Les défis de l'accès à l'information juridique
Entre annotation et recherche d'information sémantiques

Adeline Nazarenko
adeline.nazarenko@lipn.univ-paris13.fr

LIPN, Université Paris13 – Sorbonne Paris Cité & CNRS

12 décembre 2016 — Journée IA & RI

Accès à l'information juridique

- Des besoins croissants
  - Les citoyens
    "Quels arrêtés municipaux existent concernant les chemins ruraux et les véhicules à moteurs?"
  - Les personnels municipaux
    "Quels arrêtés municipaux concernant... ont fait l'objet d'un recours et ont été annulés par décision de tribunal ?"
  - Le juriste : "Quelle jurisprudence s'applique pour l'indemnisation de... ?"
  - Le législateur : "Comment la directive sur le droit d'auteur est-elle transposée ?"
- Des acteurs de plus en plus présents

  éditeurs juridiques, pouvoir publics, professionnels du droit
- Le mouvement général d'ouverture des données et de web sémantique

Les défis

- La recherche d'information sémantique
  - Quelles documents ?
  - Quelles requêtes ?
- L'annotation sémantique
  - Quel langage d'annotation ?
  - Comment annoter ?

Des réflexions plutôt que des résultats ou des conclusions...

Retour d'expérience sur quelques projets

- ONTORULE project (FP7) François Lévy, Abdoulaye Guisé (PhD)

ONTORULE
Ontologiesmeet Business Rules

- projet FUI Légilocal Sylvie Salotti, Nada Mimouni (PhD)

LégiLocal

- Labex "Empirical Foundations of Linguistics" (ANR)

Labex EFL

- Mining Scottish Legislation (Aberdeen University) François Lévy, Adam Wyner

Plan

Introduction

Accès à l'information juridique

Accéder aux sources
- Annotation sémantique au regard d'une ontologie
- Prise en compte de l'intertextualité

Accéder aux règles
- Reformulation
- Annotation à gros grain

Conclusion et perspectives

Outline

Introduction

Accès à l'information juridique

Accéder aux sources

Accéder aux règles

Conclusion et perspectives

La RI juridique

- Statut particulier du document
  - Le droit est d'abord un "discours"
  - Une décision juridique est un texte
- Enjeu de la sécurité juridique
  - Impératif d'exhaustivité
  - Nécessité de pouvoir filtrer

Exemple de Légifrance

Décret n°60-452 du 12 mai 1960 - Article 1

Masquer le panneau de navigation *<< Article précédent - Article suivant >> - Imprimer*

Navigation

Retour au texte en version initiale

Article 1

Versions de l'article:

- **Version abrogée au 21 décembre 1985**
- Version en vigueur du 13 mai 1960 au 11 janvier 1985

Version consolidée à la date du ...

Jour	Mois	Année
1	Janvier	2015

Consulter

Chemin :

Décret n°60-452 du 12 mai 1960 RELATIF A L'ORGANISATION ET AU FONCTIONNEMENT DE LA SECURITE SOCIALE

- DES REGIMES DE SECURITE SOCIALE

Article 1

Modifié par Décret n°85-38 du 10 janvier 1985 - art. 6 JORF 11 janvier 1985
Modifié par Décret 85-292 1985-03-01 art. 1 JORF 3 mars 1985 en vigueur le 1er janvier 1985
Abrogé par Décret 85-1353 1985-12-17 art. 6 JORF 21 décembre 1985

L'organisation de la sécurité sociale comprend [*composition*] :

En ce qui concerne le régime général :

- des caisses primaires et régionales de sécurité sociale ;

- des caisses d'allocations familiales ;

Systèmes existants

- Exemples
  - Légifrance
  - Normattiva
  - UK Legislation
  - ...
- Fonctionnalités
  - Recherche en plein texte : Quels sont les décrets qui traitent du bruit ?
  - Recherche par attribut : Quelles lois sont abrogées ?
  - Navigation d'un document à l'autre à travers les liens hypertextes
  - Consultation de la version consolidée d'un texte
- Mais
  - Pas de normalisation lexicale
  - Pas d'interrogation relationnelle

    Quels sont les textes qui modifient un décret sur les nuisances sonores ?
  - pas de support à la décision juridique

    Quelles règles s'appliquent sur quels cas ?

L'effort a principalement porté sur

- Les recommandations de rédaction
- L'ingénierie documentaire : des standards pour structurer l'information juridique
  - MetaLex, CEN MetaLex [Winkels *et al.*, 2003]
  - Akoma Ntoso [Sartor *et al.*, 2011a]
- La modélisation ontologique
  - Micro-ontologies [Després and Szulman, 2005]
  - LKIF-Core Ontology [Hoekstra *et al.*, 2007]
  - Patrons de conceptions [Gangemi, 2009]
- La modélisation du raisonnement ontologique AI & Law
  - Modélisation des modalités
  - Langages de règles : LegalRuleML [Athan *et al.*, 2015]

Outline

Introduction

Accès à l'information juridique

Accéder aux sources
- Annotation sémantique au regard d'une ontologie
- Prise en compte de l'intertextualité

Accéder aux règles

Conclusion et perspectives

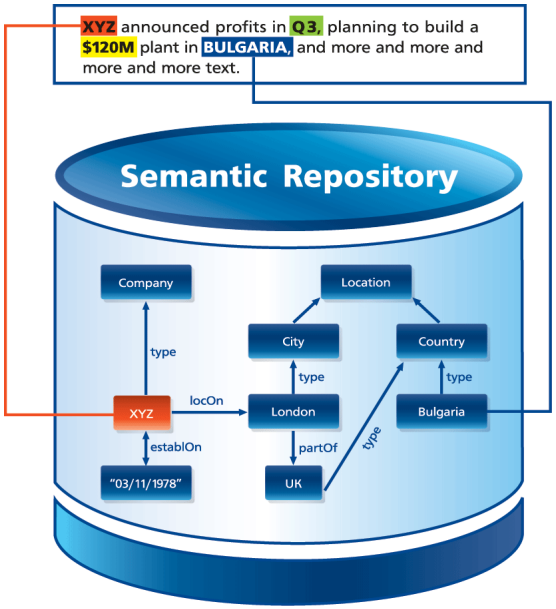
Qu'est-ce que l'annotation sémantique ?

- Annotation de textes : des métadonnées sont associées à des fragments de textes

  note, commentaire, explication, marques de mise en forme

- Annotation sémantique : les métadonnées relève d'une ressource sémantique

  vocabulaire contrôlé, terminologie, nomenclature, thesaurus, ontologie



[Popov *et al.*, 2004]

