

LEGTERM : PROJET D'UNE BANQUE DE DONNÉES MULTILINGUE DES TERMES JURIDIQUES DU DROIT NATIONAL

Jan RADIMSKÝ

Université de Bohême du Sud, České Budějovice

Abstract (En): Designing a bilingual or multilingual lexicographic tool that captures the equivalences between legal terms in different national laws and helps translators understand the intended meaning poses a complex challenge. Legal terminology is characterized by non-corresponding concepts, term polysemy, and imprecise meanings. Existing tools either focus on content for legal professionals or provide simple word-form equivalences for translators. The LegTerm project aims to bridge this gap by offering a simple and efficient terminological tool that combines the advantages of a dictionary and an encyclopedia, catering to different legal systems and languages. The article presents an overview of LegTerm, discusses terminological conventions and the relationships between concepts and terms in legal language, and describes the adopted solution. The conclusions, current project status, and prospects are also outlined.

Keywords (En): legal terminology; lexicography; law term; terminology database

Mots-clés (Fr) : terminologie juridique ; lexicographie ; terme juridique ; base de données terminologique

DOI : 10.32725/eer.2023.014

1. Introduction

Concevoir et réaliser un outil lexicographique bilingue (ou multilingue) qui rendrait compte des équivalences entre les termes juridiques de différents droits nationaux et qui, en même temps, permettrait aux traducteurs de saisir de manière simple et efficace le sens de ce qu'ils sont censés traduire s'avère une entreprise particulièrement délicate. Ceci est dû notamment à trois particularités bien connues de la terminologie juridique, à savoir la non-correspondance des concepts, la polysémie des termes et une certaine imprécision du sens de la terminologie juridique (PETRŮ, 2016). En effet, ancrés profondément dans l'ontologie conceptuelle du système juridique national, les termes juridiques renvoient à des concepts particuliers, non universels ; par conséquent, ils ont souvent des équivalents seulement imparfaits dans le système juridique cible et, parfois, le concept équivalent est totalement absent (*cf.* BOCQUET, 2008 : 13). La polysémie des termes juridiques est déterminée notamment par le fait que le vocabulaire juridique partage de nombreux mots-formes avec le langage courant (la « polysémie externe » aux termes de CORNU, 2005 : 69) et en même temps, de nombreux mots-formes sont polysémiques à l'intérieur du droit (CORNU, 2005 : 92). Finalement, il n'est pas rare que le sens des termes juridiques soit relativement flou : les concepts manquent de définitions explicites et leur extension se concrétise seulement dans le temps et dans l'usage.

Afin de mieux s'orienter dans cette complexité, les traducteurs des textes juridiques auraient besoin d'un outil lexicographique spécifique qui combinerait les avantages d'un dictionnaire (en donnant les équivalences entre mots-formes) avec ceux d'une encyclopédie (en permettant de comprendre le sens et de faire un choix terminologique raisonné). À notre meilleure connaissance, un outil semblable n'a pas encore été réalisé. Les outils existants visent à des fonctions ou à des domaines d'emploi plus spécifiques. Il s'agit notamment des dictionnaires encyclopédiques monolingues, concentrés sur la description du sens des concepts et destinés aux juristes (p. ex. *Max Planck Encyclopedia of Comparative Constitutional Law*), ou de dictionnaires et bases de données plurilingues concentrées uniquement sur les équivalences entre les formes et destinés aux traducteurs (p. ex. *United Nations Multilingual Legal Terminology Database*). Les rares projets lexicographiques plus complexes – ceux qui visent tant à donner l'équivalence formelle qu'informer sur le sens – se focalisent sur *un seul* système juridique utilisé dans un contexte multilingue, que ce soit au sein d'un pays (*Juridictionnaire*¹ au Canada) ou au sein d'une entité multinationale (IATE dans l'Union européenne). Parmi de rares exceptions figure le dictionnaire *JuriDico*², élaboré à l'Université de Montréal, qui contient environ 200 entrées (verbes en tant que termes juridiques) en anglais, en portugais et en français, avec les *frames* correspondants.

Le projet de la banque de données terminologique *LegTerm*, dont l'architecture est présentée dans cet article, se propose de combler cette lacune, en offrant tant aux traducteurs qu'aux professionnels du droit un outil terminologique simple et efficace qui permet de saisir les (non-)correspondances de forme et de sens entre les termes et concepts juridiques nationaux de pays différents dont le système juridique est également véhiculé par des langues différentes. La suite du présent article est structurée de manière suivante : le chapitre 2 donne un bref aperçu général de *LegTerm* ; le chapitre 3 présente les conventions terminologiques indispensables (3.1) et analyse la typologie des relations entre les concepts et les termes dans un domaine à concepts universels (3.2) aussi bien que dans un domaine à concepts particuliers (3.3) représenté précisément par le langage juridique ; le chapitre 4 décrit en détail la solution adoptée et ancrée dans l'architecture du *LegTerm* en procédant par les concepts (4.1), les termes (4.2) et les collocations (4.3) vers une vue d'ensemble (4.4). Les conclusions, l'état actuel du projet ainsi que les perspectives de son développement sont résumés dans le chapitre 5.

2. LegTerm : une vue d'ensemble

Développé progressivement à partir de l'année 2020 à l'Institut de langues romanes à l'Université de Bohême du Sud à České Budějovice (République

¹ Juridictionnaire : Le Juridictionnaire est un dictionnaire juridique bilingue complet disponible en anglais et en français. Il fournit des traductions et des définitions des termes juridiques utilisés dans le système juridique canadien.

² Cf. <http://olst.ling.umontreal.ca/?page_id=1357>.

tchèque), LegTerm³ est une base de données électronique de termes juridiques destinée aux traducteurs et aux professionnels du droit qui travaillent avec quatre langues et quatre systèmes juridiques différents, à savoir le tchèque (République tchèque), le français (France), l'espagnol (Espagne) et l'italien (Italie). Le contenu et l'architecture de la base de données LegTerm se concentrent spécifiquement sur les concepts de droit national, ce qui fait que, d'une part, sont omis les termes & concepts du droit international ou européen et, d'autre part, restent délibérément à l'écart également les concepts du droit national d'autres pays qui utilisent l'une des quatre langues en question, comme le Canada francophone, les pays hispanophones d'Amérique du Sud, etc. Autrement dit, le système d'équivalences entre les termes & concepts juridiques au sein du LegTerm ne s'instaure pas seulement entre différentes *langues*, mais en même temps entre différents *États* (i.e. leurs systèmes juridiques). Afin de rendre compte de ce phénomène, LegTerm introduit la notion de « langue-État » et opère avec quatre langues-États indépendants, tels le *français (France)*, etc. (cf. ci-haut, le nom de l'État est noté entre parenthèses après le nom de la langue). Ceci implique que si, dans le futur, LegTerm devait être complété par la terminologie juridique du Canada francophone par exemple, il serait indifférent que le français (en tant que langue) soit déjà présent dans LegTerm, car la nouvelle langue-État, notée *français (Canada)*, serait en principe indépendante de la langue-État *français (France)*.

