

Structuration des décisions de jurisprudence basée sur une ontologie juridique en langue arabe

Karima Dhouib*, Sylvie Després**
Faiez Gargouri***

*ISET - Sfax Tunisie, BP : 88A Elbustan ; Sfax
karima.dhouib@isets.rnu.tn,

** LimBio -74, rue Marcel Cachin 93017 Bobigny Cedex France
sylvie.despres@univ-paris13.fr

*** ISIMS-Sfax Tunisie, Route de Tunis Km 10 B.P. 242- SFAX 3021
faiez.gargouri@fsegs.rnu.tn

Résumé. L'informatique juridique, est un domaine en évolution constante. Le contexte général de notre travail est l'élaboration d'un système de recherche de jurisprudence tunisienne en langue arabe. L'objectif opérationnel de ce système est de fournir une aide aux juristes pour résoudre une situation juridique donnée en mettant à leur disposition une collection de situations similaires ce qui améliorera leur raisonnement futur. Une ontologie du domaine juridique construite à partir des documents des décisions juridiques est nécessaire dans notre contexte. Cette ontologie a pour but : (i) la structuration des décisions, (ii) la formulation des requêtes d'interrogation de la base des décisions, et (iii) la recherche des décisions. Dans cet article, nous présentons l'architecture de notre système de recherche de jurisprudence. Nous nous focalisons sur l'ontologie du domaine de jurisprudence que nous avons élaborée, ainsi que sur le module de structuration des décisions.

1 Introduction

Nos travaux de recherche se situent dans le cadre de l'informatique juridique documentaire. Parmi les objectifs figurent l'extraction et la recherche d'information dans des bases de données juridiques et l'élaboration et l'utilisation des thésaurus et des ontologies juridiques. Le contexte général de nos travaux est l'élaboration d'un système permettant de rechercher des décisions de jurisprudence tunisienne en langue arabe. Il aidera le juriste à résoudre une situation juridique particulière en mettant à sa disposition une collection de situations similaires antérieures. En Tunisie, deux sources de jurisprudence sont disponibles. Les revues et les collections sont périodiquement mises à jour. La base de données de jurisprudence tunisienne, mise en place en 2009, est constituée des 12000 décisions de la cour de cassation depuis l'année 1959. Cette base est accessible *via* le portail de la justice et des droits de l'Homme en Tunisie¹. Selon les juristes, la recherche dans les revues est très consommatrice de temps. La

1. <http://www.e-justice.tn/>

Structuration des décisions de jurisprudence basée sur une ontologie juridique en langue arabe

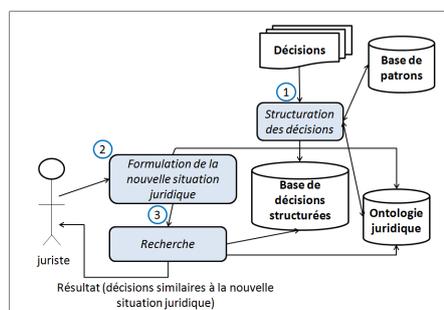


FIG. 1 – Architecture fonctionnelle du système

recherche dans la base de données de jurisprudence quant à elle, ne garantit pas la récupération de documents utiles car elle est généralement fondée sur des critères simples de recherche. Nous proposons que notre système de recherche de jurisprudence tunisienne tienne compte des avancées dans le domaine de la recherche d'information en utilisant notamment des ressources sémantiques telle qu'une ontologie juridique en langue arabe. Ce système offrira ainsi aux juristes tunisiens un gain de temps ainsi que des résultats de recherche plus pertinents. L'ontologie, servant à décrire le contenu sémantique des sources textuelles disponibles, est déployée à plusieurs niveaux afin : d'améliorer les résultats de structuration des décisions de jurisprudence, de faciliter la formulation des requêtes lors de l'interrogation de la base des décisions et finalement d'optimiser la recherche des décisions. Dans cet article, nous nous intéressons à la structuration des décisions. Ainsi, nous commençons par introduire le contexte de nos travaux, puis nous présentons notre méthode de structuration des décisions de jurisprudence. Avant de conclure ce papier, nous détaillons ensuite l'étape de construction de l'ontologie.

