answering Ontologies-based System

Ismahane KOURTIN

ELLIADD Laboratory, Bourgogne-Franche-Comté University Besançon, France

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

Kourtin_ismahane.math@yahoo.fr

Samir MBARKI

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

mbarkisamir@hotmail.com

Max SILBERTZTEIN

ELLIADD Laboratory, Bourgogne-Franche-Comté University Besançon, France

max.silberztein@gmail.com

Abdelaaziz MOULOUDI

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

mouloudi_aziz@hotmail.com

Plan

1. Introduction
2. Ontological modeling of the legal field
3. Legal Question-Answering Ontologies-based System
   - 3.1. Principle
   - 3.2. Architecture
4. Study of questions
   - 4.1. Types of questions
   - 4.2. Question Categories
   - 4.3. Linguistic analysis of the questions
5. Implementation of the system and results
6. Conclusion & perspectives

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

➢ **The mass of information in the legal field is constantly increasing.**

The need to transform legal field informations to an intelligent guide capable of providing complete and immediate responses to queries in natural language.

Question-answering system

Objective :

Set up a question-answering system based on ontologies in the legal field, capable of interrogating a corpus of laws and decrees in Arabic.

1. Introduction
2. Ontological modeling of the legal field
3. Legal Question-Answering Ontologies-based System
3.1. Principle
3.2. Architecture
4. Study of questions
4.1. Types of questions
4.2. Question Categories
4.3. Linguistic analysis of the questions
5. Implementation of the system and results
6. Conclusion & perspectives

- **Human reasoning is based on what can be called an ontology of the world, that is to say, a vision that we have on the world and categories that organize this world.**

- **This ontology is :**
  - ✓ **Reasonable.**
  - ✓ **Shared.**

- **The ontology is based on RDF model.**

Principle of RDF model : a triplet model in which resources are described by triplets (subject, predicate, object)

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

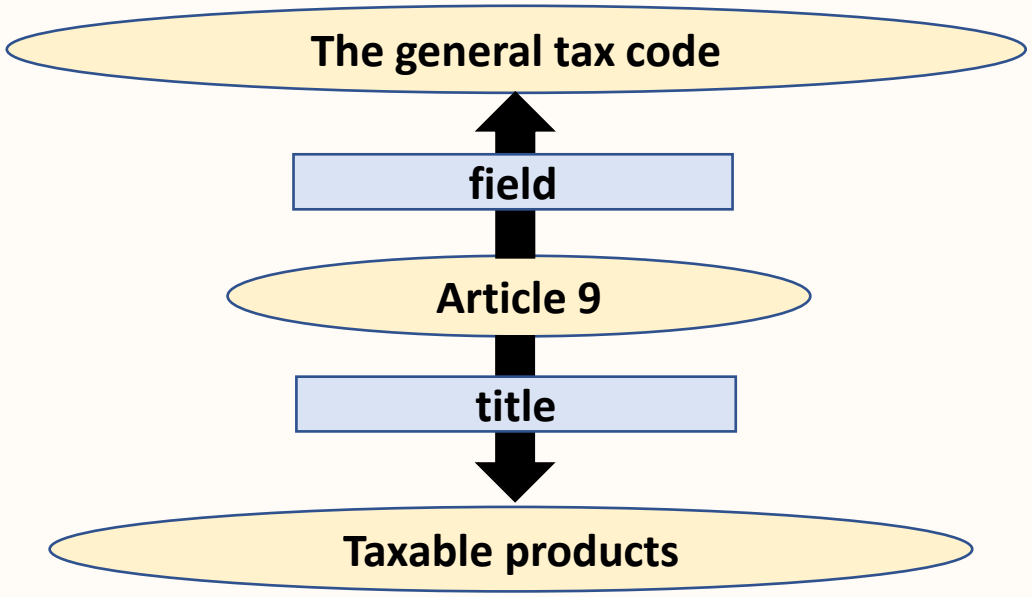
6. Conclusion & perspectives

➢ Model the statement in RDF :

« article 9 of the general tax code is entitled "taxable products" »

1) Decompose the statement in a controlled language into elementary units :
   - "article 9" has for field "the general tax code"
   - "article 9" has for title "taxable products"
2) Represent the elementary units by triplets :
   - (article 9 , field , the general tax code)
   - (article 9 , title , taxable products)

➢ This ontological knowledge is often represented in the form of trees:



1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

Concept Vs Term :

- There are several ways to call the « article 9 » :
  - « article 9 » in French and in English.
  - « 9 المادة » in Arabic.
- Let's say that there is a concept let's call it the concept #TCARTICLE9 which represents "article 9", and that we will hang on this concept:
  - the label "article 9" in French and in English;
  - the label "المادة 9" in Arabic.
- Similarly, there is a concept let's call it the concept #TC which represents "the general tax code", and that we will hang on this concept:
  - the label "le code général des impôts" in French;
  - the label "the general tax code" in English;
  - the label "المدونة العامة للضرائب" in Arabic.