Annotation sémantique : défis

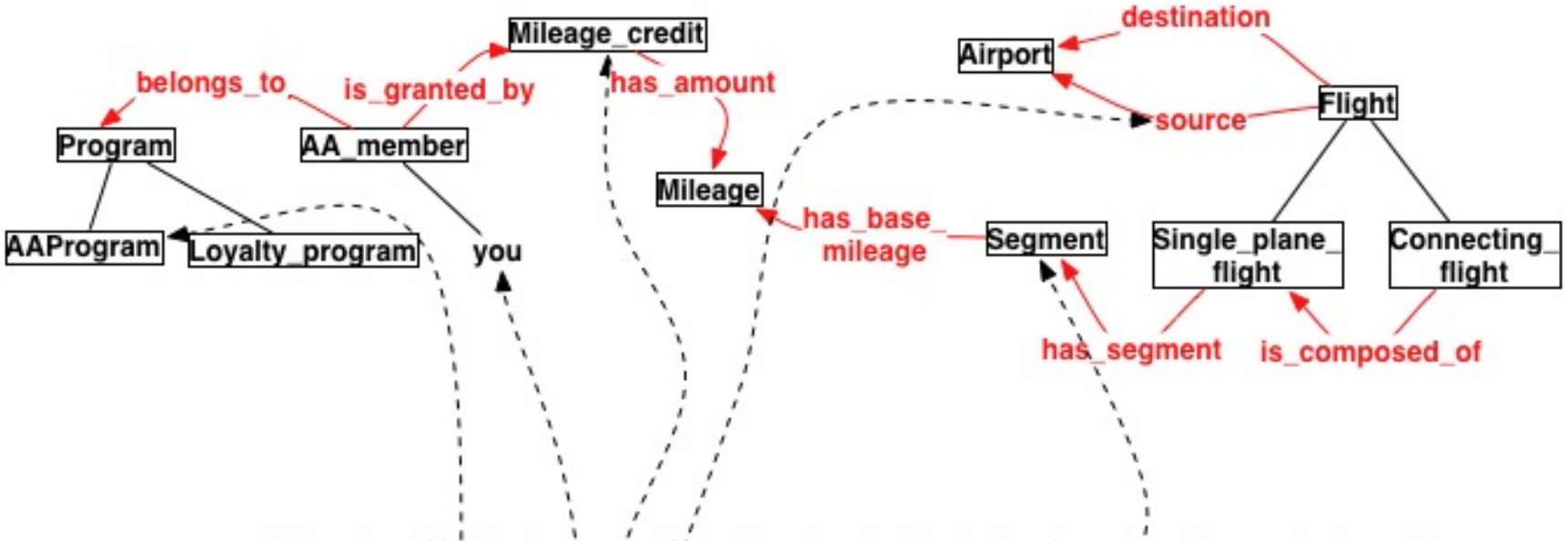
L'accent a été mis au départ sur l'annotation des instances et le peuplement d'ontologies [Vargas-Vera *et al.*, 2002] [Popov *et al.*, 2004] [Amardeilh *et al.*, 2005] [Magnini *et al.*, 2006]

mais Il faut souvent

- Coupler acquisition et annotation
- Proposer une annotation sémantique plus riche [Ma *et al.*, 2010]
  - Avec un grain variable
  - Exploitant toute la sémantique des ontologies [Lévy *et al.*, 2010b]

    individuals, concepts, roles/relations, rules
- Annoter automatiquemet

Annotation au regard d'une ontologie



AAdvantage **flight mileage credit** is determined on the basis of nonstop distances between the **airports** where your **flight** originates and terminates.
On **connecting flights**, **you**'ll receive **mileage credit** for each **segment** of your **trip** ;
on **single-plane flights**, **you**'ll receive the nonstop origin-destination **mileage**.

Mileage credit cannot be earned for the same **flight** in more than one of the following **programs**:
the **AAdvantage program** or any other **loyalty program** in which **American Airlines** participates.

Navigation dans le texte annoté

SemEx

Rule Editor | Navigator

Hierarchy

- Space
- Source
- Agent
- Function
- Attribute
  - Safety
  - Humidity
  - Temperature
- Number
- Device
- Dimension
- Conditioning
- Method
  - VirtualMethod
  - PhysicalMethod
    - ResistanceTest
    - PhysicalSeatBeltTes
      - TrolleyTest
      - TensileTest
      - BreakingStrengtl
      - ExposureTest
      - MicroSlipTest
      - CorrosionTest
    - FixtureTest

Concepts | Properties

Corpus

/Users/fl/flhome/1recherche/LN/corpus/ontorules/2011-02-24et25Review meeting/demoQuery4/resources/semantics/AudiMiniRFDa.html

Ontological entity

7.3. Micro-slip test (see Annex 11, figure 3 to this Regulation).

7.3.1. [R1]The samples to be submitted to the micro-slip test shall be kept for a minimum of 24 hours in an atmosphere having a temperature of 20 + 5 C and a relative humidity of 65 + 5 per cent. [R13]The test shall be carried out at a temperature between 15 and 30 C.

7.3.2. It shall be ensured that the free section of the adjusting device points either up or down on the test bench, as in the vehicle.

7.3.3. [R18]A 5 daN load shall be attached to the lower end of the section of strap.[R18] The end shall be subjected to a back and forth motion, the total amplitude being 300 + 20 mm (see figure).

7.3.4. If there is a free end serving as reserve strap, it must in no way be fastened or clipped to the section under load.

7.3.5. It shall be ensured that on the test bench the strap, in the slack position, descends in a concave curve from the adjusting device, as in the vehicle. The 5 daN load applied on the test bench shall be guided vertically in such a way as to prevent the load swaying and the belt twisting. The attachment shall be fixed to the 5 daN

[Lévy *et al.*, 2010a]

L'annotation sémantique juridique...

- Un incontournable
  - Structure du texte
  - Vocabulaire et tournures juridiques
  - Concepts-clés (du droit mais pas seulement)
- Un effort collectif
  - Ressources génériques
  - Ontologies de domaine
  - Publication des sources et de leurs métadonnées

... toujours en chantier

Plan

Introduction

Accès à l'information juridique

Accéder aux sources
- Annotation sémantique au regard d'une ontologie
- Prise en compte de l'intertextualité

Accéder aux règles

Conclusion et perspectives

L'intertextualité

- Définition : [Genette, 1982]
  « relation de coprésence entre deux ou plusieurs textes, c.-à-d. éidétiquement et le plus souvent, par la présence effective d'un texte dans un autre »
  - Plagiat — Illicite → détection
  - Citation — explicite → analyse, exploitation
  - Allusion, pastiche — stylistique, génétique des textes

  [Ganascia *et al.*, 2014]
- Rôle : un texte ne s'interprète pas isolément mais en fonction de tous les autres textes auxquels il fait référence et qui lui font référence

→ Prendre en compte l'intertextualité dans l'analyse de contenu

Caractéristiques de l'intertextualité juridique

- Explicite, codifiée et normative
- Multiforme : liens de citation, jurisprudence, transposition, etc.
- Omniprésente dans le raisonnement juridique

Décret n°60-452 du 12 mai 1960 - Article 1

Masquer le panneau de navigation *<< Article précédent - Article suivant >> - Imprimer*

Navigation

Retour au texte en version initiale

Article 1

Versions de l'article:

- **Version abrogée au 21 décembre 1985**
- Version en vigueur du 13 mai 1960 au 11 janvier 1985

Version consolidée à la date du ...

Jour	Mois	Année
1	Janvier	2015

Consulter

Chemin :

Décret n°60-452 du 12 mai 1960 RELATIF A L'ORGANISATION ET AU FONCTIONNEMENT DE LA SECURITE SOCIALE

- DES REGIMES DE SECURITE SOCIALE

Article 1

Modifié par Décret n°85-38 du 10 janvier 1985 - art. 6 JORF 11 janvier 1985
Modifié par Décret 85-292 1985-03-01 art. 1 JORF 3 mars 1985 en vigueur le 1er janvier 1985
Abrogé par Décret 85-1353 1985-12-17 art. 6 JORF 21 décembre 1985

L'organisation de la sécurité sociale comprend [*composition*] :

En ce qui concerne le régime général :

- des caisses primaires et régionales de sécurité sociale ;

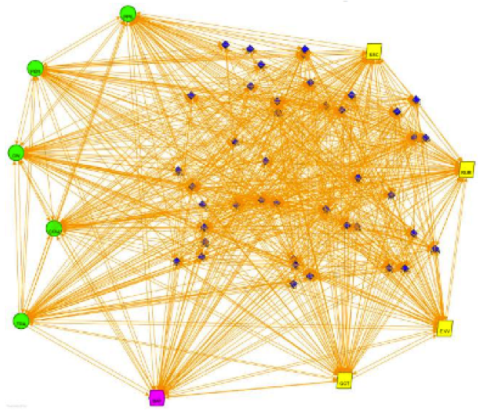
- des caisses d'allocations familiales ;

Accéder aux sources

Traitement de l'intertextualité juridique

- Détection, normalisation et formatage [Palmirani *et al.*, 2009][Sartor *et al.*, 2011b]
  - Fonctionnalités de navigation
  - Risque de désorientation [Conklin, 1987]
- Analyse des réseaux de citations (normes, codes, jurisprudence)
  - Analyse des la structure des réseaux [Bommarito and Katz, 2009] [Winkels and de Ruyter, 2011][Boulet *et al.*, 2011]
  - Calcul de notoriété [Fowler *et al.*, 2007]
  - Aide à la visualisation [Gultemen and van Engers, 2013]
- Quid de l'interrogation ?