2 Architecture du système de recherche de jurisprudence

Dans notre système, quatre composantes sont intégrées : (i) un corpus de décision de jurisprudence, (ii) une ontologie du domaine juridique en langue arabe servant à la description du contenu sémantique du corpus, (iii) une base de patrons rassemblant des marqueurs qui délimitent le début de chaque partie d'une décision et (iiii) finalement une base de décisions structurées qui constitue le résultat de l'étape de structuration des décisions. La construction du système comporte trois phases qui sont :

- **phase 1** : la structuration des décisions. Au cours de cette phase, une analyse des décisions est effectuée afin d'identifier une structure organisationnelle commune et construire les patrons pouvant délimiter les différentes parties de la structure retenue. Il s'agit, ensuite, de structurer les décisions, en se fondant sur les patrons ainsi que sur une ontologie juridique.
- **phase 2** : le nouveau problème juridique est décrit en utilisant une ontologie juridique.
- **phase 3** : un processus de recherche fondé sur l'ontologie sera élaboré afin de renvoyer les décisions les plus similaires au nouveau problème.

3 Structuration des décisions

Une décision de jurisprudence se présente sous la forme d'un document textuel de deux à cinq pages. Les décisions partagent une régularité de contenu. Quand le juriste parcourt une décision, il met beaucoup de temps pour délimiter ses différentes parties et déterminer celles pouvant l'aider à résoudre d'un point de vue juridique son nouveau problème. De ce fait, et contrairement aux systèmes de recherche de jurisprudence existant actuellement, nous envisageons à travers notre système, de fournir au juriste des décisions non dans leurs états textuels initiaux, mais plutôt, sous forme structurée. Dans cette partie, nous présentons notre approche pour la structuration des décisions. La méthodologie adoptée dans ce travail consiste en un ensemble d'étapes ordonnées. La première étape est la constitution du corpus d'apprentissage. Il est composé d'un ensemble de décisions de différents types, annotées manuellement par les experts. L'étape2 est celle de l'analyse de ce corpus pour identifier et définir des patrons. L'étape3 consiste à structurer les décisions.

3.1 Analyse des décisions

Notre corpus d'étude est composé de 600 décisions rendues par la cour de cassation, disponibles sur le portail de justice et des droits de l'Homme en Tunisie². L'étape d'analyse a été réalisée en étroite collaboration avec les experts juristes afin de déterminer la manière de structurer les documents textuels contenant les décisions. Lire une décision, n'est pas toujours évident. L'aisance ne s'acquiert qu'au fil de l'expérience³. Au cours de la lecture, le juriste examine la décision de manière globale, et tente de reconnaître les blocs descriptifs et les blocs de raisonnement. Une cinquantaine de décisions (arrêt) de différents types (pénal, civil, administratif, criminel) ont été analysées afin de repérer les régularités dans la structuration de leur contenu. Une structure organisationnelle des décisions a été définie. Une décision est composée de :

- données générales : elles comportent le numéro de l'arrêt, son type, sa date de publication, le volume dans lequel il a été publié, la matière, les références légales citées ainsi que quelques mots-clés.
- le principe⁴ : il contient souvent l'essentiel de l'arrêt. Il est généralement décrit sous une forme abstraite sans prendre en compte des faits d'espèce. Il peut toutefois être spécialement adapté à la solution et apporter de ce fait des précisions.
- Introduction : Elle décrit la décision présentée au tribunal.
- Les faits d'espèce⁵ : Ils représentent les faits ayant motivé la saisine d'un juge.
- les moyens au pourvoi⁶ : ils exposent de manière concise les critiques adressées à la décision attaquée.
- le raisonnement juridique : il constitue la partie essentielle de l'arrêt. Le juge donne son argumentation et la justification de la solution adaptée en partant des faits du litige et en donnant les références juridiques sur lesquelles il s'est appuyé Farzindar et al. (2004).

2. <http://jurisprudence.e-justice.tn/wwwisis/juris.09/form.htm>

3. <http://site.juristudiant.com/Analyse-de-decision.html>

4. [://site.juristudiant.com/Analyse-de-decision.html](http://site.juristudiant.com/Analyse-de-decision.html)

5. <http://www.dictionnaire-juridique.com/definition/espece.php>

6. <http://www.scribd.com/le-pourvoi-en-cassation/d/409507>

Structuration des décisions de jurisprudence basée sur une ontologie juridique en langue arabe

– conclusion : l'arrêt de cassation se termine systématiquement par un paragraphe bref, le dispositif, prononçant la cassation, la cassation partielle ou le rejet du pourvoi. Pour identifier ces différentes parties de la décision, l'expert juriste s'appuie sur des marqueurs du domaine juridique. Ces marqueurs vont nous servir de base pour élaborer les patrons qui sont utilisés pour structurer les décisions.