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

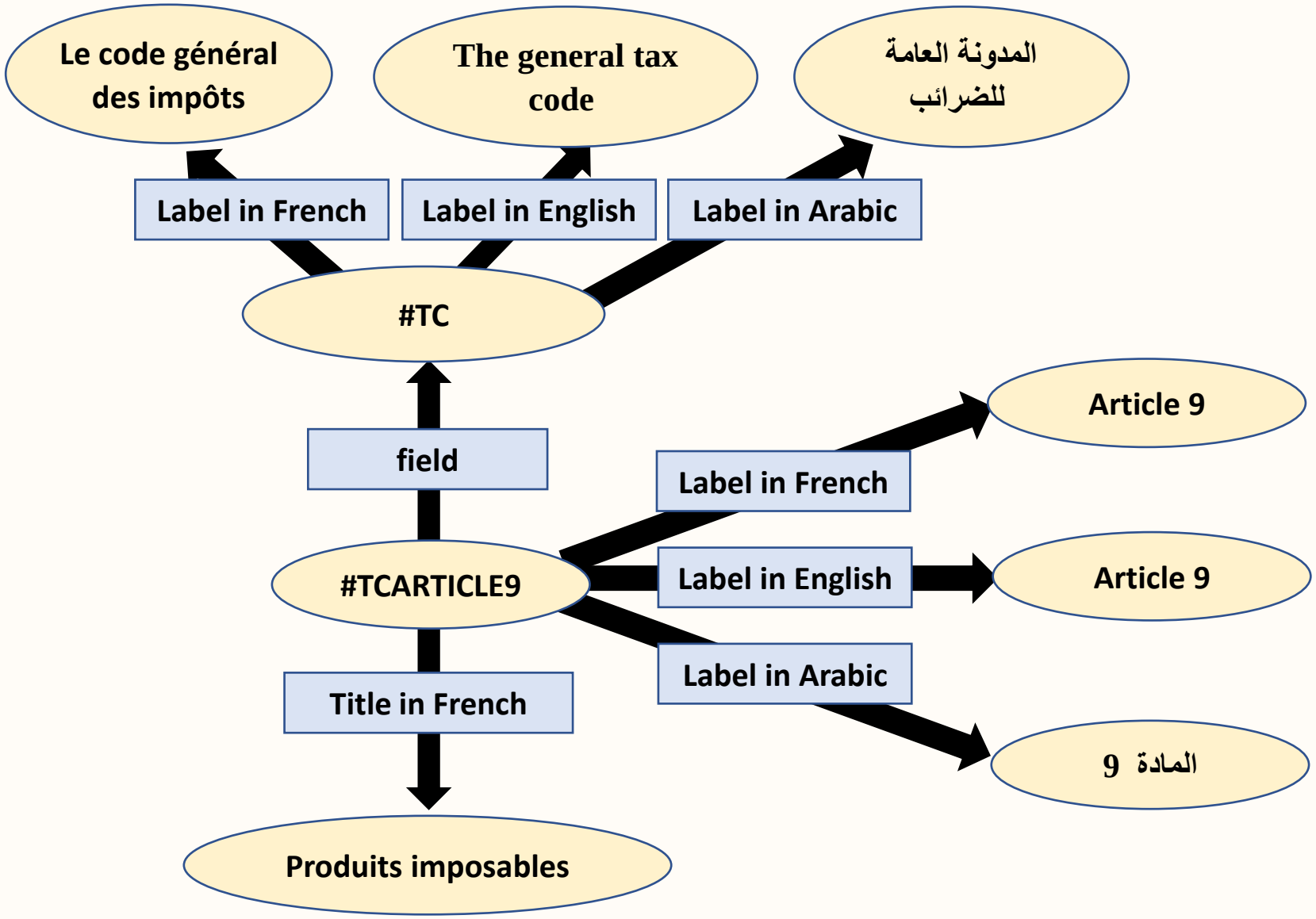
4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

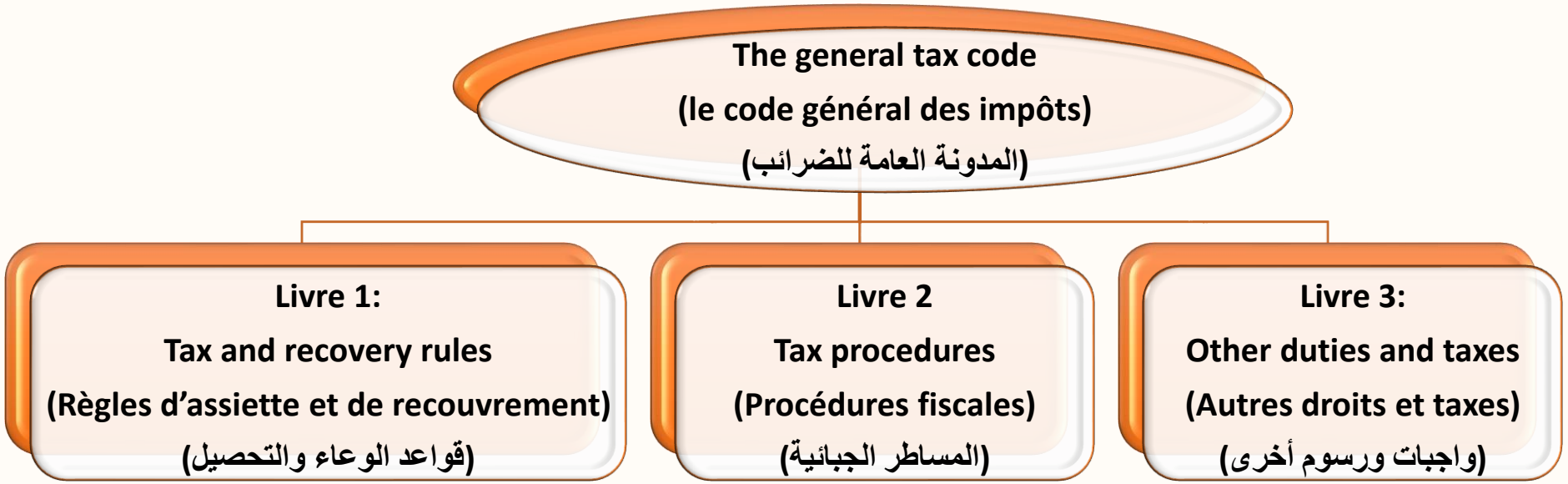
6. Conclusion & perspectives

Formal ontology:



1. Introduction
2. Ontological modeling of the legal field
3. Legal Question-Answering Ontologies-based System
3.1. Principle
3.2. Architecture
4. Study of questions
4.1. Types of questions
4.2. Question Categories
4.3. Linguistic analysis of the questions
5. Implementation of the system and results
6. Conclusion & perspectives

- It is this knowledge in logic of the legal field that we will machine so that the machine is able to simulate the inferences that we naturally make.
- We opted for a modular construction of the legal field by code.
- We decided to start with the general tax code.
- The general tax code divides into three main branches:



- Each of the branches is also broken down into more specialized branches to arrive at the end with a more precise subject with a set of articles which deals with it.