Dans son architecture interne, LegTerm repose sur la conception onomasiologique (i.e. *knowledge – based* ou *concept-based*, au sens de L'HOMME, 2020 : 7-18). Autrement dit, elle est basée sur les *concepts* en tant qu'unités abstraites de sens juridique, ce qui la différencie des dictionnaires (juridiques) qui, dans une perspective sémasiologique, s'appuient sur les *mots* (simples ou complexes), *termes* ou *formes*. Au sein de LegTerm, toutes les informations liées à un *concept* sont regroupées sur une (et une seule) *fiche terminologique*. Ainsi, tous les *termes* (i.e. formes) désignant des concepts identiques ou proches au sein de chaque système juridique sont concentrés sur la même fiche terminologique, ce qui implique qu'ils peuvent être considérés comme *synonymes* au sein d'une langue-État et *équivalents* entre différentes langues-États.

L'architecture de LegTerm est donc en principe similaire à celle de la base de données terminologique canadienne *TERMIUM Plus*, mais elle comporte un certain nombre d'innovations importantes qui permettent de saisir de manière efficace les problèmes spécifiques liés à la traduction des termes de droit national, dont notamment celui du caractère non universel des concepts juridiques. En effet, une conception onomasiologique, comme celle de *TERMIUM Plus*, présuppose et requiert l'existence de concepts universels. Dans l'élaboration de LegTerm, nous avons dû affronter la question de savoir comment structurer une base de données onomasiologique plurilingue dans un domaine où l'universalité des concepts fait défaut. L'essentiel de la solution repose sur une structuration particulière des concepts, des termes et de leurs attributs majeurs en relation à une langue-État donnée. En effet, l'architecture de LegTerm prévoit un stockage structuré de

³ Cf. <<https://legterm.cz>>.

nombreuses informations concernant les termes et les concepts, tels les définitions, les collocations, les domaines d'emploi, différentes notes d'usage ou l'information sur l'acceptabilité des termes, dont l'organisation sera présentée dans la suite du présent texte.

3. Les concepts et les termes juridiques dans une perspective terminographique

3.1 Conventions terminologiques

L'analyse de l'organisation des données terminologiques requiert une définition rigoureuse des entités et termes de base et une observation scrupuleuse des conventions de notation dont les principes feront l'objet de la présente section. Nous avons déjà introduit plus haut la notion de *langue-État*, dont les valeurs notées *français (France)*, *italien (Italie)*, etc. représentent les combinaisons indissociables de langues et de systèmes juridiques, entre lesquels les équivalences formelles et conceptuelles s'établissent. Il reste à définir deux autres entités-clés, celle de *concept* et de *terme*, ainsi que la manière dont nous allons nous y référer pour les décrire.

Conformément à la tradition du paradigme de connaissance (*Knowledge paradigm*, cf. L'HOMME, 2020 : 5-18), ancrée dans la théorie générale de terminologie de Wüster, nous entendons par *concept* une unité atomique de sens technique. Cette vue est en principe compatible avec celle donnée par la norme ISO 1087:2019 qui définit le concept comme une « unité de connaissance créée par une combinaison unique de caractéristiques » (*unit of knowledge created by a unique combination of characteristics*), sous réserve du fait qu'il serait illusoire de postuler la possibilité d'énumérer explicitement toutes les caractéristiques du concept juridique au sein d'une définition qui en délimiterait les contours en termes de conditions nécessaires et suffisantes. Elle est également compatible avec une vue mentaliste qui, dans la lignée de la pensée saussurienne, appréhende le concept comme un ensemble abstrait de représentations mentales inconscientes lié à une ou plusieurs formes phonétiques et/ou graphiques (JACKENDOFF, 2012). Ces perspectives convergent vers l'observation que le concept, en tant qu'unité de sens, ne peut jamais être appréhendé, saisi ou identifié d'une manière directe. Dans une base de données comme LegTerm, il est représenté techniquement par un code abstrait unique qui le distingue des autres concepts. Comme ce type de représentation serait peu pratique dans un texte scientifique, ou plus généralement dans toute communication entre humains, les concepts seront désignés dans la suite du présent article par une forme linguistique (i.e. par un terme) écrite en majuscules, comme p. ex. : ACIDE_SULFURIQUE ou DROIT.

Un *terme* (ou *désignation*) est conçu au sein de LegTerm comme une forme linguistique destinée à la représentation verbale d'un concept, ce qui correspond à la définition de la *désignation* donnée par la norme ISO 1087:2019 (*représentation d'un concept par un signe qui le dénote dans un domaine ou sujet*). Comme l'architecture de LegTerm ne prévoit pas une distinction entre les concepts généraux et individuels, elle n'opère pas non plus avec la distinction entre les différents types

de désignations prévus par ISO 1087:2019 (comme *terme*, *appellation*, *nom propre*, *symbole*, etc.) : un *terme* dans LegTerm peut désigner un concept quelconque, qu'il soit général ou individuel. D'un point de vue formel, un *terme* peut revêtir différentes formes linguistiques ayant un degré de complexité différent, tels un mot simple ou construit (« vitriole »), un composé syntagmatique (« acide sulfurique »), une abréviation ou un symbole (« H_2SO_4 »), etc.⁴ D'un point de vue sémantique, il est important de souligner qu'un *terme* fait toujours référence à un et un seul concept au sein de LegTerm. Pour donner un exemple concret, ce principe d'homonymie radicale implique que « droit₁ » (en tant qu'*ensemble de règles de conduite socialement édictées et sanctionnées* – comme dans *droit civil*, *droit pénal*, etc.) et « droit₂ » (*prérogative individuelle qui existe sur la tête d'une personne* – comme dans *droit de vote*, etc.) représentent deux termes différents, car rattachés à deux concepts différents. La manière dont ces données sont organisées au sein de LegTerm est illustrée dans le tableau 1. Dans la suite de cet article, les termes seront notés entre guillemets français (« droit »), sans l'utilisation des index (i.e. sans spécifier « droit₁ » vs. « droit₂ »).

ID du concept	Concept (label)	Terme	Définition
2 ⁵	DROIT	« droit »	Ensemble de règles de conduite socialement édictées et sanctionnées, qui s'imposent aux membres de la société. Ex. : <i>droit civil</i> , <i>droit pénal</i>
8 ⁶	DROIT_SUBJECTIF	« droit » « droit subjectif » « prérogative »	[...] prérogative individuelle qui existe sur la tête d'une personne, titulaire du droit, par opp. à Droit objectif Ex. : <i>droit de vote</i>

Tableau 1 : Concepts DROIT et DROIT_SUBJECTIF avec les termes correspondants

L'exemple donné dans le tableau 1 montre que le terme « droit » rattaché au concept DROIT_SUBJECTIF est synonyme des termes « droit subjectif » et « prérogative ». En revanche, ceci n'est pas le cas du terme « droit » rattaché au concept DROIT.

