

УДК 004.822

Представлення правових знань на основі онтологій

*Карасюк В.В., к.т.н., доц.,
Національний юридичний університет імені Ярослава Мудрого,
м. Харків vl_karasuk@ukr.net*

В роботі розглядається методологія та програмний засіб для створення онтології правових знань. Особливості правової інформації наклали свої вимоги на структуру бази знань. Сучасна реалізація пакету програм на основі запропонованої методики передбачає web-інтерфейс і можливість наповнення бази багатьма користувачами.

Вступ

Правники основним об'єктом своєї праці мають документи. Вони використовують різні документи, що відносяться до певних юридичних ситуацій: нормативні документи, підзаконні акти, правозастосовчі документи, рішення, ухвали, поточні документи тощо. Доступний обсяг поширених правових баз, які пропонуються користувачам, містить біля 500 тисяч документів. Це дуже велика кількість. Тому не дивно, що серед них мають місце колізії, досить заплутана структура підпорядкування понять в окремих галузях права, невідповідність документів один одному. Ці проблеми ставлять задачу формалізації опису правових знань. До того ж, рухаючись у напрямку побудови інформаційного суспільства, ми мусимо мати впорядковане доступне електронне представлення системи правових знань.

У багатьох країнах світу вже створені бази правових знань для їх практичного використання. За основу їх опису майже у всіх випадках взята онтологічна структура. Як приклади, можна назвати такі відомі правові онтології.

Цікавими системами є FOlaw (Functional Ontology of Law) [1], LRI Core [2], Frame-based Ontology, CLO (Core Legal Ontology), Jurwordnet [3]. Слід також згадати системи юридичного виводу на основі прецедентів LCBR (legal case-based reasoning) [4,5]. Для владнання онлайн суперечок використовується платформа Ontomedia, яка працює з Core Mediation Ontology і онтологічний субдомен використовується для посередницьких функцій з окремими субдоменами (наприклад, проблемна область споживача, проблемна область моделі сім'ї, здоров'я і т.ін.). Ця модель є осно-

вою для представлення понять і відношень у посередництві [6]. В Україні відомо про роботи у напрямку створення онтології правового спрямування на базі технології METHONTOLOGY [7] та правової онтології учбового спрямування JurOnt [8].

В цілому завдання побудови онтологічних структур, що описують знання обраної предметної області, є привабливою і актуальною для багатьох сфер діяльності і досліджень, особливо для тих, які мають наповнення, що динамічно змінюються.

Мета даної роботи – розглянути особливості онтологічних моделей правової інформації для систем електронного навчання та надати характеристику підходу до наповнення системи.

Онтологія як основа формування структури правових знань

Є низка причин, за яких доцільно використовувати онтологічне уявлення знань [9]: сумісне використання людьми або програмними агентами загального розуміння структури інформації; можливість повторного використання знань проблемної області; створення явних припущень у предметній області; виокремлення знань у предметній області від оперативних знань; можливість об'єктивного аналізу знань у предметній області.

У роботі [10] вказано, за рахунок чого онтології в праві мають серйозне обґрунтування для свого застосування. Це взаємозв'язок правових норм між собою; важливість юридичних рішень; ступінь мінливості правових норм; подібність різних гілок права та інші. Тому онтологічний інжиніринг у правовій сфері має характерні риси: 1) велика кількість загальноприйнятих понять зі своєю специфікою застосування; 2) відмінність в структурах різних областей права; 3) наявність проміжного загальнотеоретичного правового рівня між онтологією верхнього рівня і онтологією предметної області; 4) велика кількість теоретичних припущень і абстрактних конструкцій, що залежать від специфіки правових поглядів. Тому варто вважати, що найбільш ефективним є підхід, заснований на вивченні природної мови та побудові онтологій на великих текстових корпусах.