1. Introduction
2. Ontological modeling of the legal field
3. Legal Question-Answering Ontologies-based System
3.1. Principle
3.2. Architecture
4. Study of questions
4.1. Types of questions
4.2. Question Categories
4.3. Linguistic analysis of the questions
5. Implementation of the system and results
6. Conclusion & perspectives

- **The first part of the built ontology includes :**
  - **The architecture of the general tax code.**
  - **The decomposition of the general tax code into more specific sub-fields.**
  - **The articles which deal with each sub-field.**

- **To represent this machine knowledge :**
  - **RDF language syntax Turtle.**
  - **RDFS, OWL vocabularies.**
  - **The two ontologies:**
    - ✓ **SKOS ontology to represent thesauri**
    - ✓ **Dublin Core ontology to describe documents.**

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

- **Describe the general tax code architecture :**

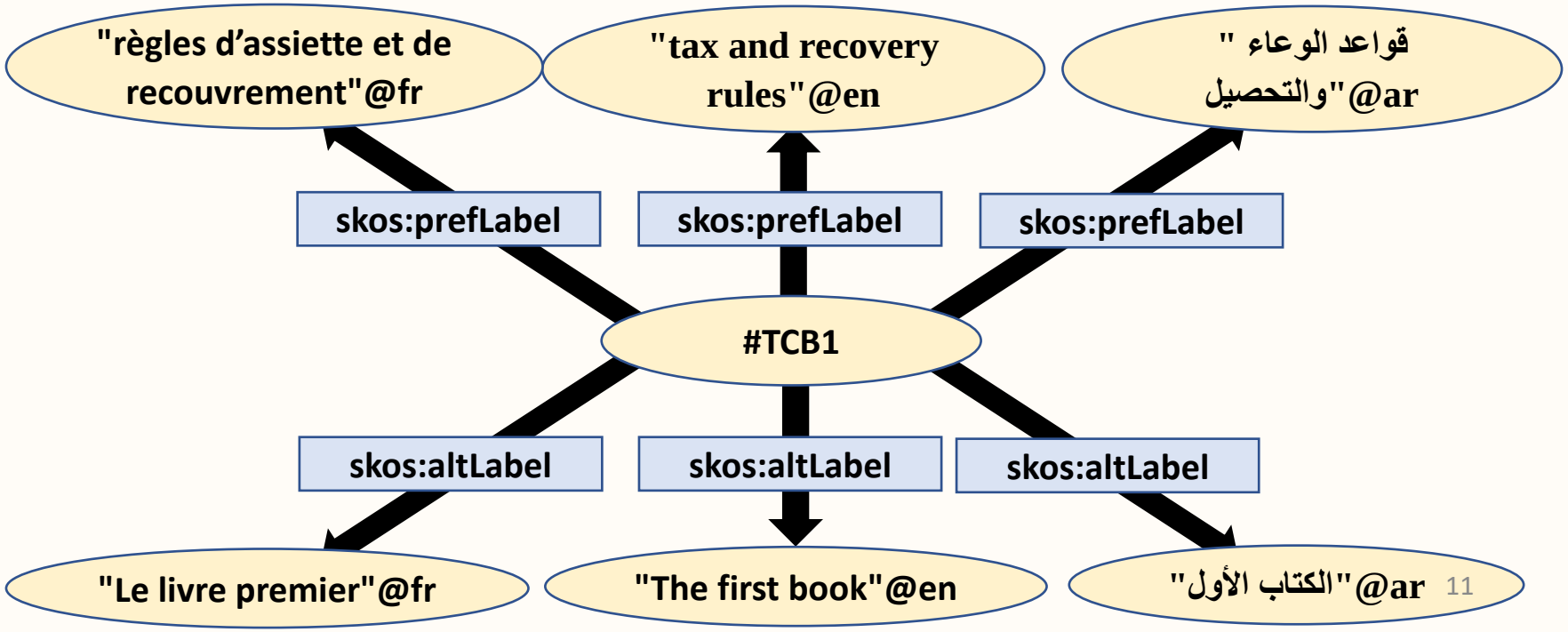
  → SKOS ontology

- **The SKOS vocabulary is defined in OWL and allows to define in RDF, simple knowledge organization;**

- **The classes and properties of the SKOS schema make it possible:**
  - **to declare concepts;**
  - **to associate them with different types of labels;**
  - **to document them in natural language;**
  - **to declare different types of relationships between these concepts;**
  - **to declare schemas of concepts which group together a set of concepts.**

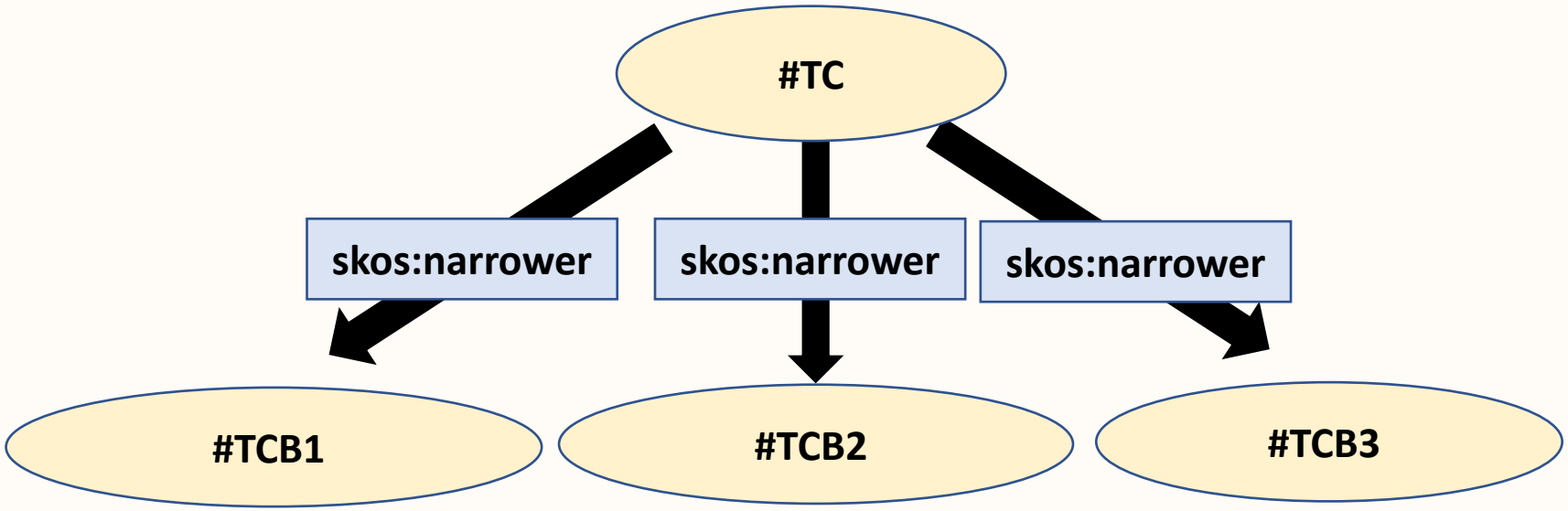
1. Introduction
2. Ontological modeling of the legal field
3. Legal Question-Answering Ontologies-based System
3.1. Principle
3.2. Architecture
4. Study of questions
4.1. Types of questions
4.2. Question Categories
4.3. Linguistic analysis of the questions
5. Implementation of the system and results
6. Conclusion & perspectives