Normes, codes, jurisprudence, etc.



Intertextualité et recherche d'information

- Recherche d'information généraliste
  - Analyse de de la structure du web [Broder *et al.*, 2000]
  - Calcul de notoriété [Brin and Page, 1998][Kleinberg, 1999]
- Recherche d'information juridique Légifrance, UK Legislation, etc.
  - Visualisation du lien et de son type

Article 1

Modifié par Décret n°85-38 du 10 janvier 1985 - art. 6 JORF 11 janvier 1985

  - Navigation (cliquer sur le lien)
  - Recherche par attributs: trouver les documents « modifiés »

→ Interrogation relationnelle

Analyse des besoins des juristes

- Juristes : Meritxell Fernandez (CERSA), Ève Paul (Victoires Éditions)
- Exemples de requêtes : types de doc. + descripteur sém. + lien intertextuel
  - *Je cherche la décision qui a fait l'objet du jugement A de la Cours de cassation*
  - *J'aimerais voir les arrêtés municipaux concernant les chemins ruraux qui ont fait l'objet d'un appel et ont été annulés par décision de jurisprudence*
  - *Quels sont les articles de code cités par les arrêtés municipaux qui concernent les chemins ruraux et qui ont été confirmés?*
  - *Je me demande si les textes visés par les arrêtés municipaux portant sur les chemins ruraux sont également cités par ceux concernant les véhicules à moteur*

→ Intégrer l'interrogation relationnelle à la recherche d'information

Formulation simple du problème

- Hypothèses de travail
  - Les documents sont annotés
  - Les liens intertextuels sont identifiés
  - Les documents et les liens sont typés
- Approche à base de graphes
  - Collection documentaire = Réseau sémantique de documents
  - Requêtes = Graphes
  - Appariement de graphes

$\rightarrow$ Intégrer l'interrogation relationnelle à la recherche d'information

Collection

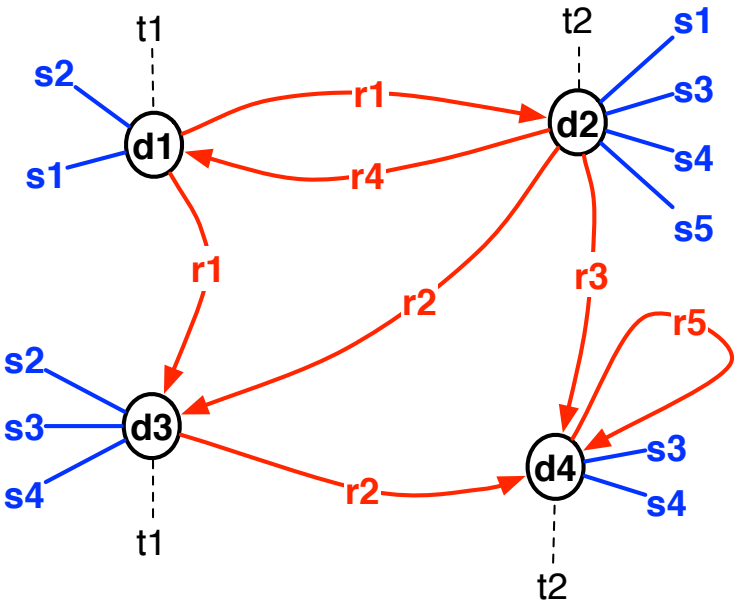
[Mimouni, 2015]

Multigraphe attribué et orienté

- Noeuds : documents
- Attributs
  - types de doc.
  - descripteurs sémantiques
- Arêtes : relations intertextuelles typées

Propriétés

- Nombreux documents
- Nombreux descripteurs
- Plusieurs types de liens
- Pas de contrainte topogique



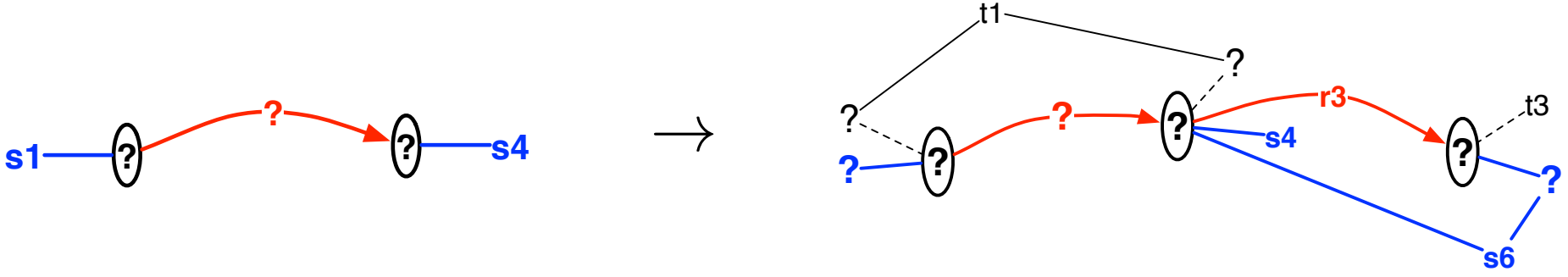
Modélisation: hiérarchies de propriétés

- Les descripteurs sont structurés en thesaurus
- Le raisonnement sur la hiérarchie des normes et des documents juridiques est au coeur du raisonnement juridique
- Les relations intertextuelles sont elles-mêmes organisées en hiérarchie

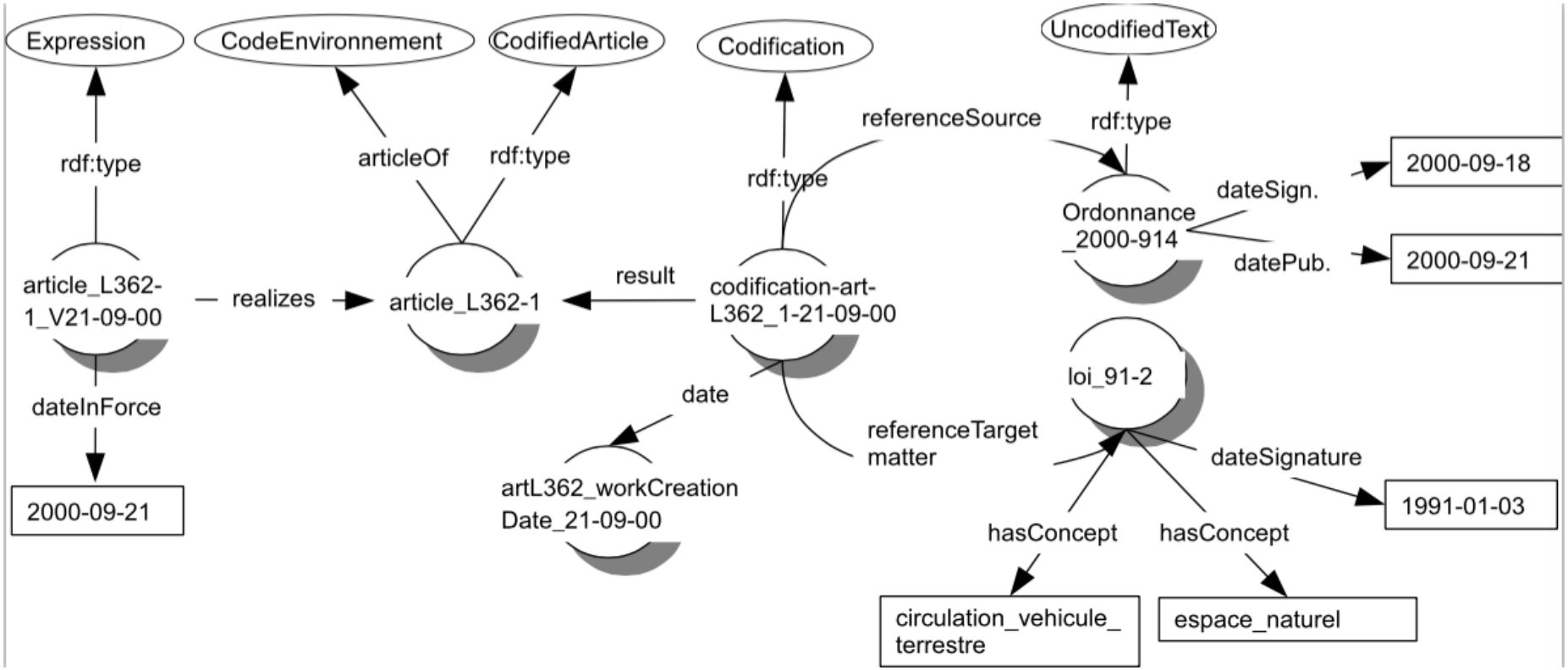
la décision de la cour X statue sur la décision antérieure de la cour Y

Confirmation ou annulation ?