Pour ceux qui, notamment parmi les informaticiens, souhaiteraient mieux comprendre le système interne de l'identification des concepts, il convient de donner quelques précisions relatives aux deux premières colonnes du tableau. L'identifiant unique de chaque concept est un code arbitraire qui a deux, voire trois versions matérielles. La première correspond à un numéro ID court et simple, donné dans la colonne de gauche (2 et 8 dans notre exemple), et permet une identification rapide, simple et univoque du concept dans la communication humaine entre les personnes qui préparent les données. La seconde version du numéro ID, plus étendue (p. ex. 2b7d3c60-a27b-4940-98dc-b60799b26626), interne à la base des

⁴ P. ex. *tétraoxosulfate de dihydrogène*, ou des désignations archaïques comme *acide anglais*, *acide des chambres*, *acide méta-sulfurique*, *acide vitriolique*, *esprit acide*, *huile glaciale*.

⁵ Cf. <<https://legterm.cz/concept/detail/2b7d3c60-a27b-4940-98dc-b60799b26626>>.

⁶ Cf. <<https://legterm.cz/concept/detail/66fac967-73fe-4280-8088-70827eb1d41b>>.

données et plus difficilement saisissable par l'humain, sert de clé primaire (*primary key*) à l'intérieur du système et permet de créer un lien HTML qui renvoie directement à la fiche terminologique complète du concept en question, comme il est indiqué dans les notes en bas de page. Les deux versions de l'identifiant sont publiquement accessibles et visibles à tous les utilisateurs de LegTerm. La troisième version de l'identifiant du concept, schématisée en abrégé dans la colonne n° 2 est, en revanche, plus particulière. Dénommé *label* au sein de LegTerm, cet identifiant se compose d'une désignation écrite en majuscules, suivie d'une brève définition (en tchèque) comme dans (1).

- (1) PRÁVO : soubor právních norem
DROIT : ensemble de règles juridiques

Le label permet de saisir l'identité d'un concept dans la communication entre terminologues humains, il n'est accessible qu'aux utilisateurs enregistrés qui préparent les données pour les saisir dans LegTerm. Créé obligatoirement au moment où le concept en question est introduit dans le système, il ne peut plus jamais être modifié ou effacé, ce qui permet de préserver l'identité de chaque concept depuis son introduction dans LegTerm sous une forme appréhendable par le terminologue humain.

3.2 Relation concept – terme(s) : le cas des concepts universels

Après avoir défini les entités-clés *concept* et *terme*, nous nous proposons d'étudier les types des relations qui s'instaurent entre elles au niveau théorique ainsi que l'impact de cette situation sur l'organisation interne de la base des données terminologique. Considérons d'abord le cas plus simple qui implique les concepts à caractère universel, donc indépendants de la langue, du pays, ou de la culture, illustré par l'exemple dans le schéma 1. Comme suggéré par la nature de l'exemple, les concepts universels sont relativement courants dans les sciences naturelles, mais bien moins courants dans les sciences humaines et sociales.

Au sein d'une langue, la relation entre un concept (ici : ACIDE_SULFURIQUE) et un terme (ici p. ex. : « acide sulfurique ») devrait être biunivoque, c'est-à-dire du type 1 : 1, du moins selon le souhait des terminologues (*cf.* TEMMERMAN, 2000 : 10). Néanmoins, ceci est très rarement le cas. En réalité, un concept peut être désigné par plusieurs termes comme « acide sulfurique », « tétraoxosulfate de dihydrogène », « vitriol », « H₂SO₄ », ce qui fait que nous avons affaire à une relation du type 1 : N. Du fait d'être rattachés au même concept, ces termes peuvent être définis comme *synonymes* au sein d'une langue, bien évidemment sous réserve que leur interchangeabilité effective dans un contexte concret soit limitée et conditionnée par d'autres facteurs dont il ne sera pas question ici.

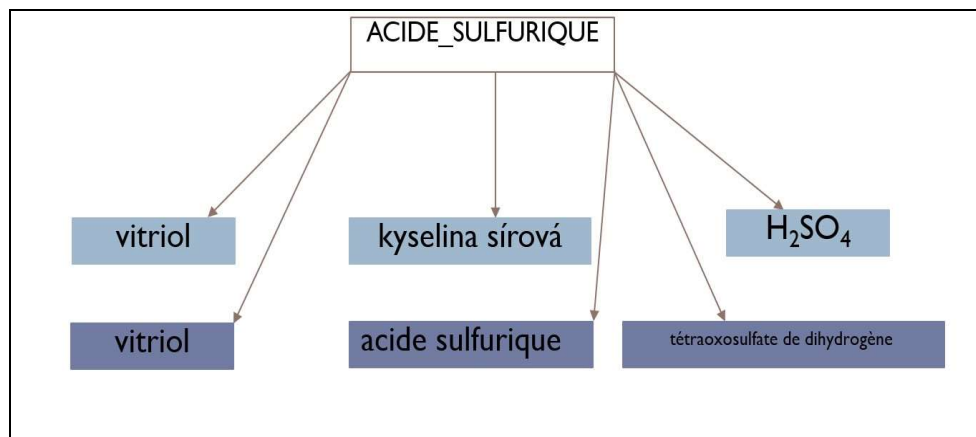


Schéma 1 : Relation concept – terme(s) dans l'hypothèse des concepts universels
(termes français et tchèques)

Transposée vers une situation plurilingue – avec le français et le tchèque en l'occurrence –, la situation où 1 concept universel est rattaché à N termes dans chaque langue implique une relation du type N : N entre les (au moins 3) termes tchèques et les (au moins 3) termes français. Du fait d'être rattachés au même concept, les termes tchèques et français dans le schéma 1 peuvent être considérés comme *équivalents*, toujours sous réserve que leur interchangeabilité effective dans un contexte précis ne soit pas sans conditions supplémentaires.

Les relations illustrées par le schéma 1 impliquent une organisation relativement simple d'une base de données terminologique à caractère onomasiologique. En effet, toutes les entités rattachées au même concept sont réunies sur une seule *fiche terminologique*. Cette fiche peut comporter une définition du concept en question (dans une langue ou en plusieurs versions linguistiques), ainsi que les N termes susceptibles de désigner ce concept dans chaque langue. C'est précisément le type d'architecture adopté par la banque *TERMIUM Plus* qui, autour d'entrées universelles, réunit les données terminologiques pour l'anglais, le français et l'espagnol (cf. la fiche ACIDE_SULFURIQUE⁷ dans *TERMIUM Plus*).