- Each area, sub-area, or article, is defined as a concept with scos:Concept.
- Three properties in skos allow to declare concept labels more finely than with the rdfs:label property:
  - skos:prefLabel : allows to indicate the preferred label and used to designate the concept title
  - skos:altLabel : allows to indicate alternative labels and used to designate the concept name
  - skos:hiddenLabel : allows to indicate hidden labels.



1. Introduction
2. Ontological modeling of the legal field
3. Legal Question-Answering Ontologies-based System
3.1. Principle
3.2. Architecture
4. Study of questions
4.1. Types of questions
4.2. Question Categories
4.3. Linguistic analysis of the questions
5. Implementation of the system and results
6. Conclusion & perspectives

- The skos:definition property is used to associate a natural language definition with a concept.
- For the semantic relationships between the concepts, I used the properties:
  - skos:narrower : the value concept of the property skos:narrower is more precise than the subject concept;
  - skos:broader : the reverse property of skos:narrower.



1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

➢An extract from the ontology in turtle:

```
<#TCB1> a skos:Concept;
        skos:narrower <#TCB1P1>;
        skos:narrower <#TCB1P2>;
        skos:narrower <#TCB1P3>;
        skos:prefLabel "قواعد الوعاء والتحصيل"@ar;
        skos:prefLabel "règles d’assiette et de recouvrement"@fr;
        skos:prefLabel "tax and recovery rules"@en;
        skos:altLabel "الكتاب الأول"@ar;
        skos:altLabel "Le livre premier"@fr;
        skos:altLabel "The first book"@en;
        skos:definition " يضم قواعد الوعاء والتحصيل والجزاءات المتعلقة بالضريبة على
الشركات والضريبة على الدخل والضريبة على القيمة المضافة وواجبات التسجيل"@ar;
        skos:definition "regroupant les règles d'assiette, de recouvrement et
des sanctions en matière d’I.S, d’I.R, de T.V.A et de D.E."@fr;
        skos:definition "regrouping the rules of assessment, recovery and
sanctions in matters of SSI, IR, T.V.A and D.E."@en;
        dc:date "2020".
```

1. Introduction
2. Ontological modeling of the legal field
3. Legal Question-Answering Ontologies-based System
3.1. Principle
3.2. Architecture
4. Study of questions
4.1. Types of questions
4.2. Question Categories
4.3. Linguistic analysis of the questions
5. Implementation of the system and results
6. Conclusion & perspectives

- The ontology is based on the RDF model which makes it possible to represent and exchange knowledge in triplets form.

(subject, predicate, object)

- SPARQL is a query language that allows to query RDF graphs.

if we know at least one element of a triplet, then we can query the RDF graph via SPARQL queries to retrieve the rest of the data from the triplet

- In the example of the triplet: (article 9 , title , taxable products)
  - If we know the name of the subject that is “article 9” and the property of this object that interests us that is the “title”,
  - Then we can recover the value of the title that is "taxable products" Via a SPARQL query which will query the ontology.

- This answers a question like: What is the title of article 9 of the general tax code?

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

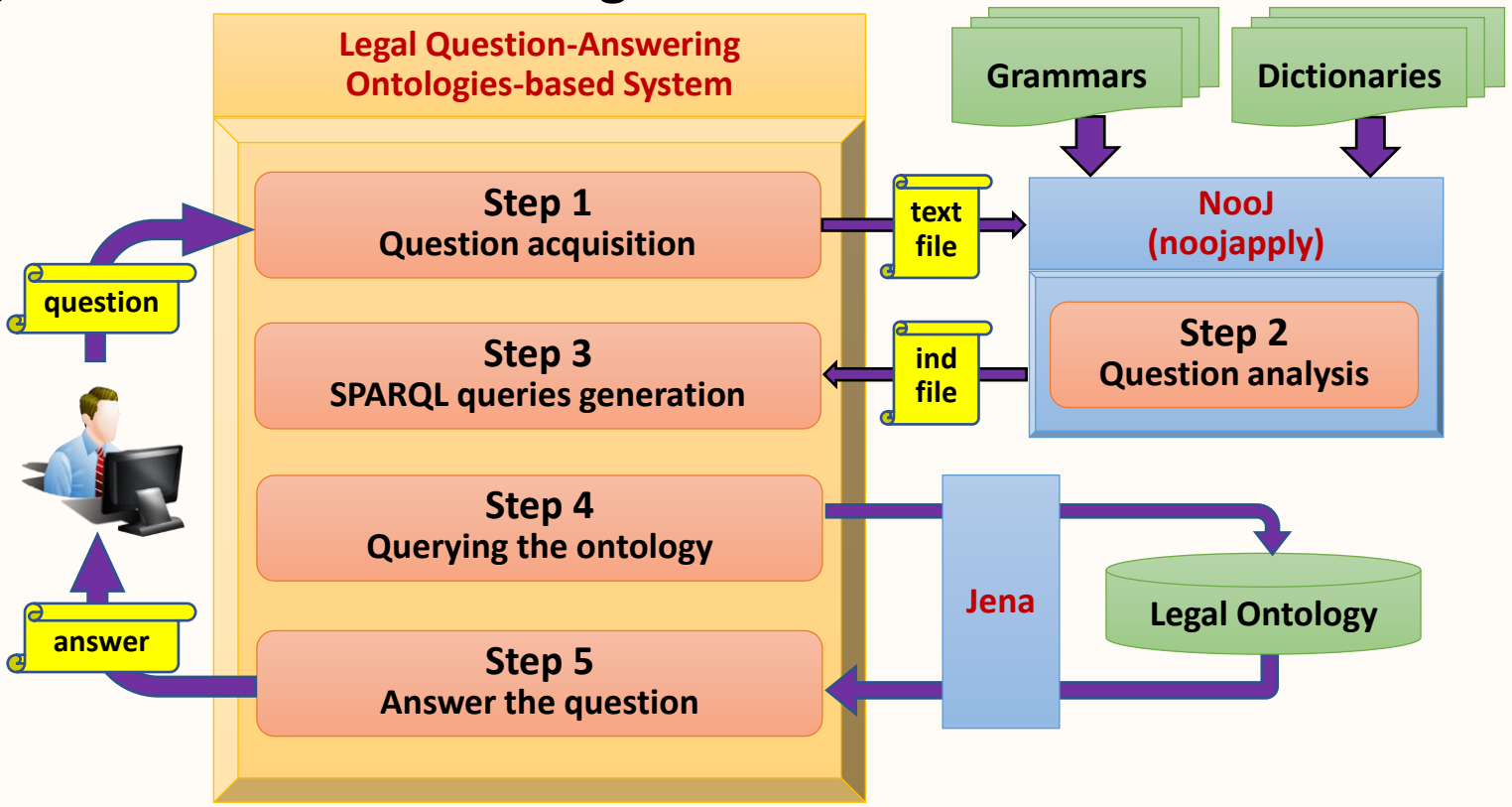
5. Implementation of the system and results