3.2 Définition des patrons et problématique de structuration des décisions

Suite à l'analyse des documents du corpus, nous avons construit, manuellement et avec l'aide de l'expert, les différents patrons possibles afin d'entamer l'étape de structuration des décisions. Nous avons classés ces patrons en fonction des parties auxquelles ils appartiennent (introduction, faits d'espèce,...). Un patron est un ensemble de descripteurs indispensables pour marquer le début d'une partie d'une décision. Un terme est choisi de sorte qu'il figure dans tous les marqueurs d'une même partie. Dans cet article, nous limitons la présentation du patron à la partie faits d'espèce. En examinant des extraits de décisions, nous pouvons constater que certaines phrases, marquant le début de la partie "faits d'espèces" et ayant le même sens tel que "attendu, selon l'arrêt attaqué et les faits d'espèces", peuvent avoir des formulations différentes. Soit une instance d'un patron P que nous avons construit, trois problèmes peuvent survenir, si nous voulons contrôler sa présence dans les textes des décisions et qui sont présentés dans la figure 2.

Soit P= الحث + أتضح + الحكم + المنتقد + الوقائع (P= attendu que + arrêt+ attaqué+faits d'espèce)
1- Le terme est désigné par ses synonymes dans des textes différents. Par exemple, pour désigner l'expression « il s'avère que », plusieurs termes synonymes peuvent être utilisés : تبين , تفيد , إتضح :
2- Le même terme apparaît sous différentes formes morphologiques. Par exemple l'expression « il s'avère que » peut être désignée avec le terme sous les formes أفاد , تفيد , يستفاد .
3- On peut trouver des termes ayant le même sens juridique et cités différemment mais qui ne sont pas forcément des synonymes tels que الوقائع (les faits d'espèce) et الأبحاث في القضية.

FIG. 2 – Problèmes rencontrés lors du contrôle de présence des patrons

3.2.1 Solutions proposées

La complexité du domaine juridique est constituée par la spécificité des connaissances et des vocabulaires juridiques. En effet un même terme peut ne pas avoir de sens autre que juridique. Mais on peut également trouver des termes appartenant au vocabulaire de sens commun ayant non seulement au moins un sens juridique mais aussi au moins un sens non juridique Cornu (2005). Nous pouvons donner ici deux exemples et qui sont présentés dans la figure 3. Nous proposons de construire une ontologie juridique. Cette ontologie permettra d'élaborer un consensus entre les experts juristes et de résoudre l'ambiguïté de la variabilité terminologique. En outre, cette ontologie pourra être implémentée avec un langage formel tel qu'OWL afin de l'intégrer à notre système de recherche de jurisprudence et ainsi faire des inférences. Quant à la prise en compte des différentes formes morphologiques que peut prendre un terme, nous allons utiliser l'outil morph2. C'est un analyseur morphologique de la langue arabe qui

<p>-Le mot طعن peut signifier blesser ou causer une lésion, humilier mais juridiquement signifie attaquer un arrêt. - Le terme وجه désigne couramment le visage mais juridiquement peut exprimer un moyen au pourvoi.</p>
--

FIG. 3 – Exemple de termes juridiques

permet de reconnaître la composition des mots et de fournir des informations morphologiques spécifiques sur les mots Chaâben Kammoun et al. (2001).