3.3 Relation concept – terme(s) : le cas des concepts particuliers

Dans le domaine du droit national, nous devons faire face une situation tout à fait différente. En effet, les concepts juridiques du droit national sont étroitement liés à l'ordre juridique en question, il s'agit donc de concepts particuliers, non universels. Comme il est démontré par des exemples dans le schéma 2, certains concepts sont similaires dans une certaine mesure, comme MARIAGE_CIVIL et MANŽELSTVÍ dans les ordres juridiques français et tchèque respectivement, sans

⁷ Cf. <https://www.btb.termiumpplus.gc.ca/tpv2alpha/alpha-eng.html?lang=eng&i=1&srchtxt=acide+sulfurique&index=alt&codom2nd_wet=1>.

se chevaucher pour autant⁸. Qui plus est, il y a même des concepts pour lesquels un homologue dans l'autre ordre juridique fait défaut : le concept français de MAGISTRAT ne trouve pas d'homologue dans l'ordre juridique tchèque.

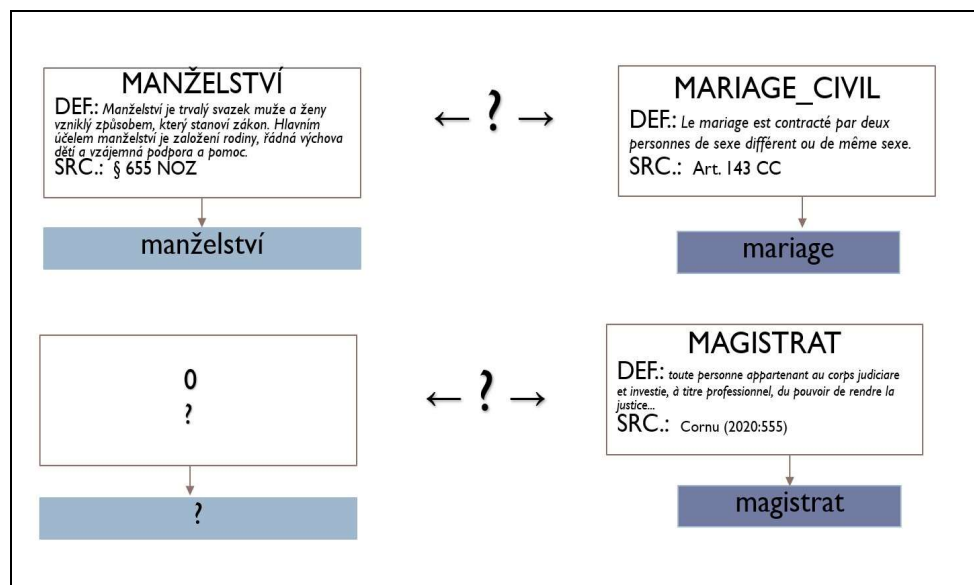


Schéma 2 : Relation concepts – terme(s) dans l'hypothèse des concepts particuliers
(termes français et tchèques)

L'existence des concepts particuliers pour l'ordre juridique en question a de sérieuses implications théoriques, notamment en ce qu'elle rend impossible l'équivalence entre les termes. En effet, nous avons défini plus haut les relations de synonymie (au sein d'une langue) et d'équivalence (entre deux langues) en tant que relations entre deux ou plusieurs termes qui désignent le même concept. Dans un domaine à concepts particuliers, cette notion d'équivalence ne pourra pas être maintenue telle quelle : il faudra considérer que le degré de l'équivalence des termes est fonction de la similitude des concepts. Avant de postuler l'équivalence, il est donc nécessaire d'établir la typologie des différentes relations qui peuvent s'instaurer entre les concepts dans les différents systèmes juridiques. En même temps, on conçoit que la traduction met en jeu non seulement deux langues différentes, mais deux systèmes juridiques différents, ce qui justifie la notion de *langue-État* définie plus haut.

Pour établir une typologie des relations qui peuvent s'instaurer entre les concepts juridiques de deux (ou plusieurs) langues-États nous nous appuyons sur une réanalyse de la typologie des équivalences entre les termes juridiques français et tchèques exposée dans RADIMSKÝ (2004). Considérant que pour être efficace et applicable dans la pratique, cette typologie doit être avant tout simple, nous distinguons trois types de relations entre les concepts provenant de langues-États

⁸ À l'heure actuelle, le mariage peut être contracté par des personnes du même sexe dans l'ordre juridique français, ce qui n'est pas le cas dans l'ordre juridique tchèque.

différents, en fonction de si, dans le système juridique cible, le concept en question est : compatible (2a), inexistant (2b) ou obsolète (2c). La typologie exposée dans le schéma 2 est donc complétée par le cas du concept obsolète (2c).

(2a) **concept compatible** dans la langue-État cible

Le concept existe dans le système juridique cible, mais son extension peut être différente

Ex. : GOUVERNEMENT, MINEUR, MARIAGE_CIVIL

(2b) **concept inexistant** dans la langue-État cible

Le concept n'existe pas dans le système juridique cible ou l'extension du concept le plus proche diffère trop pour que l'on puisse établir une correspondance

Ex. : MAGISTRAT, SOCIÉTÉ_PAR_ACTIONS_SIMPLIFIÉE

(2c) **concept obsolète** dans la langue-État cible

Le concept n'existe plus dans le système juridique cible du fait des changements législatifs

Ex. : FILIATION_NATURELLE, DÉLAI_DE_VIDUITÉ

Le cas (2a), commenté plus haut, est central pour le langage juridique. Dans la pratique il peut présenter deux situations différentes, mais proches. D'une part, les systèmes juridiques tchèque et français connaissent les concepts GOUVERNEMENT ou MINEUR, dont les définitions ne diffèrent pas de manière significative. Leur extension est donc (pratiquement) identique, ce qui fait qu'ils pourraient être considérés techniquement comme *concepts universels*. D'autre part, les deux systèmes juridiques connaissent le concept de MARIAGE_CIVIL, mais leur extension diffère : à l'heure actuelle, le mariage civil peut être contracté par des personnes du même sexe dans l'ordre juridique français, ce qui n'est pas le cas dans l'ordre juridique tchèque. Il s'agit donc de *concepts particuliers*, mais la différence d'extension, relativement mineure, n'est pas de nature à établir une non-correspondance absolue entre les deux concepts. Du point de vue terminographique, il n'est pas nécessaire d'effectuer une distinction entre les concepts qui fonctionnent comme GOUVERNEMENT d'une part et ceux comme MARIAGE_CIVIL de l'autre. Techniquement parlant, il s'agit de concepts particuliers dont l'extension est plus ou moins proche, ce qui permet de les identifier comme *concepts correspondants*.

Le cas du concept inexistant (2b) recouvre toutes les situations où pour un concept juridique de la langue-État de départ il n'est pas possible de trouver un concept correspondant dans la langue-État cible. Par exemple le système juridique français connaît le concept de MAGISTRAT (personne exerçant les fonctions de

juge ou de procureur)⁹ ou de SOCIÉTÉ_PAR_ACTIONS_SIMPLIFIÉE¹⁰, mais l'ordre juridique tchèque ne connaît pas ces concepts.