6. Conclusion & perspectives

- **Challenge**: analyze the question of the user and extract the data from the triplet(s) and the unknowns targeted by the question.
- These data and unknowns of the triplet(s) will be transformed into a SPARQL query which will be responsible for retrieving the values of the unknowns of the triplet(s) from the ontology.

→ Get the answer to the user's question

- The proposed architecture of the question-answering system in the legal field based on ontologies :



1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

- **No matter the field, no matter the language, we can group the questions into two types according to the expected answer :**
  - **Boolean questions whose answer is either yes or no: AskQuestion ;**
  - **Questions to be answered : SelectQuestion.**

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

➢Six categories of legal questions are identified according to the purpose of the question :

Category	Example
YesNoHavingProp : questions allowing to know if the value of a property corresponds or not	- Est-ce que les véhicules agricoles sont taxables ? -هل السيارات الزراعية خاضعة للضريبة؟ - Are agricultural vehicles taxable?
ValueOfProp : questions allowing to find out the value of a property	- Quel est l'intitulé de l'article 15 du code général des impôts ? -من المدونة العامة للضرائب؟ 15ما هو عنوان المادة - What is the title of article 15 of the general tax code?
EntitiesHavingProp : questions for obtaining the list of entities with a certain property value	- Quels sont les véhicules taxables ? -ما هي المركبات الخاضعة للضريبة؟ - What vehicles are taxable?
ArticlesPrescribingClass : questions allowing to find the articles relating to a given subject	- Quelles sont les articles relatifs à l'impôt sur les sociétés ? -ما هي المواد المتعلقة بالضريبة على الشركات؟ - What are the articles relating to corporate tax?
ConditionOnProp : questions used to identify conditions or contexts on a certain property value	- Dans quelles conditions les coopératives sont exonéré des impôts ? -ما هي شروط إعفاء التعاونيات من الضرائب؟ - Under what conditions are cooperatives exempt from taxes?
TargetedEntities : questions allowing to identify the entities targeted by a given legal text	- Quels types de véhicules sont concernés par l'article 259 du code général des impôts ? -من المدونة العامة للضرائب؟ 259ما هي أنواع المركبات التي تغطيها المادة - What types of vehicles are covered by article 259 of the general tax code?

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

- The patterns of the proposed questions are based on a formal language which will be represented by a regular grammar in the NooJ automatic language processing platform.
- One question is the sequence of:

QuestionType	QuestionBody	QuestionMark
a sequence of words coding the type of the question	the body of the question	the end question mark

- For the QuestionType, there are two types depending on the expected answer:
  - Boolean questions whose answer is either yes or no (AskQuestion) ;
  - Questions to be answered (SelectQuestion).
- The QuestionBody is formed by one of the six categories of questions mentioned above : YesNoHavingProp, ValueOfProp, EntitiesHavingProp, ArticlesPrescribingClass, ConditionOnProp and TargetedEntities.
- QuestionMark is the symbol « ? » in French and in English, and the symbol « ؟ » in Arabic.

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

Extracting patterns from ValueOfProp category questions :

The ValueOfProp category questions, which allow to know the value of a property, are represented by the sequence: Property, Of, Subject, and OfRestriction and Restriction which are factual.

	Sample question
French	QuestionType Property Of Subject OfRestriction Restriction QuestionMark Quel est l' intitulé de l' article 15 du code général des impôts ?
Arab	QuestionMark Restriction OfRestriction Subject Of Property QuestionType ؟ المدونة العامة للضرائب من 15 مادة ال عنوان ما هو
English	QuestionType Property Of Subject OfRestriction Restriction QuestionMark What is the title of article 15 of the general tax code ?