→ Etendre le langage de requête



Représentation des collections (RDF)



Interrogation (SPARQL)

Je voudrais la dernière version (ou la version en vigueur) de l'article L362-1 du code de l'environnement.

```
SELECT ?article
WHERE {
  ?article rdf:type :Article .
?article dct:title "Code de l'environnement - Article L. 362-1" .
?article :dateInForce ?date .
} ORDER BY DESC(?date) LIMIT 1
```

```
SELECT ?version
WHERE {
  ?version rdf:type lido:DocumentaryUnitExpression .
?version metalex:realizes :article_L362-1 .
?version lido:dateInForce ?date .
} ORDER BY DESC(?date) LIMIT 1
```

Interrogation relationnelle : bilan

- Un mécanisme simple reposant sur
  - les technologies du web sémantique
  - une forte structuration des sources juridiques (en cours)
  - l'annotation systématique des références (existant)
- Un savoir faire à généraliser

Outline

Introduction

Accès à l'information juridique

Accéder aux sources

Accéder aux règles
- Reformulation
- Annotation à gros grain

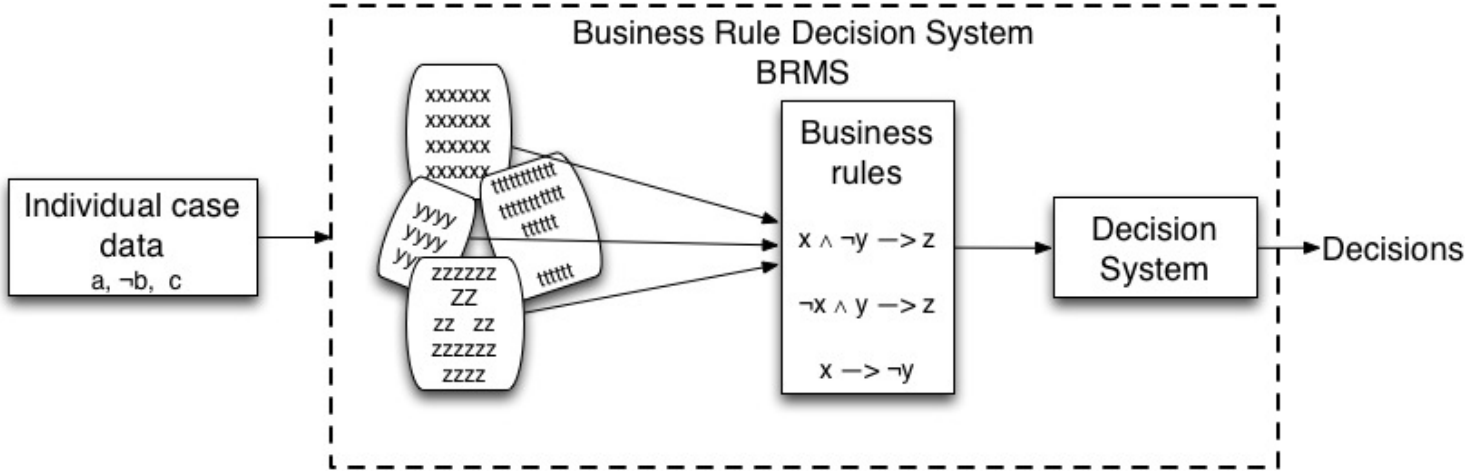
Conclusion et perspectives

L'accès aux règles

Un enjeu de sécurité juridique

- Contrôler la sécurité juridique des décisions
- Gérer les droits attachés aux objets, données, personnes
- Comparer et confronter des réglementations entre elles
- Vérifier et maintenir la cohérence des documents sources

→ Intégrer les textes juridiques et réglementaires à la prise de décision

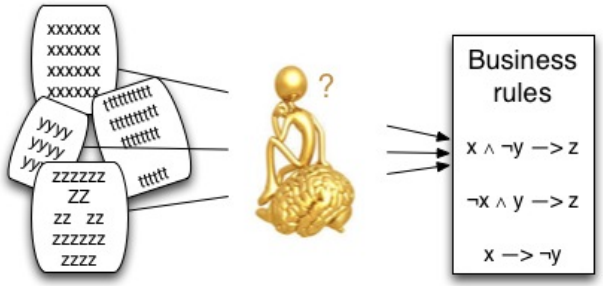


L'accès aux règles

Deux approches

- La reformulation des règles discursives en règles formelles
- L'annotation à gros grain des règles présentes dans les textes