Le cas d'un concept obsolète (2c) peut être considéré comme une particularité du langage juridique. Comme la loi est sujette aux changements diachroniques, un concept juridique peut devenir désuet à la suite d'une décision du législateur. Ainsi, les concepts comme FILIATION_NATURELLE¹¹ ou DÉLAI_DE_VIDUITÉ¹² ne font plus partie du système juridique français, alors que c'était le cas dans le passé. À ce point il convient de souligner que la notion de *concept obsolète* (p. ex. DÉLAI_DE_VIDUITÉ) diffère en principe de la notion de *terme obsolète*. (p. ex. « vitriol » pour désigner le concept ACIDE_SULFURIQUE). Un terme obsolète (comme « vitriol ») peut désigner un concept qui n'est pas obsolète (ACIDE_SULFURIQUE). En revanche, si le concept est effectivement obsolète (comme DÉLAI_DE_VIDUITÉ), les termes qui le désignent (« délai de viduité ») seront nécessairement obsolètes eux aussi. D'un point de vue terminographique la notion de *concept obsolète* paraît importante dans le domaine du langage juridique, car la base de données terminologique devrait fournir les informations sur les changements diachroniques dans le système conceptuel et terminologique.

Dans le chapitre suivant, nous allons montrer comment cette typologie des relations qui peuvent s'instaurer entre les concepts juridiques provenant de plusieurs langues-États différentes se traduit dans l'architecture de la base de données terminologiques.

4. Structuration interne de LegTerm

4.1 Concepts

Dans un système à concepts particuliers, il serait naturel de postuler que la base de données terminologique devrait contenir une liste de concepts spécifiques pour chaque langue-État. Par la suite, il faudrait déterminer le type de relation qui s'instaure entre les différents concepts des différentes langues-États, comme dans l'exemple donné dans le schéma 3. En fonction de la typologie énoncée dans la section 3.3, la décision concernant le type de relation entre concepts serait binaire : soit le concept compatible dans la langue-État cible existe et la relation peut être établie – ce qui est le cas du *concept compatible* (2a) ou celui du *concept obsolète*

⁹ « Au sens strict, toute personne appartenant au corps judiciaire et investie, à titre professionnel, du pouvoir de rendre la justice (magistrat du siège) ou de la requérir au nom de l'État (magistrat du parquet). » CORNU *et al.* (2020 : 555).

¹⁰ « Société commerciale ne faisant pas publiquement appel à l'épargne constituée entre au moins deux personnes morales dotées d'un capital important [...]. » CORNU *et al.* (2020 : 873).

¹¹ La filiation naturelle est le rapport qui unit par un lien de parenté une personne (un enfant) à ceux dont elle est issue (ses parents), dans la situation où les parents ne sont pas mariés, mais pourraient l'être légalement. (Cf. DHÉNIN & LAMADON, 1995 : 213).

¹² « Délai d'attente (de 300 jours, sauf abrégement judiciaire) destiné à éviter la confusion de paternité [...] que la veuve, et par extension la femme divorcée, doit laisser s'écouler avant de contracter un nouveau mariage. » CORNU *et al.* (2020 : 279).

(2c) – soit il n'existe pas et la relation ne peut pas être établie (ce qui le cas du *concept inexistant* en (2b)).

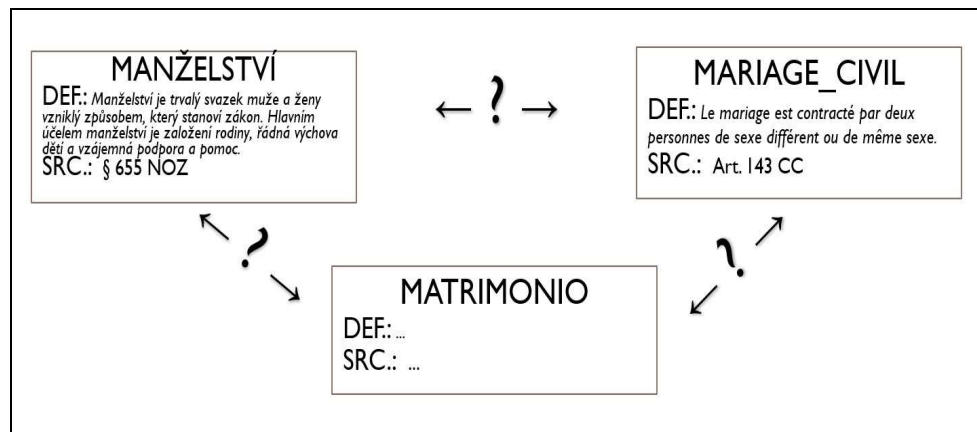


Schéma 3 : Relations entre les concepts particuliers de trois langues-États

Bien qu'intuitivement logique, cette solution s'avère inutilement compliquée dans la pratique, notamment du fait qu'en présence de 3, voire 4 langues-États, le système devient trop complexe. En effet, avec 3 langues-États, nous obtenons 3 tables de concepts liées par 3 relations binaires, avec 4 langues-États, les 4 tables de concepts doivent être liées par 6 relations binaires, etc. De plus, pour les concepts non existants au sein de certaines langues-États il faudrait postuler un équivalent fictif pour qu'il soit possible d'y attribuer un terme équivalent (en effet, même si le concept n'existe pas dans le système juridique cible, il faut trouver une solution pour traduire le terme correspondant, donc un concept auquel ce terme se rattacherait).

La solution adoptée dans l'architecture de LegTerm est donc différente et plus simple. Elle prévoit une seule liste de concepts juridiques. Celle-ci est basée en premier lieu sur le système juridique tchèque, mais si un concept y fait défaut et qu'il est nécessaire de le faire figurer, il peut être pris d'un autre système juridique. Autrement dit, il s'agit en principe d'une liste de concepts *universels*, bien que pris, dans la pratique, des différents systèmes juridiques qui font objet d'un traitement au sein de la base de données. Afin de savoir si le concept fait effectivement partie du système juridique en question, il faut que chaque concept prenne pour chaque langue-État l'un des *statuts* suivants : *valable*, *non valable* ou *obsolète*. Ces valeurs traduisent directement la typologie des équivalences introduite dans 3.3, mais les relations ne s'établissent pas entre chaque paire de langues-États, mais entre l'item de la liste des concepts universels et la situation particulière d'une langue-État. Ainsi, le statut *valable* signale qu'il y a *compatibilité* du concept « universel » (celui dans la liste des concepts de LegTerm) avec celui du système juridique en question ; le statut *non valable* traduit l'*inexistence* du concept dans le système juridique en question ; et le statut *obsolète* est réservé aux concepts obsolètes.

À titre d'exemple les tableaux 2-4 montrent trois concepts différents et leurs statuts dans les différentes langues-États.