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

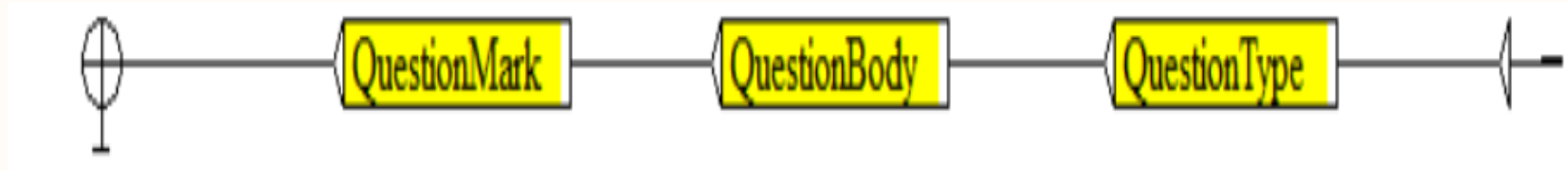
5. Implementation of the system and results

6. Conclusion & perspectives

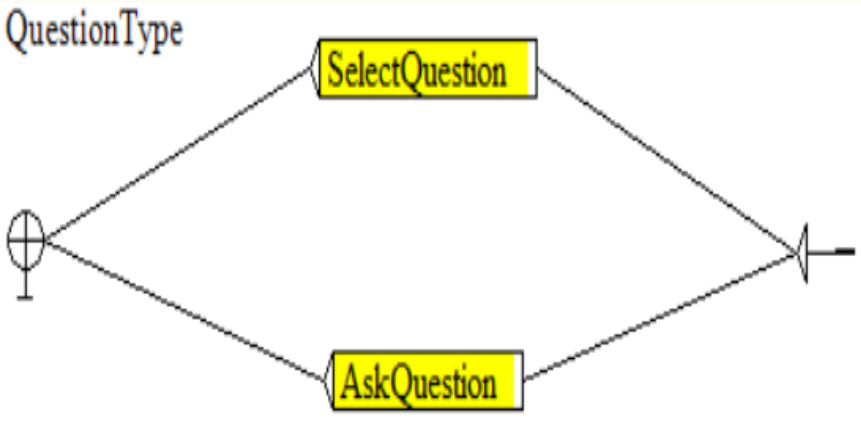
The syntax grammars of the ValueOfProp category questions:

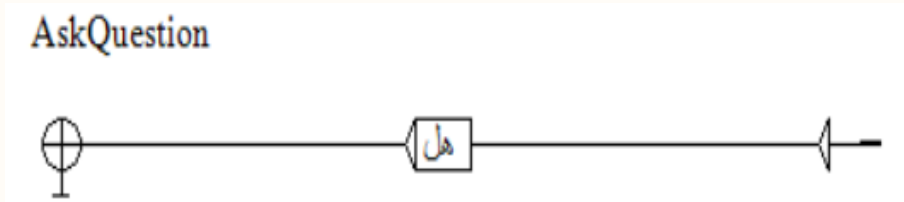
- The grammars recognize and extract the components of the RDF triplet (s) from the user's question.

A. Established syntactic grammars in Arabic : 1/4



- This transducer describes the main components of the question.





- This transducer allows to recognize the type of the question.

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

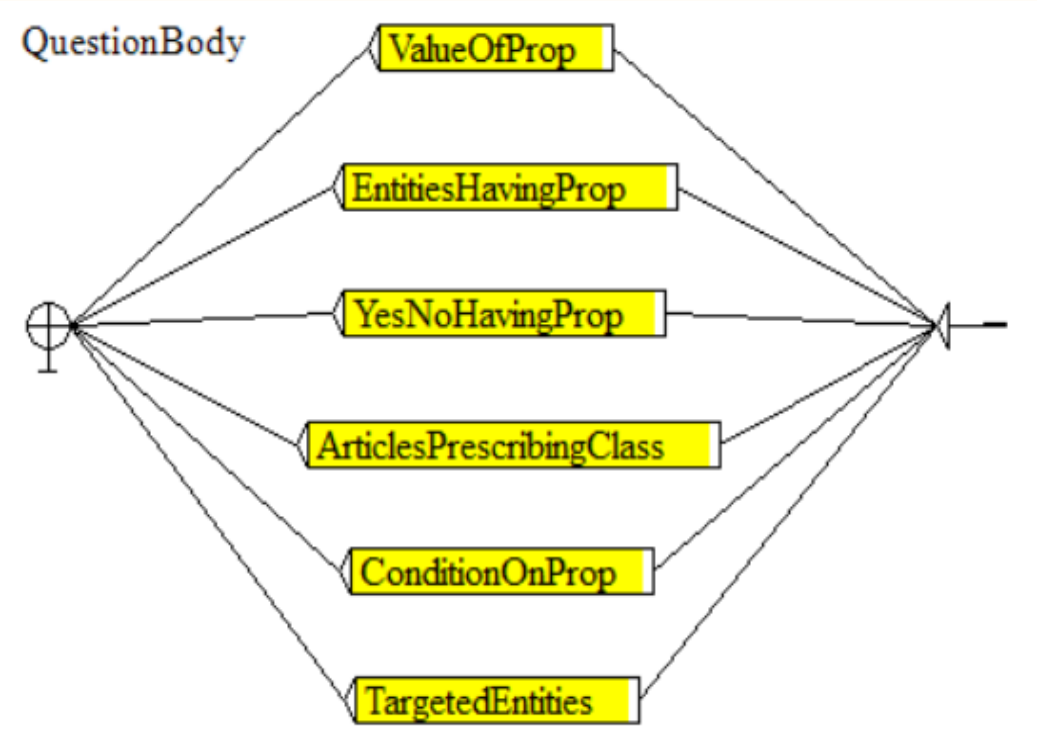
4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

The syntax grammars of the ValueOfProp category questions :

A. Established syntactic grammars in Arabic : 2/4



- **This transducer describes the different paths allowing the definition of the category of the question. Each path represents a set of rules allowing to recognize a category of questions.**