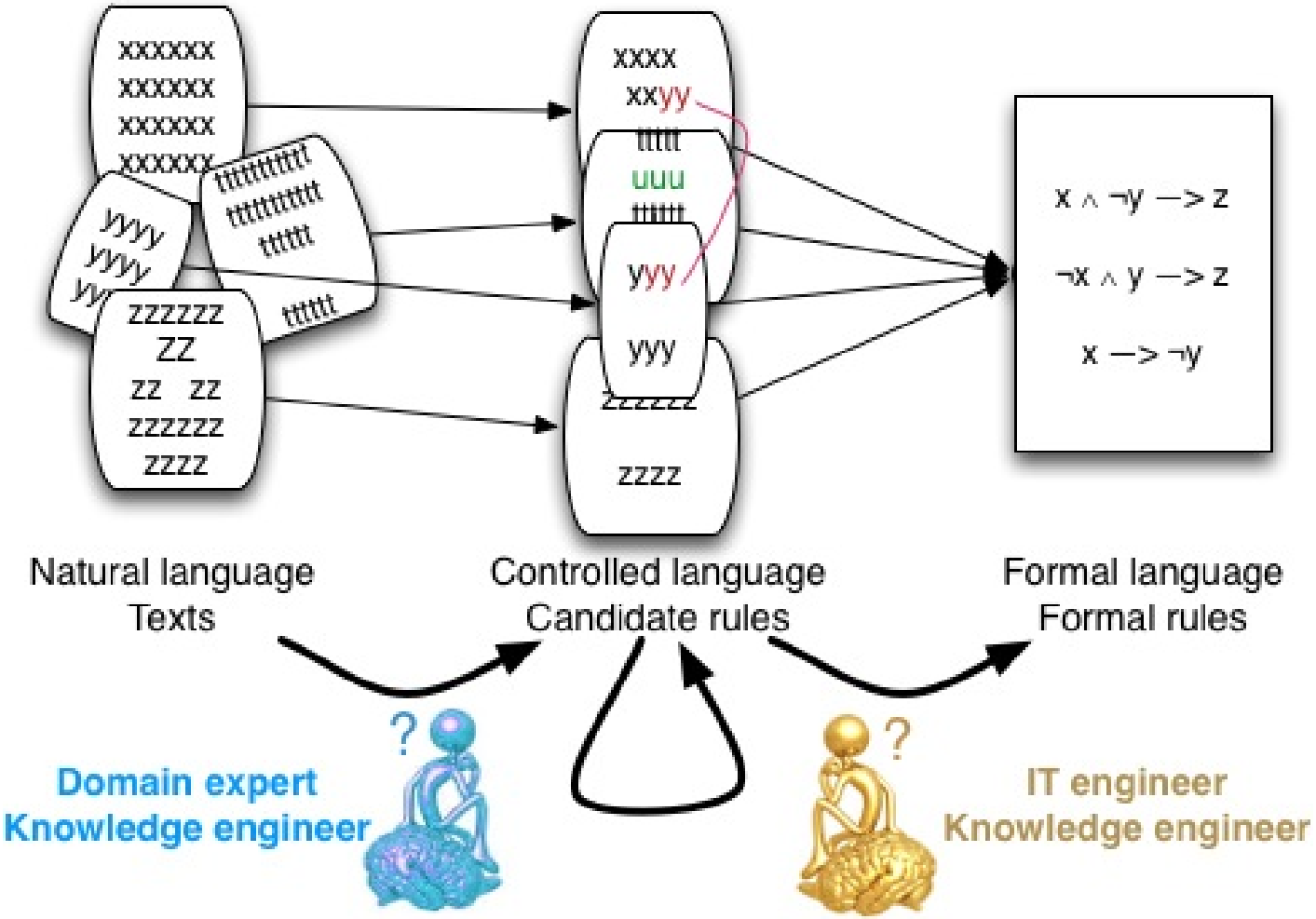
Le problème de la formalisation



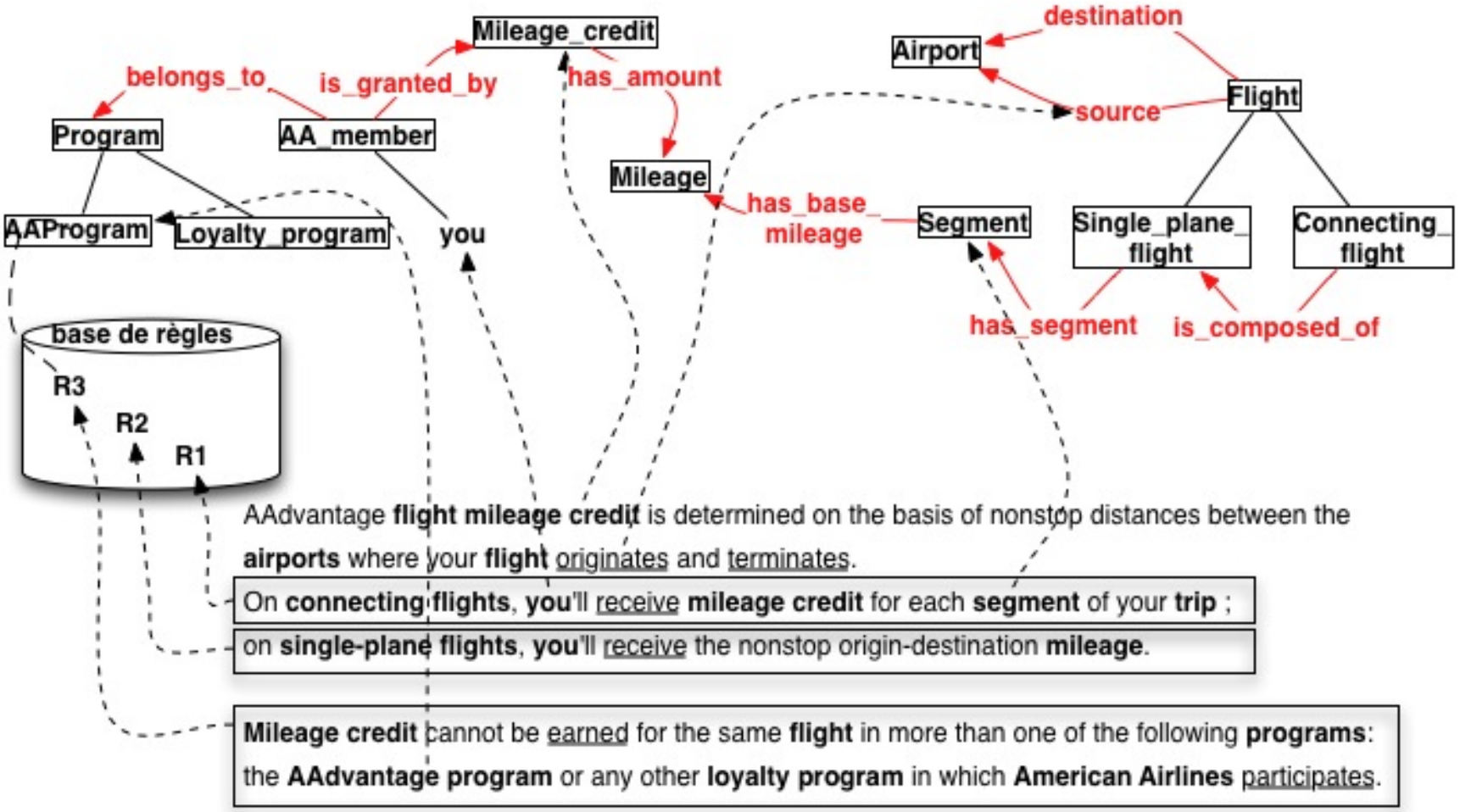
- Les langages naturel et formel se situent aux deux extrémités du continuum de formalisation [Baumeister *et al.*, 2011]
- les langues naturelles sont complexes
  - Information factuelle et éléments rhétoriques
  - Information redondante et implicite
  - Ambiguïté lexicale and structurelle
- La traduction directe et automatique en langage formel est impossible
  - Le choix du langage formel dépend du système cible
  - Les approches existantes s'attaquent à des problèmes simplifiés

[Unger *et al.*, 2012] [Dinesh *et al.*, 2008] [Bajwa *et al.*, 2011]

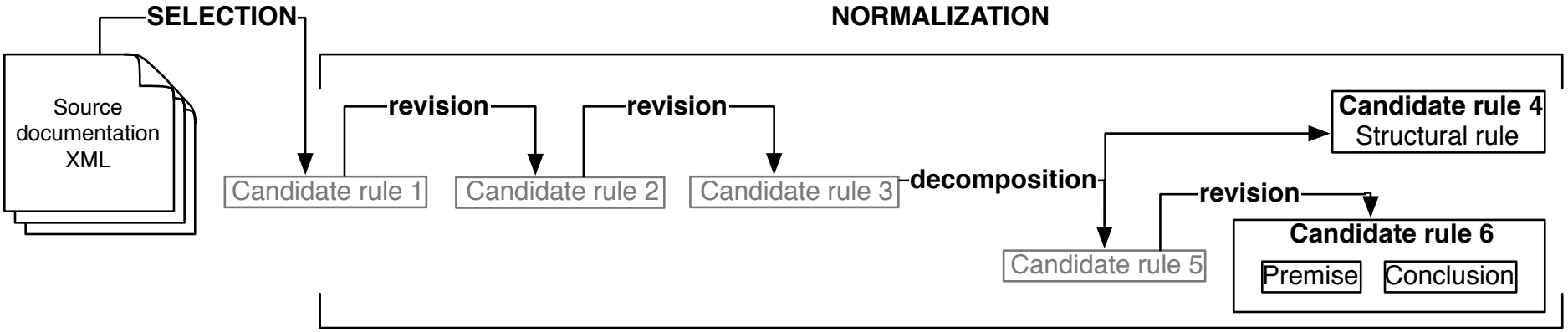
La formalisation, un processus collaboratif



Annotation r/ ontologie et base de règles



Un processus itératif



Input Documentation source

Process Des fragments de règles sont extraits des textes sources (langage naturel)
Des opérations de normalisation sont appliquées en séquence (contrôle man.)