Concept	Langue-État	Statut du concept	Définition
MARIAGE_CIVIL	tchèque (CS)	valable	[définition CS]
	français (France)	valable	[définition FR]
	italien (Italie)	valable	[définition IT]

Tableau 2 : Le concept MARIAGE_CIVIL et ses statuts dans LegTerm

Concept	Langue-État	Statut du concept	Définition
MAGISTRAT	tchèque (CS)	non valable	[---]
	français (France)	valable	[définition FR]
	italien (Italie)	valable	[définition IT]

Tableau 3 : Le concept MAGISTRAT et ses statuts dans LegTerm

Concept	Langue-État	Statut du concept	Définition
DÉLAI_DE_VIDUITÉ	tchèque (CS)	obsolète	[définition CS]
	français (France)	obsolète	[définition FR]
	italien (Italie)	obsolète	[définition IT]

Tableau 4 : Le concept DÉLAI_DE_VIDUITÉ et ses statuts dans LegTerm

Le tableau 2 montre que les trois langues-États (i.e. systèmes juridiques) ont un concept compatible avec le concept universel de MARIAGE_CIVIL. Les termes rattachés à ce concept peuvent donc être considérés comme synonymes (au sein d'une langue-État) ou équivalents (entre les différentes langues-États). Autrement dit, l'architecture en question permet de maintenir la notion de *synonymie* et *équivalence* telle que définie dans la section 3.1 – et telle qu'elle fonctionne au sein d'un système à concepts universels, comme p. ex. *TERMIUM Plus*. Mais en même temps, cette architecture n'oblige pas de postuler UNE définition universelle pour les concepts, ce qui ne serait pas souhaitable. Au contraire, chaque concept peut bénéficier d'une définition particulière à une langue-État, comme il est signalé dans la dernière colonne du tableau 2. La confrontation de ces définitions permettra de faire ressortir les différences d'extension du concept en question entre les différents systèmes juridiques.

L'exemple du tableau 3 montre la situation d'un concept dit *inexistant*. En effet, le concept de MAGISTRAT est valable (donc, existant) au sein des systèmes juridiques français et italien, mais fait défaut dans le système juridique tchèque. De ce fait, le statut de ce concept universel pour la langue-État tchèque (CS) est *non valable*, ce qui implique également qu'il ne pourra pas avoir une définition au sein de la langue-État *tchèque (CS)*.

L'exemple dans le tableau 4 montre un concept qui est désormais obsolète dans les trois systèmes juridiques en question.

En principe, rien n'empêche qu'on puisse avoir un concept obsolète dans une langue-État, et valable ou inexistant dans d'autres langues-États. Toutes les combinaisons de statuts pour un concept entre les différentes langues-États sont envisageables.

4.2 Acceptabilité des termes

Nous avons vu dans la section 3.1. que les termes se rattachent aux concepts par une relation du type 1 : N, ce qui veut dire qu'un concept peut être désigné par plusieurs termes. Bien que synonymes, les termes référant au même concept ne sont pas pour autant interchangeables. Afin de nuancer le choix terminologique, la norme ISO 1087:2019 (sous 3.4.19-22) introduit pour chaque terme l'attribut *acceptabilité* avec les valeurs *privilegié*, *toléré*, *désuet* et *déconseillé* (en anglais *preferred*, *admitted*, *obsolete*, *deprecated*). LegTerm reprend ce système, comme le montre l'exemple des termes associés au concept ASSOCIATION dans le tableau 5.

Concept	Langue-État	Statut du concept	Termes et leur acceptabilité
ASSOCIATION	tchèque (CS)	valable	« zapsaný spolek » (preferred) « spolek » (admitted) « občanské sdružení » (obsolete)
	français (France)	valable	« association » (preferred) « association loi de 1901 » (admitted) « association à but non lucratif » (admitted)

Tableau 5 : Le concept ASSOCIATION avec les termes tchèques et français correspondants

En principe, tous les termes considérés comme corrects dans le système juridique en question sont notés comme *tolérés* et l'un deux – celui qui est utilisé directement par le législateur ou, à défaut, celui qui est le plus explicite, est noté comme *privilegié*. Autrement dit, les termes *tolérés* et *privilegiés* devraient être interchangeables, le choix est laissé au traducteur en fonction du contexte concret. L'usage des termes marqués comme désuets dans la traduction doit en revanche être justifié par des raisons particulières. Le tableau 5 montre que le terme tchèque « občanské sdružení » est une ancienne version du terme « zapsaný spolek » (introduit par la réforme du Code civil tchèque en 2012), mais les concepts juridiques auxquels ils renvoient sont compatibles – ce qui motive le fait que ces termes soient rattachés au même concept universel ASSOCIATION. Dans une traduction allant du français au tchèque, le terme « občanské sdružení » va représenter l'équivalent correct des trois termes français uniquement dans un texte juridique antérieur à la réforme du Code civil tchèque. L'étiquette *déconseillé* est réservée aux termes incorrects mais néanmoins utilisés pour un concept donné, par exemple, par le grand public, dans le journalisme, dans le langage substandard, etc.

Les valeurs d'acceptabilité des termes, telles que définies par ISO 1087:2019, sont applicables dans le cas des concepts compatibles (tableau 5). Toutefois, si un concept est *obsolète* ou *inexistant* dans le système juridique cible, les mêmes critères d'acceptabilité ne peuvent pas être appliqués, comme le montrent les exemples dans le tableau 6.

Concept	Langue-État	Statut du concept	Termes et leur acceptabilité
SAS	tchèque (CS)	non valable	« zjednodušená akciová společnost » (foreign concept denomination)
	français (France)	valable	« société par actions simplifiée » (preferred)
DÉLAI_DE_VIDUITÉ	tchèque (CS)	obsolète	« vdovský rok » (obsolete)
	français (France)	obsolète	« délai de viduité » (obsolete)

Tableau 6 : Les concepts SOCIÉTÉ_PAR_ACTIONS_SIMPLIFIÉE (SAS) et DÉLAI_DE_VIDUITÉ avec les termes tchèques et français correspondants

Dans le cas de concept obsolète (p. ex. : DÉLAI_DE_VIDUITÉ), le statut des termes correspondants sera par définition désuet, car il est difficile d'imaginer qu'un « sens obsolète » puisse être véhiculé par un terme qui, lui, ne soit pas désuet. Le choix des étiquettes possibles est donc réduit à une seule valeur. Un concept *inexistant* dans un système juridique, i.e. doté du statut *non valable* pour une langue-État spécifique (p. ex. SAS en tchèque (CS)), doit être désigné par un terme artificiel, forgé pour véhiculer un contenu en principe étranger au système juridique local. De ce fait, aucune des étiquettes prévues par la norme ISO 1087:2019 ne paraît convenable. Au sein de LegTerm, nous avons prévu pour ce type de terme une nouvelle étiquette *dénomination d'un concept étranger* (*foreign concept denomination*). Elle indique que le terme en question ne désigne aucun concept dans le droit national donné, mais désigne un concept provenant d'un système juridique étranger.

L'interdépendance des attributs *état du concept* et *acceptabilité du terme* est schématisée dans le tableau 7.