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

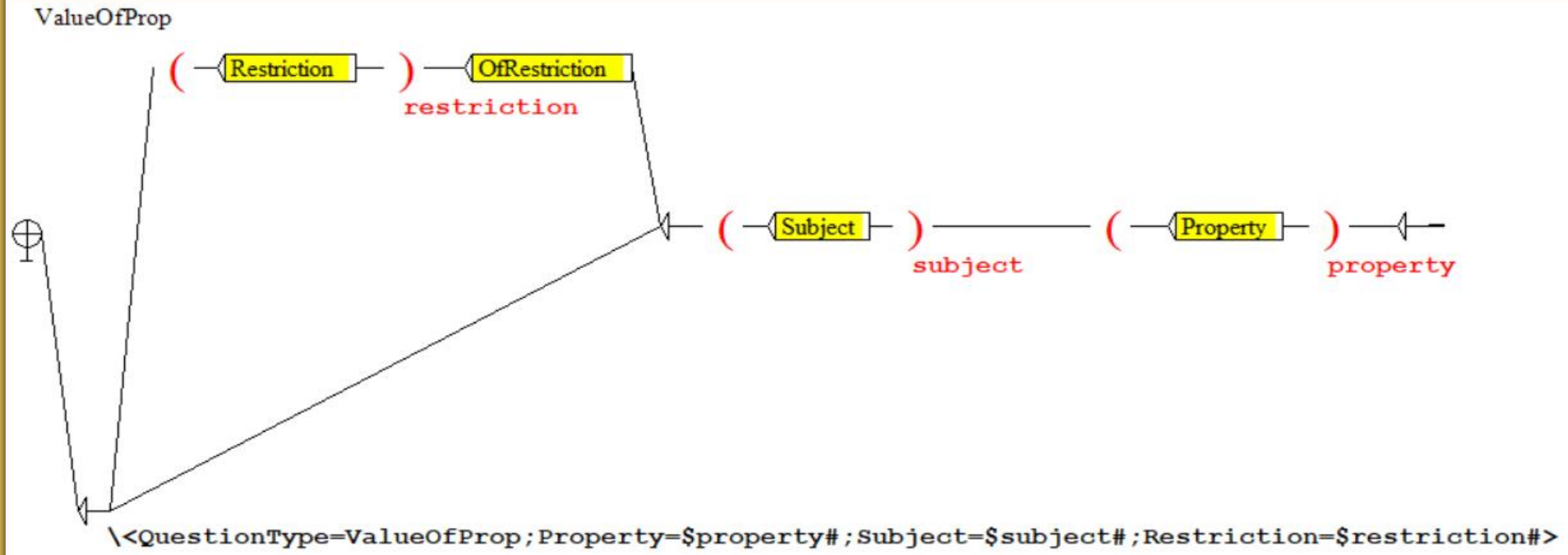
4.3. Linguistic analysis of the questions

5. Implementation of the system and results

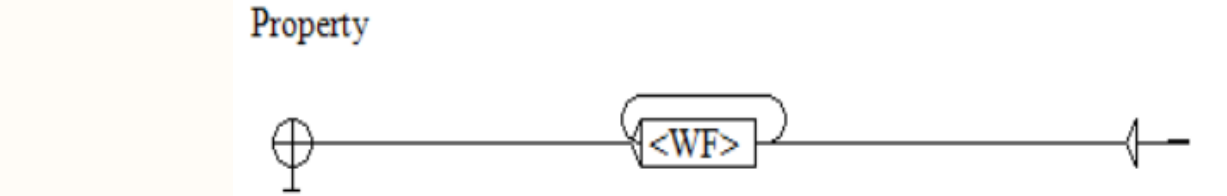
6. Conclusion & perspectives

The syntax grammars of the ValueOfProp category questions :

A. Established syntactic grammars in Arabic : 3/4



- This transducer makes it possible to recognize the RDF components of the ValueOfProp category questions.



- This transducer makes it possible to recognize the property of the RDF triplet of questions in the ValueOfProp category.

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

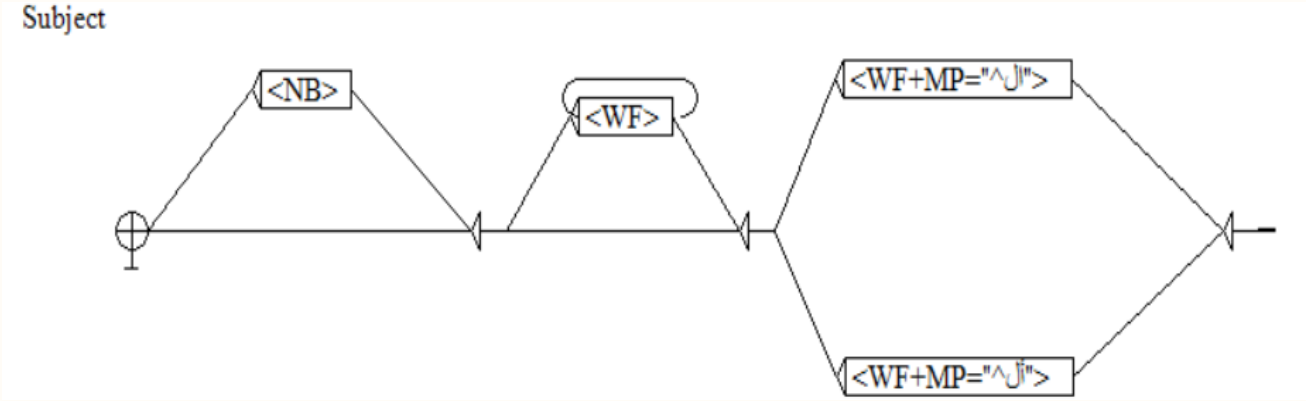
4.3. Linguistic analysis of the questions

5. Implementation of the system and results

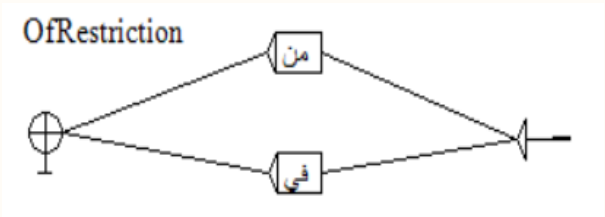
6. Conclusion & perspectives

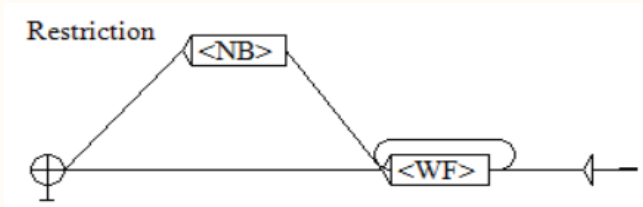
The syntax grammars of the ValueOfProp category questions :

A. Established syntactic grammars in Arabic : 4/4

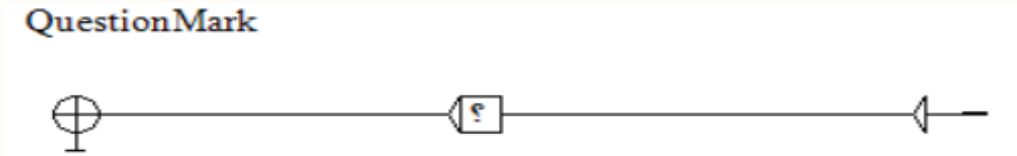


➢ This transducer allows to recognize the subject of the RDF triplet of ValueOfProp category questions.