Output Un ensemble de règles indépendantes, décontextualisées, désambiguïsées et éventuellement structurées en prémisses et conclusions (langage contrôlé)

IF a test has a duration less than 6 hours, THEN the test is InvalidTest

SemEx rule editor

SemEx

File Edit View Settings

Search Engine | Annotator | Rule Editor | Navigator

Base Rules | Pattern base

- Rule
  - R24[In that case, when the dynamic test ha
    - R25[In that case, when the dynamic_test
      - R26[if the SafetyBeltAssembly is inte
        - R27[if the SafetyBeltAssembly is i
          - R28[if the SafetyBeltAssembly i
  - R18[A 5 daN load shall be attached to the lo
    - R19[For the micro_slip_test a 5 daN load
      - R20[For the micro_slip_test a 5 daN lo
        - R23[For the micro_slip_test the tot
        - R22[For the micro_slip_test the up
        - R21[For the micro_slip_test a 5 da
  - R13[The test shall be carried out at a tempe
    - R14[The micro_slip_test shall be carried
      - R15[The temperature of the micro_sli
        - R16[The temperature of the micro_
        - R17[The temperature of the micro_
  - R5[If the strap slips or breaks at or within 1
    - R6[If the strap slips or breaks at or withi
      - R7[If the strap slips or breaks at or wi
        - R10[If the strap breaks at or within
        - R9[If the strap slips then the TestO
      - R8[If the strap slips or breaks at or wi
        - R12[If the strap breaks at or within

View a rule | Edit a rule | Browse the rules

Initial version

In that case, when the dynamic test has been carried out for a type of **vehicle** it need not be repeated for other types of **vehicle** where each **anchorage point** is less than 50 mm distant from the corresponding **anchorage point** of the tested **belt**.

Editor

In that case, when the dynamic test has been carried out for a type of **vehicle** it need not be repeated for other types of **vehicle** where each **anchorage point** is less than 50 mm distant from the corresponding **anchorage point** of the tested **belt**.

Parameters

Type :

Pattern :

Premise :

Conclusion :

Revisions :

Refers to : R24, R25, R26

Subrule of : R24, R25, R26

User name :

Editing date : 06/02/2011

Informations

Type : Concept

[Guissé, 2013]

Sélection des règles à annoter

La sélection des règles requiert une bonne connaissance du contexte métier et de la documentation source

- L'annotation sémantique facilite la lecture des textes sources
- La recherche de mots-clés permet d'identifier les fragments de texte pertinents

  ex. marqueurs modaux, connecteurs logiques

- L'expert peut appliquer des requêtes Sparql combinant des expressions régulières et des étiquettes sémantiques

Opérations de normalisation

- Normalisation lexicale
  - Remplacement par les "termes préférés"
  - Désambiguisation
- Décontextualisation : pronoms, termes génériques, références
- Normalisation syntaxique :

  ré-ordonnancement des phrases, segmentation des énumérations

  décomposition des règles
- Restoration sémantique : entités implicites, intervalles de contraintes

Expériences

Analyse détaillée de 2 ensemble de règles candidates

- American Airlines: 95 règles candidates

Opérations de normalisation	% de règles candidates normalisées
Lexicale	65%
Contextuelle	64%
Syntaxique	100%
$\rightarrow$*Décomposition*	*30%*
$\rightarrow$*Restructuration*	*68%*

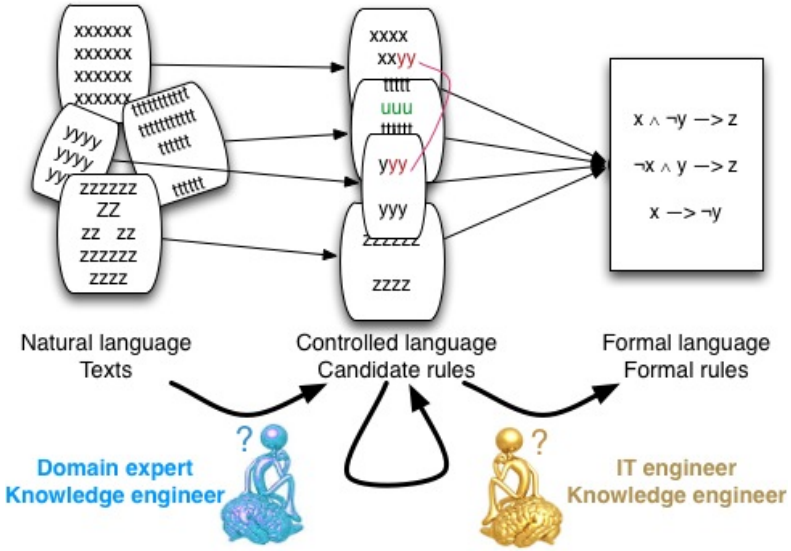
- Audi : 100 règles candidates

Opérations de normalisation	% de règles candidates normalisées
Lexicale	61%
Contextuelle	57%
Syntaxique	100%
$\rightarrow$*Décomposition*	*40%*
$\rightarrow$*Restructuration*	*32%*

Bilan

- Un outil pour guider l'extraction et la normalisation des règles (SemEx)

  [Lévy *et al.*, 2010a][Guissé, 2013]
- Une approche de normalisation difficile à mettre en oeuvre
  - Variabilité des corpus et des patrons d'extraction
  - Langage cible mal défini SBVR Structured English
- Une étape cruciale pour la formalisation des règles métiers
- Approche coûteuse dédiée à des contextes applicatifs précis

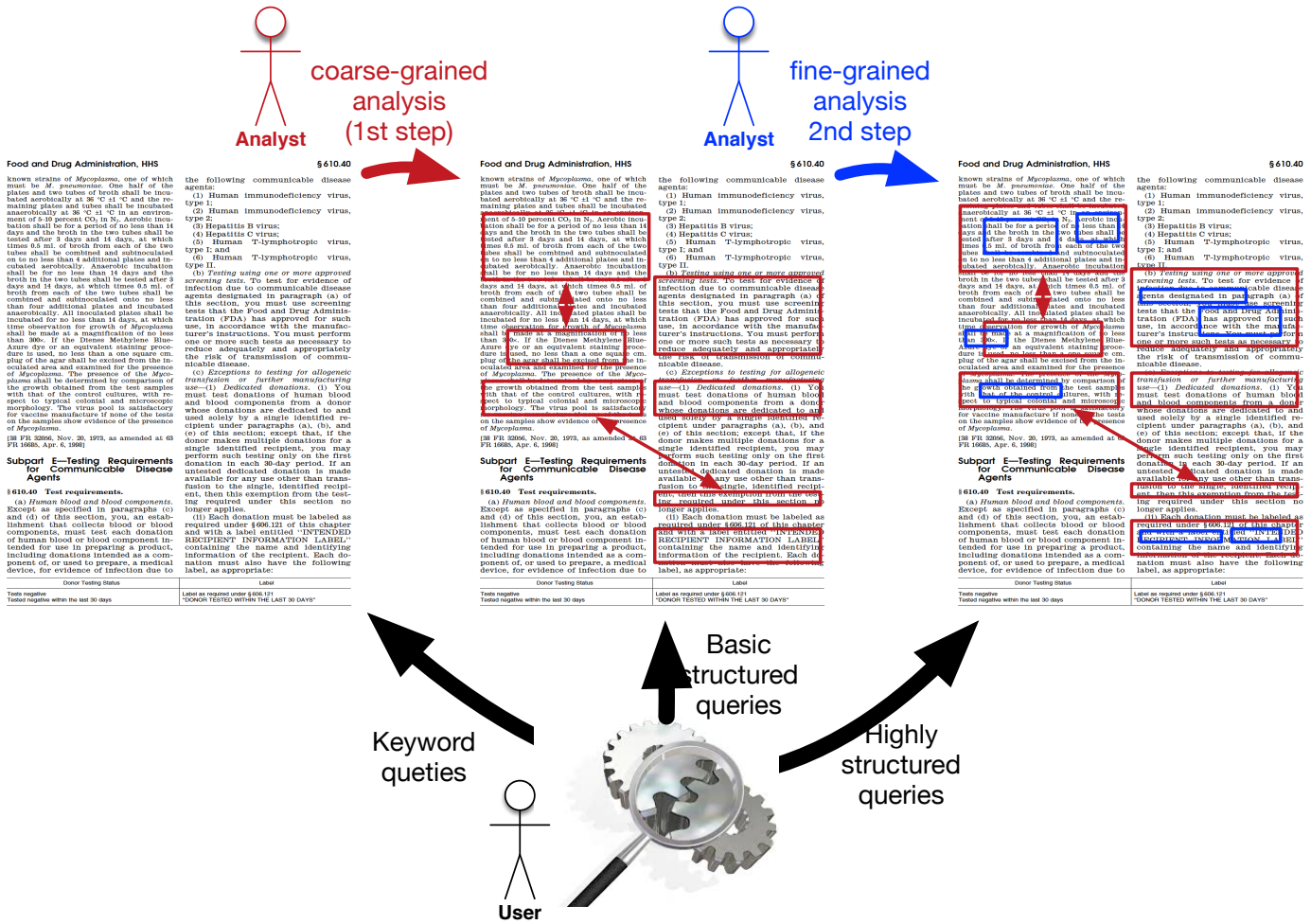


Annotation à gros grain

- Objectif [Nazarenko *et al.*, 2016][Wyner *et al.*, 2016]
  - Analyser un corpus réglementaire
  - Raisonner automatiquement sur les règles