Statut du concept (dans la langue-État en question)	Acceptabilité du terme (valeurs possibles)
valable	préféré (preferred) toléré (admitted) désuet (obsolete) déconseillé (deprecated)
non valable	dénomination d'un concept étranger (foreign concept denomination)
obsolète	désuet (obsolete)

Tableau 7 : Interdépendance des attributs *état du concept* et *acceptabilité du terme* au sein de LegTerm

L'information véhiculée par la paire des étiquettes *état du concept* et *acceptabilité du terme* est riche et simple en même temps. Elle permet au traducteur de savoir non seulement si – et dans quelle mesure – le terme est correct, mais aussi s'il désigne une réalité interne ou externe au système juridique en question, éventuellement en prenant en compte la dimension diachronique tant sur l'axe formel que sur l'axe conceptuel.

4.3 Collocations

Dans les textes juridiques, les termes apparaissent parfois dans des tournures figées¹³ qui peuvent prendre différentes formes (verbo-nominale, nominale, nominale-adjectivale, etc.), mais partagent deux propriétés essentielles : d'une part, elles ont une forme figée, constante, fréquente dans les textes ; d'autre part, cette forme n'est pas prévisible (i.e. le même contenu pourrait être exprimé théoriquement par plusieurs alternatives dont seulement certaines sont communes et, par conséquent, correctes). Ces tournures figées sont enregistrées dans LegTerm en tant que *collocations*, chaque collocation étant rattachée à un terme. Les collocations du terme tchèque « manželství » (fr. « mariage ») sont par exemple : *vstoupit do manželství*, *rozpad manželství*, *po dobu trvání manželství*, *šťastné manželství*, etc., leurs équivalents français seraient : *entrer en mariage*, *rupture du mariage*, *pendant le mariage*, *mariage heureux*.

Dans LegTerm, un terme peut être lié à N collocations au sein d'une langue-État. En même temps, il est possible de lier manuellement les collocations équivalentes entre différentes langues-États, si elles sont rattachées aux termes relevant du même concept. Si l'équivalence de collocations respectives est établie entre deux paires de langues-États, l'équivalence entre la troisième paire est automatique. Par exemple, lorsque la collocation espagnole *derecho a contraer matrimonio* est liée à la collocation française *droit au mariage* et, en même temps, aux collocations tchèques *právo na uzavření manželství* et *právo uzavřít manželství*, la collocation française devient automatiquement liée aux deux collocations tchèques.

Il faut toutefois signaler que la distinction entre *termes* et *collocations* peut s'avérer précaire dans la pratique du fait que de nombreux concepts juridiques sont désignés par des termes composés qui revêtent la forme de syntagmes nominaux. À l'heure actuelle, l'enregistrement des collocations dans LegTerm se trouve encore dans une phase expérimentale. Par principe, lorsqu'une forme composée est susceptible de renvoyer à un concept juridique spécifique qui peut être défini, elle est de préférence enregistrée dans LegTerm en tant que *terme*¹⁴.

¹³ Appelées aussi *patrons linguistiques* (« jazykové šablony ») par le jurilinguiste Michal TOMÁŠEK (2003 : 52).

¹⁴ La priorité que nous accordons à la relation concept-terme est en ce sens absolue. Par conséquent, LegTerm n'adhère pas à la définition suivante qui conçoit les *collocations* comme combinaisons syntagmatiques de plusieurs termes : « Terms – even long noun-adjective strings – designate individual concepts, whereas technical collocations identify situations or propositions that are for the most part made up of several mutually related concepts. The purpose of both elements is the same: to firmly delineate the conceptual and linguistic space occupied by strictly defined individual or closely associated concepts and units of knowledge » (WRIGHT & BUDIN, 1997 : 15-16). En d'autres termes, une unité complexe comme « droit de succession » sera classifiée comme *terme* du fait qu'elle renvoie à une réalité conceptuelle dans le système juridique (i.e. au concept DROIT_DE_SUCCESSION, définissable et ayant une place dans le système conceptuel des droits), indépendamment du fait qu'elle soit construite de manière compositionnelle à partir de termes atomiques « droit » et « succession » qui renvoient à leur tour aux concepts atomiques respectifs.

4.4 Vue d'ensemble

Même si les concepts (4.1), les termes (4.2) et les collocations (4.3) représentent la partie la plus visible de LegTerm, cette base de données contient de nombreuses autres entités reliées entre elles de manière à fournir à l'utilisateur une information en même temps complète et simple à saisir. Une vue d'ensemble de l'architecture interne de LegTerm est donnée dans le schéma 4.

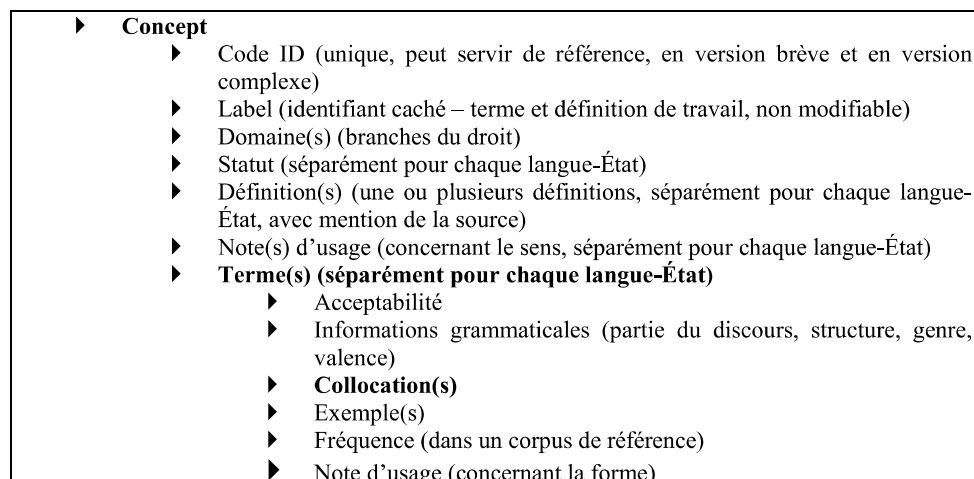


Schéma 4 : Structuration des entités dans LegTerm

Conformément à la conception onomasiologique, l'unité de base est un *concept* qui a un identifiant unique et non modifiable en trois versions matérielles (deux *codes ID* et un *label* – voir le tableau 1 avec le commentaire correspondant).

Un concept peut appartenir à un ou plusieurs *domaines* qui représentent des branches du droit, comme *la théorie du droit*, *le droit pénal*, *le droit constitutionnel*, etc., sélectionnés de la liste des domaines¹⁵ qui est universelle, c'est-à-dire unique pour toutes les langues-États. Comme d'usage dans les ressources terminologiques, la mention du *domaine* (ou *domaine d'emploi*) facilite avant tout la désambiguïsation des termes homonymes rattachés à des concepts différents, son identification permet également de définir le champ d'application des relations conceptuelles (*cf.* RONDEAU, 1984 : 85).

À chaque concept il faut attribuer un *statut* pour chaque langue-État – ses valeurs possibles étant : *valable*, *non valables* ou *obsolète* – qui permet de saisir le rôle du concept dans le système juridique en question (*cf.* la section 4.2.).