➢ This transducer allows to recognize the restriction on the questions of ValueOfProp category.



➢ This transducer allows to recognize the end of the question.

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

➢The system is capable of answering questions in Arabic, French and English such as:

- Quel est l'intitulé de l'article 15 du code général des impôts ?
- ما هو عنوان المادة 15 من المدونة العامة للضرائب؟
- What is the title of article 15 of the general tax code?

1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

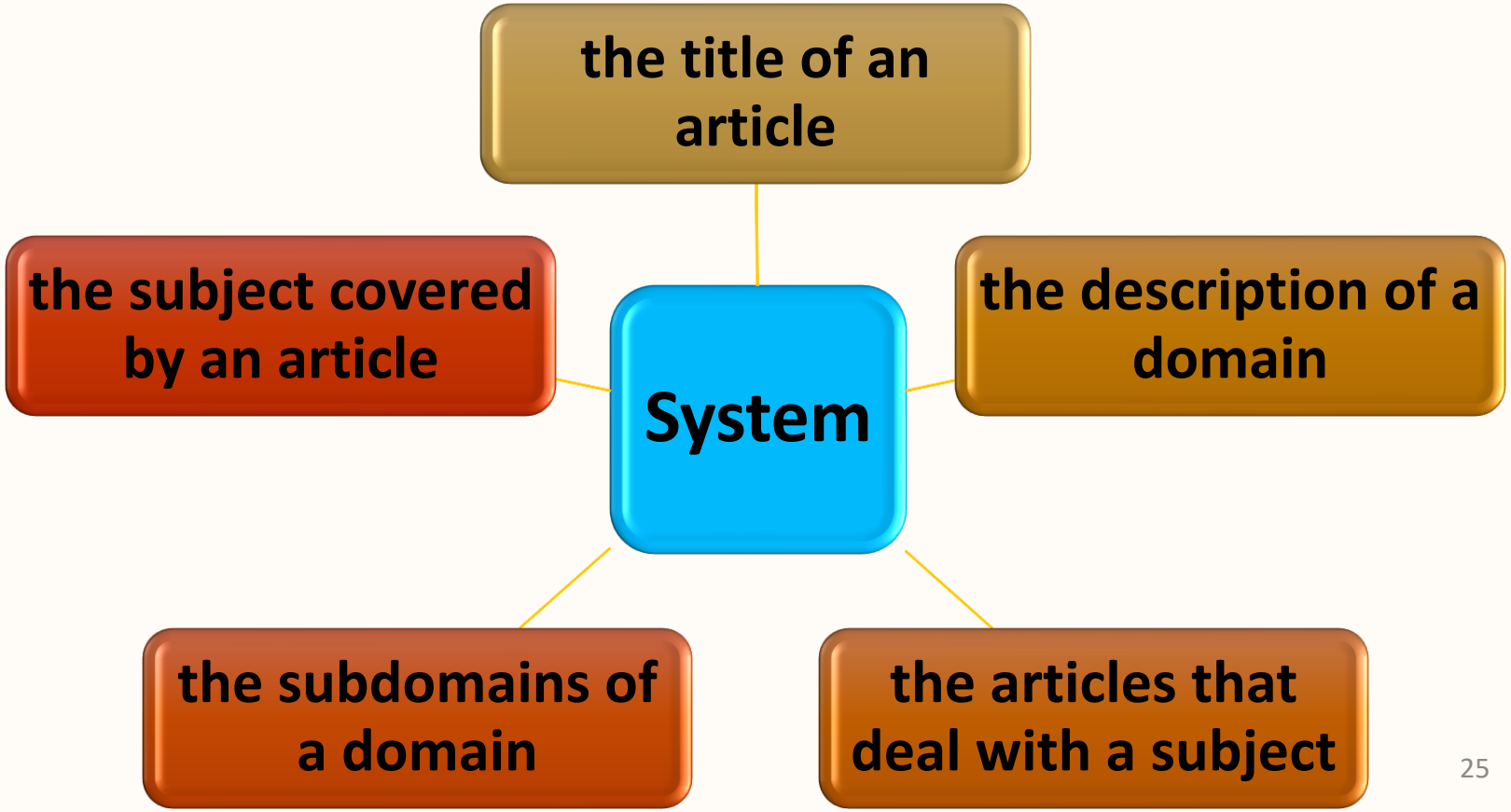
4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

Conclusion :

- A legal ontology of the architecture of the general tax code is constructed;
- A question-answering system in the legal field based on this ontology is developed.



1. Introduction

2. Ontological modeling of the legal field

3. Legal Question-Answering Ontologies-based System

3.1. Principle

3.2. Architecture

4. Study of questions

4.1. Types of questions

4.2. Question Categories

4.3. Linguistic analysis of the questions

5. Implementation of the system and results

6. Conclusion & perspectives

Perspectives:

- Enrich the ontology with the modeling of the content of articles ;
- Build a corpus of questions ;
- Enrich the patterns of questions.

Ismahane KOURTIN

ELLIADD Laboratory, Bourgogne-Franche-Comté University Besançon, France

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

Kourtin_ismahane.math@yahoo.fr

Samir MBARKI

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

mbarkisamir@hotmail.com

Max SILBERTZTEIN

ELLIADD Laboratory, Bourgogne-Franche-Comté University Besançon, France

max.silberztein@gmail.com

Abdelaaziz MOULOUDI

MISC Laboratory, Faculty of Science, Ibn Tofail University Kenitra, Morocco

mouloudi_aziz@hotmail.com

Thank you for your attention !