    Qui a le droit de faire quoi et où ?
  - Explorer et comprendre la loi
  - Aider les rédacteurs et le législateur à faire la loi
  - Faciliter l'accès à la loi pour les citoyens, les acteurs économiques et publics
- Expériences [Wyner *et al.*, 2017]
  - Scottish Legal Intruments
  - Réglementation sur la vente et la distribution de tabacs manufacturés

Annotation à gros grain



Legal RuleML

- Bénéficier de l'interopérabilité d'un langage du web sémantique

  RuleML

- S'interfacer avec un langage de règles dédié

  Legal RuleML[1]

  [Athan *et al.*, 2015]

  - Atom, if, then, Rule, Rel(ation), Ind(ividual), Formula, Var, slot, Time, Data, And, Or, Negation
  - Agent, Authority, AuxiliaryParty, Bearer, ConstitutiveStatement, Context, DefeasibleStrength, FactualStatement, Jurisdiction, LegalSource, Obligation, Override, PenaltyStatement, Permission, Prohibition, Reparation, Right, Role, StrictStrength, TemporalCharacteristics, Violation

- Utiliser un LRML "à gros grain" pour l'annotation

[1] OASIS open access/open-source standard XML markup language

Exemple de texte

Smoking Prohibition (Children in Motor Vehicles) (Scotland) Act 2016

2016 asp 3

The Bill for this Act of the Scottish Parliament was passed by the Parliament on 17th December 2015 and received Royal Assent on 21st January 2016

An Act of the Scottish Parliament to prohibit smoking in private motor vehicles in the presence of children, subject to limited exceptions; and for connected purposes.

(1) It is an offence for an adult to smoke in a private motor vehicle when—

(a) there is a child in the vehicle, and

(b) the vehicle is in a public place.

(2) Subsection (1) does not apply to a private motor vehicle that is designed or adapted for use as living accommodation and which, at the time the smoking occurs, is parked and is being used as living accommodation.

(3) A person who commits an offence under subsection (1) is liable on summary conviction to a fine not exceeding level 3 on the standard scale.

2 Fixed penalties

The schedule provides for fixed penalties for offences under section 1.

3 Powers to enter and require information

(1) An authorised officer of a local author[illegible] search a private motor vehicle which is

Objectifs

- Répondre aux questions des utilisateurs juristes
  - What are the offences and associated penalties or defences?
  - What prohibitions apply to tobacco products?
  - What obligations have been placed on what entities, *e.g.* shop owners?
  - What permissions are given to Scottish Ministers
  - Given a provision, what are related overriding or reparation provisions?
- Annotation
  - Comment catégoriser les différentes portions de textes ?

    prohibition, obligation, penalty, etc.
  - Comment relier ces portions de textes entre elles ?

    exception, réparation, etc.

Démarche

- Etapes
  - Annotation manuelle du corpus dans un langage d'annotation "simple"
    - Segmentation du texte en fragments
    - Catégorisation des fragments pertinents
    - Mise en relation des fragments
  - Transformation en LegalRuleML
  - Interrogation du corpus formellement annoté
  - Analyse de la règlementation
- Expérience
  - Approche agile, en ligne et collaborative
  - 6 étudiants ont annoté une dizaine de textes
  - Mise au point incrémentale du guide d'annotation
- Résultat
  - Corpus annoté
  - Guide d'annotation
  - Outil d'interrogation

Annotations

Le "smoking corpus" a été annoté avec[2]

- des éléments ($\sim$ 600)
  - PrescriptiveStatement
    - Permission
    - Obligation
    - Prohibition
  - ConstitutiveStatement : definition
  - Penalty
- des relations entre éléments
  - OverrideStatement : exception
  - ReparationStatement

relation entre une pénalité et la prescription associée

[2] Pour les definitions, voir le document LegalRuleML Core Specification

Outil d'interrogation

2010ActAsAmended_EO_AF_AW_

lrml.ddns.net/doc/show/2010ActAsAmended_EO_AF_AW_RS1.xml

Search

LRML Search | Documents | Search | Login

2010ActAsAmended_EO_AF_AW_RS1.xml

Go Back | Download | Delete

Tobacco and Primary Medical Services (Scotland) Act 2010 asp 3 (Scottish Act) Part 1 TOBACCO AND NICOTINE VAPOUR PRODUCTS ETC. Chapter 1 DISPLAY, SALE AND PURCHASE Display of tobacco products etc.

1 Prohibition of tobacco displays etc.

[#ps1]

(1) A person who in the course of business displays or causes to be displayed tobacco products or smoking related products in a place where tobacco products are offered for sale commits an offence.

Overriden by: [#ps2], [#ps3]

Has reparation: [#rs5]

[#ps2]

(2) A person does not commit an offence under subsection (1) if the display - (a) is in a specialist tobacconist, (b) does not include cigarettes or hand-rolling tobacco, – and (c) Complies with any prescribed requirements.

Overrides: [#ps1]

[#ps3]

(3) A person does not commit an offence under subsection (1) if (a) the tobacco products or smoking related products are displayed in the course of a business involving the sale of tobacco

- Constitutive
- Factual
- Penalty
- Prescriptive
  - Obligation
  - Permission
  - Prohibition
- Reparation

Bilan

- 1ère étape
  - On peut répondre aux questions des utilisateurs
  - L'approche globale est validée : langage et guide d'annotation assimilables par des juristes, outil de consultation, etc.
  - Les difficultés d'interprétation sont identifiées
    - Sémantique des modalités : *must*, *may*, *shall not*
    - Portée des opérateurs
    - Formulation des exceptions
    - Modélisation des réparations
- Mise en production
  - Intégrer la structure des documents — LegalDocML
  - Mettre les outils aux mains des juristes
  - Evaluer la qualité des annotations
  - Etendre l'expérience à d'autres corpus juridiques
- Travaux futurs
  - Apprendre à annoter automatiquement ?
  - Affiner la granularité du langage d'annotation ?

Outline

Introduction

Accès à l'information juridique

Accéder aux sources

Accéder aux règles

Conclusion et perspectives

Conclusion

- L'accès à l'information juridique est un enjeu
  - Retrouver des documents, tous les documents sur un sujet donné
  - Raisonner avec les règles exprimées dans les textes juridiques
- Cela passe par l'annotation sémantique
  - Structuration des documents
  - Annotation des relations entre documents
  - Annotation lexicale et conceptuelle (vocaculaires & ontologies)
  - Identification, normalisation, structuration des règles ($\neq$ Formalisation)
- Mais les défis restent nombreux
  - Complexité et variabilité de la langue juridique
  - Diversité des usages
  - Ouverture et partage de données, ressources et interprétations

Le Legal Semantic Web se construit mais il reste beaucoup à faire !

Merci de votre attention

References I

Florence Amardeilh, Philippe Laublet, and Jean-Luc Minel.

Document annotation and ontology population from linguistic extractions.

In *Proceedings of the 3rd international conference on Knowledge capture (K-CAP '05)*, pages 161–168, New York, NY, USA, 2005. ACM.

Tara Athan, Guido Governatori, Monica Palmirani, Adrian Paschke, and Adam Z. Wyner.

Legalruleml: Design principles and foundations.

In Wolfgang Faber and Adrian Paschke, editors, *Reasoning Web. Web Logic Rules - 11th Int. Summer School, Berlin, Germany, 2015, Tutorial Lectures*, pages 151–188. Springer, 2015.

References II

Imran Sarwar Bajwa, Mark G. Lee, and Behzad Bordbar.

Sbvr business rules generation from natural language specification.