Ensuite, chaque concept peut avoir une ou plusieurs *définitions* qui sont spécifiques pour chaque langue-État. Les définitions sont accompagnées d'une mention de la source. LegTerm fait appel aux trois principaux types de définitions juridiques, à savoir aux définitions légales, aux définitions jurisprudentielles et aux

¹⁵ La plupart des concepts appartiennent à un domaine, mais les concepts très généraux tels que DROIT, LOI, JUGE peuvent avoir plusieurs domaines.

définitions dites « doctrinales » ou « scientifiques » (PETRŮ, 2015 : 129) qui apportent des perspectives variées et complémentaires afin de saisir le sens d'un concept. Dans la pratique, contrairement à ce qu'on serait amené à croire, peu de concepts juridiques trouvent une définition explicite dans la législation (définition légale) ; les manuels universitaires de droit s'avèrent en revanche une source riche et informative, tant pour l'identification de concepts que comme source de définitions de qualité. La confrontation des définitions permet de rendre compte des différences d'extension du concept dans les différentes langues-États, ces différences peuvent également être explicitement décrites dans les *notes d'usage concernant le sens*. Toutes ces informations permettent de saisir, d'identifier et de comprendre le *concept* en tant qu'unité de sens juridique.

Un concept est associé à un ou plusieurs *termes* dans chaque langue-État. Les termes portent une marque d'*acceptabilité* (préféré, toléré, déconseillé, désuet, dénomination d'un concept étranger) et les informations grammaticales de base. À chaque terme il est possible d'associer une ou plusieurs *collocations*, *exemples*, *notes d'usage (concernant la forme)* et informations sur la *fréquence relative* dans des corpus spécialisés. Autrement dit, il s'agit d'informations qui permettent de comprendre l'usage du *terme* en tant que forme linguistique.

5. Conclusions et état du projet

Lancée en 2021, la base de données LegTerm est librement accessible en ligne pour consultation à l'adresse <<https://legterm.cz/>>. L'interface informatique, réalisée par la société Neatous, permet aux utilisateurs enregistrés d'effectuer l'ajout, la modification et la mise à jour manuelle des données. En ce qui concerne le contenu (les données terminographiques), il s'agit d'un outil expérimental encore en germe. LegTerm contient actuellement quelque 700 concepts juridiques avec les termes et les définitions correspondants pour la langue-État tchèque (CS), mis à jour progressivement par Ivo Petrů. Depuis 2021, l'équipe composée d'enseignants, doctorants et étudiants de l'Institut de langues romanes à l'Université de Bohême du Sud¹⁶ prépare et teste les procédures de saisie des données terminologiques pour les autres langues-États et, en même temps, vise à vérifier empiriquement l'efficacité de l'architecture de LegTerm.

D'autres travaux en vue du développement du projet sont prévus ou déjà en cours. En premier lieu, un modèle qui permettrait de saisir les relations sémantiques entre les concepts juridiques dans LegTerm vient d'être proposé par Kateřina Hodková dans sa thèse de doctorat, dont la soutenance a eu lieu en septembre 2023. Dans le cadre des mémoires de master, d'autres étudiants contribuent progressivement à la création de corpus de textes juridiques spécialisés qui

¹⁶ Ivo Petrů, juriste, assure l'entretien de la structure conceptuelle et, assisté par Kateřina Hodková, veille à la terminologie juridique tchèque et française ; Jana Pešková entretient la terminologie juridique espagnole ; Jan Radimský est responsable de la conception terminographique de la base de données, coordonne les travaux de l'équipe et s'occupe, avec Alessandra Bonsignori, de la terminologie juridique italienne.

pourraient être reliés à LegTerm dans le futur. Finalement, nous envisageons également de créer dans LegTerm une interface informatique pour l'exportation des glossaires multilingues selon les critères choisis, en vue de la traduction assistée par ordinateur. L'obstacle majeur auquel le projet doit faire face réside actuellement dans le manque des « ressources humaines » appropriées : en effet, la préparation matérielle des données demande un travail manuel considérable et requiert une expertise avancée dans le domaine du droit des langues-États en question.

BIBLIOGRAPHIE

- BOCQUET Claude (2008), *La traduction juridique : fondement et méthode* (Traducto), Bruxelles, De Boeck.
- CORNU Gérard *et al.* (2020), *Vocabulaire juridique*, 13^e éd., Paris, Quadrige.
- CORNU Gérard (2005), *Linguistique juridique*, 3^e éd., Paris, Montchrestien.
- DHÉNIN Jean-François ; LAMADON Stéphane (1995), *Droit civil I*, Paris, Bréal.
- ISO 1087:2019(fr) : Travail terminologique et science de la terminologie – Vocabulaire.
- JACKENDOFF Ray (2015), *A user's guide to thought and meaning*. 1st publ. in paperback. Oxford, Oxford University Press.
- L'HOMME Marie-Claude (2020), *Lexical semantics for terminology: an introduction* (Terminology and lexicography research and practice (TLRP) volume 20), Amsterdam / Philadelphia, John Benjamins Publishing Company.
- PETRŮ Ivo (2015), Codification de la terminologie juridique dans les définitions légales, *Écho des études romanes*, vol. 11, n° 1, p. 125-135. DOI : 10.32725/eer.2015.008.
- PETRŮ Ivo (2016), La traduction juridique : entre équivalence fonctionnelle et équivalence formelle, *Études romanes de Brno* vol. 37, n° 2, p. 177-189. DOI : 10.5817/ERB2016-2-14.
- RADIMSKÝ Jan (2004), Dans quelle mesure est-il possible de traduire un terme juridique ? *Études romanes de Brno*, vol. 34, n° 1, p. 37-44.
- RONDEAU Guy (1984), *Introduction à la terminologie*, 2^e éd., Boucherville, Gaëtan Morin Éd.
- TEMMERMAN Rita (2000), *Towards new ways of terminology description: the sociocognitive-approach* (Terminology and Lexicography Research and Practice 3), Amsterdam, Benjamins.
- TOMÁŠEK Michal (2003), *Překlad v právní praxi*, 2^e éd. complétée, Praha, Linde.
- WRIGHT Sue Ellen ; BUDIN Gerhard (éds. ; 1997), *Handbook of terminology management*, Amsterdam / Philadelphia, J. Benjamins.

LATE

< <https://iate.europa.eu/> >

Juridictionnaire

< <https://www.noslangues-ourlanguages.gc.ca/fr/juridictionnaire/index-fra> >

LegTerm

< <https://legterm.cz> >

Max Planck Encyclopedia of Comparative Constitutional Law

<<https://oxcon.oup.com/home/MPECCOL>>

TERMIUM Plus

<<https://www.btb.termiumplus.gc.ca>>

United Nations Multilingual Legal Terminology Database

<<https://unterm.un.org>>

JuriDico

<http://olst.ling.umontreal.ca/?page_id=1357>