4 Construction de l'ontologie juridique

D'après Gruber et Thomas (1993), une ontologie est la spécification explicite et formelle d'une conceptualisation partagée d'un domaine de connaissance. Les ontologies applicatives sont produites par des ingénieurs de la connaissance en vue de réduire la complexité du langage juridique et de faciliter le stockage et l'accès aux documents juridiques. Elles sont extraites à partir de larges corpus textuels Bourcier (2006). Nous proposons de construire une ontologie applicative pour le domaine juridique en langue arabe. Elle permettra de rassembler le matériau terminologique afin d'optimiser les étapes de structuration et de recherche des décisions. Un cadre méthodologique en quatre étapes (constitution d'un corpus de document, analyse linguistique du corpus, conceptualisation, opérationnalisation de l'ontologie) est commun à la plupart des méthodes de construction d'ontologies à partir de textes Mondary (2011). De ce fait, nous avons commencé par définir, à partir de notre corpus, les concepts juridiques et les relations les liant. L'étape d'extraction de concepts, qui vont être utilisés pendant la phase de structuration, a été réalisée manuellement. Le repérage de ces concepts était relativement facile car ils apparaissent dans les mêmes endroits des textes des décisions. La validation des concepts a ensuite été réalisée par un expert juriste. Nous avons utilisé Protégé⁷ 2000 version 4.1.0, pour l'édition de notre ontologie car c'est l'outil le plus puissant qui supporte la langue arabe. La première version de notre ontologie contient 30 concepts. La relation que nous avons utilisée est la relation de subsomption. Cette ontologie doit être enrichie dans nos prochains travaux afin d'intégrer tous les concepts du corpus.

5 Structuration des décisions basée sur l'ontologie juridique

Une fois tous les patrons des différentes parties des décisions élaborés, nous avons attribué un poids pour chaque terme d'un patron suivant son importance. Puis, nous avons cherché, pour chaque terme du patron, ses différents équivalents à partir de l'ontologie. Ainsi chaque patron sera enrichi par une liste de nouveaux termes. Une analyse morphologique sera par la suite réalisée pour dégager pour chaque terme, ses différentes formes morphologiques. L'identification de la présence d'un patron dans une décision se fera par le contrôle de l'existence des termes du patron au niveau de chaque phrase. Ainsi un score sera calculé pour chaque phrase

7. <http://protege.stanford.edu>

en prenant en compte les poids des termes. La phrase de la décision ayant le score le plus élevé par rapport à un patron donné sera ainsi la phrase candidate pour délimiter le début de la partie correspondante de la décision. Nous avons pu mettre en place un prototype qui met en oeuvre la méthode que nous avons suivie pour la structuration des décisions. Le prototype prend en entrée un document textuel contenant une décision et fournit en résultat un document XML bien structuré dont les balises sont les différentes parties d'une décision. Les résultats fournis par ce prototype ont été évalués et validés par les experts juristes.

6 Conclusion

Dans ce papier, nous avons présenté le contexte de nos travaux de recherche, qui est le développement d'un système de recherche de jurisprudence tunisienne en langue arabe. Nous avons ensuite décrit une première étape du processus de développement de ce système qui est la structuration des décisions fondée sur une ontologie juridique. Nous envisageons, dans nos travaux futurs, l'enrichissement de cette ontologie et son utilisation au niveau du processus de recherche des décisions.

Références

- Bourcier, D. (2006). Susciter la construction interdisciplinaire d'ontologies juridiques : bilan d'une expérience. In *SDC 2006*, Paris.
- Chaâben Kammoun, N., L. H. Belguith, et A. B. Hamadou (2001). The morph2 new version : A robust morphological analyzer for arabic texts. In *Actes des 10èmes journées internationales d'analyse statistique des données JADT'2010*, Rome, Italy.
- Cornu, G. (2005). *Linguistique juridique*. France : Montchrestien.
- Farzindar, A., G. Lapalme, et J. Desclés (2004). Résumé de textes juridiques par identification de leur structure thématique. In *Traitement automatique de la langue (TAL)*, Volume 45, pp. 39–64.
- Gruber, T. et R. Thomas (1993). A translation approach to portable ontology specifications. *Knowl. Acquis.* 5, 199–220.
- Mondary, T. (2011). *Construction d'ontologies à partir de textes. L'apport de l'analyse de concepts formels*. Thèse de doctorat, Université Paris-Nord - Paris XIII.

Summary

The computerized legal information is an evolving field. The present research sets out to develop a retrieval system of Tunisian jurisprudence in Arabic. The main objective of this system is to provide assistance to lawyers to solve a given legal situation by making available a collection of similar situations which will improve their future reasoning. In this paper, we shall focus on the first stage of the system in question, which is the modeling and structuring of case law decisions based on legal ontology.