In *AAAI Spring Symposium 2011 Artificial Intelligence 4 Business Agility*, pages 541–545, San Francisco, USA, 2011. AAAI.

Joachim Baumeister, Jochen Reutelshoefer, and Frank Puppe.

Engineering intelligent systems on the knowledge formalization continuum.

International Journal of Applied Mathematics and Computer Science (AMCS), 21(1), 2011.

Michael J. Bommarito and Daniel Martin Katz.

Properties of the united states code citation network.

ArXiv e-prints, November 2009.

References III

Romain Boulet, Pierre Mazzega, and Danièle Bourcier.

A network approach to the french system of legal codes- part i: Analysis of a dense network.

Journal of Artificial Intelligence and Law, 19:333–355, 2011.

Sergey Brin and Lawrence Page.

The anatomy of a large-scale hypertextual web search engine.

In *Proceedings of the seventh international conference on World Wide Web (WWW7)*, pages 107–117, Amsterdam, The Netherlands, 1998. Elsevier Science Publishers B. V.

A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener.

Graph structure in the Web.

Comput. Netw., 33(1-6):309–320, June 2000.

References IV

Jeff Conklin.

Hypertext: An introduction and survey.

IEEE Computer, 20(9):17–41, 1987.

Sylvie Després and Sylvie Szulman.

Merging of Legal Micro-Ontologies from European Directives.

In *IAAIL Workshop series*, pages 149–155, 2005.

Nikhil Dinesh, Aravind Joshi, Insup Lee, and Oleg Sokolski.

Reasoning about conditions and exceptions to laws in regulatory conformance checking,.

In *Proc. of the 9th Int. Conf. on Deontic Logic in Computer Science*, page 16 pages, 2008.

References V

James H. Fowler, Timothy R. Johnson, James F. Spriggs II, Sangick Jeon, and Paul J. Wahlbeck.

Network analysis and the law : Measuring the legal importance of precedents at the u.s. supreme court.

Political Analysis, 15:324–346, 2007.

Jean-Gabriel Ganascia, Pierre Glaudes, and Andrea Del Lungo.

Automatic detection of reuses and citations in literary texts.

CoRR, abs/1404.2997, 2014.

Aldo Gangemi.

Introducing pattern-based design for legal ontologies.

In *Proceedings of the 2009 Conference on Law, Ontologies and the Semantic Web: Channelling the Legal Information Flood*, pages 53–71, Amsterdam, The Netherlands, The Netherlands, 2009. IOS Press.

References VI

Gérard Genette.

Palimpsestes.

Poétique. Le Seuil, 1982.

Abdoulaye Guissé.

Une plateforme d'aide à l'acquisition et à la maintenance des règles métier à partir de textes réglementaires.

PhD thesis, Université Paris 13, 2013.

Thèse de doctorat dirigée par Lévy, François, Informatique.

Dincer Gultemen and Tom van Engers.

Graph-based linking and visualization for legislation documents (glvd).

In *Network Analysis in Law Workshop (NAiL2013@ICAIL) associated with the International Conference on AI and Law (ICAIL 2013*, Rome, Italy,, June 2013.

References VII

Rinke Hoekstra, Joost Breuker, Marcello Di Bello, and Alexander Boer.

The LKIF core ontology of basic legal concepts.

In Pompeu Casanovas, Maria Angela Biasiotti, Enrico Francesconi, and Maria Teresa Sagri, editors, *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*, June 2007.

Jon M. Kleinberg.

Authoritative sources in a hyperlinked environment.

Journal of the acm, 46(5):604–632, 1999.

François Lévy, Adeline Nazarenko, Abdoulaye Guissé, Nouha Omrane, and Sylvie Szulman.

An environment for the joint management of written policies and business rules.

In *Proceedings of the International Conference on Tools with Artificial Intelligence (IEEE-ICTAI 10)*, pages 142–149, 2010.

References VIII

François Lévy, Adeline Nazarenko, and Abdoulaye Guissé.

Annotation, indexation et parcours de documents numériques.

Revue des Sciences et Technologies de l'Information (Série IDocument Numérique), 13(3/2010):121–152, December 2010.

Yue Ma, Adeline Nazarenko, and Laurent Audibert.

Formal description of resources for ontology-based semantic annotation.

In *Proceedings of the International Conference on Language Resources and Evaluation*, Malta, May 2010.

Bernardo Magnini, Emanuele Pianta, Octavian Popescu, and Manuela Speranza.

Ontology population from textual mentions: Task definition and benchmark.

In *Proceedings of the 2nd Workshop on Ontology Learning and Population: Bridging the Gap between Text and Knowledge*, pages 26–32, Sydney, Australia, July 2006. Association for Computational Linguistics.

References IX

Nada Mimouni.

Interrogation d'un réseau sémantique de documents : l'ntertextualité dans l'accès à l'information juridique, Thèse de Doctorat en Informatique.

PhD thesis, Université Paris 13 – Sorbonne Paris Cité, 2015.

Adeline Nazarenko, François Lévy, and Adam Wyner.

Towards a methodology for formalizing legal texts in legalruleml.

In Floris Bex and Serena Villata, editors, *Legal Knowledge and Information Systems - JURIX 2016: The Twenty-Ninth Annual Conference*, volume 294 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, 2016.

Monica Palmirani, Luca Cervone, and Fabio Vitali.

Legal metadata interchange framework to match cen metalex.

In *ICAIL*, pages 232–233, 2009.

References X

- Borislav Popov, Atanas Kiryakov, Damyan Ognyanoff, Dimitar Manov, and Angel Kirilov.

  Kim – a semantic platform for information extraction and retrieval.

  Nat. Lang. Eng., 10(3-4):375–392, 2004.

- G. Sartor, M. Palmirani, and E. Francesconi.

  Legislative XML for the Semantic Web: Principles, Models, Standards for Document Management.

  Law, governance and technology series. Springer, 2011.

- Giovanni Sartor, Monica Palmirani, Enrico Francesconi, and Maria Angela Biasiotti.

  Law, Governance and Technology: Legislative Xml for the Semantic Web: Principles, Models, Standards for Document Management.

  Law, Governance and Technology Series, 4. Springer London, Limited, 2011.

References XI

Christina Unger, Lorenz Bühmann, Jens Lehmann, Axel-Cyrille Ngonga Ngomo, Daniel Gerber, and Philipp Cimiano.

Sparql template based question answering.

In *21st International World Wide Web Conference (WWW2012)*, April 2012.

Maria Vargas-Vera, Enrico Motta, John Domingue, Mattia Lanzoni, Arthur Scutt, and Fabio Ciravegna.

Mnm: Ontology-driven tool for semantic markup.

In Siegfried Handschuh, Niegel Collier, Rose Dieng, and Steffen Staab, editors, *Proceedings Workshop on Semantic Authoring, Annotation & Knowledge Markup (SAAKM 2002)*, pages 43–47, Lyon, France, 2002.

Radboud Winkels and Jelle de Ruyter.

Survival of the fittest: Network analysis of dutch supreme court cases.

In *AICOL*, pages 106–115, 2011.

References XII

Radboud Winkels, A Boer, and R Hoekstra.

Metalex: An xml standard for legal documents.

In *Proceedings of the XML Europe Conference, London (UK)*, 2003.

Adam Wyner, Adeline Nazarenko, and Francois Lévy.

Towards a high-level controlled language for legal sources on the semantic web.

In Brian Davis, J. Gordon Pace, and Adam Wyner, editors, *Proc. of the 5th Int. Workshop on Controlled Natural Language (CNL2016)*, pages 92–101, Aberdeen, UK, July 2016. Springer.

References XIII

Adam Wyner, Fraser Gough, Francois Lévy, Matt Lynch, and Adeline Nazarenko.

On annotation of the textual contents of scottish legal instruments.

In Adam Z. Wyner and Giovanni Casini, editors, *Legal Knowledge and Information Systems - JURIX 2017: The Thirtieth Annual Conference, Luxembourg, 13-15 December 2017*, volume 302 of *Frontiers in Artificial Intelligence and Applications*, pages 101–106. IOS Press, 2017.