При створенні онтологій важливе значення має оцінка їх придатності та якості на основі системи критеріїв. Всебічна типологія критеріїв якості онтології передбачає [11]: епістемологічну адекватність (епістемологічна ясність, епістемологічна інтуїція, епістемологічна доречність, епістемологічна закінченість); експлуатаційна адекватність (послідовність, обчислюваність); можливість багаторазового використання (завдання і методу,

предметної області). Ці критерії доцільно використовувати як стандарт для опису всіх онтологій.

Аналіз розглянутих проєктів, в яких застосовані онтологічні принципи опису правових знань та методик їх побудови наводить на думку, що в даний час не напрацьовані оптимальні принципи побудови онтологічних інформаційних систем.

Особливості структури правової онтології

Пропонується структуру правових знань уявити онтологією, яка описує зв'язок пов'язаних понять і належних до них описів. Для проєктування структури онтології правової інформації слід врахувати ряд особливостей, притаманних правовій інформації, і в особливості звернути увагу на наступні:

- синонімічність визначень (вузлів онтології);
- обмеженість конкретних формулювань нормативних документів у часі;
- наявність обов'язкового зв'язку визначень (вузлів онтології) зі строгими формулюваннями (законодавчими визначеннями) у нормативних документах.

Тобто база даних, що реалізує онтологію, включає такі частини:

- поняття і зв'язки;
- зв'язки між групами понять;
- тексти - джерела;
- словесне відображення понять і зв'язків;
- індекси вживання понять і зв'язків у тексті.

Розглянемо кожну частину.

- поняття та зв'язки: записуються з унікальними ідентифікаторами, з рядком назви для виведення в графічному інтерфейсі.

- тексти - джерела: зберігаються у вигляді набору рядків - речень, з фіксацією належності до розділу і тексту, тобто серед них можна виділити наступні сутності:

- текст - упорядкована множина розділів, тип, назва, список авторів та коментарі;

- розділ - назва, упорядкована множина підрозділів і впорядкована множина речень;

- речення - текст у вигляді рядка і типу речення (з вказівкою належності до основного тексту розділу або його заголовка).

- словесне відображення понять і зв'язків схематично можна зобразити так, як показано на рис. 1.

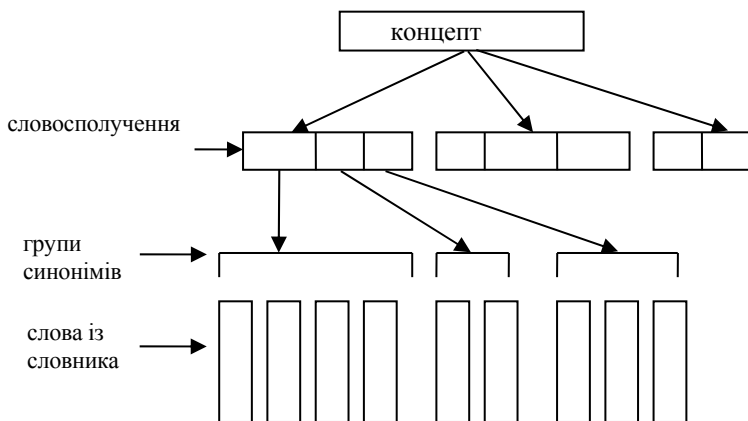


Рис.1. Структурне відображення понять і їх зв'язків у базі даних онтології.

Практична реалізація онтології правової інформації

Запропонована структура моделі знань реалізована у вигляді програмного комплексу, названого JurOnt. Комплекс виконаний у вигляді чотирьох підсистем, з використанням сучасних технологій об'єктно-орієнтованого візуального програмування, в середовищі Eclipse 3.4 на мові Java з підтримкою JDK версії 1.6. Передбачений web інтерфейс користувача і автоматизований режим роботи з базою знань, у тому числі автоматизоване наповнення онтології з доступних текстових документів.

Передбачаються наступні режими роботи з системою.

- Режим створення онтологічної системи. У цьому режимі фахівці - когнітологи розробляють ядро онтологічної системи і наповнюють середу первинними знаннями .

- Режим експлуатації. Передбачає одночасне використання системи багатьма користувачами з доступом до знань в єдиній базі.

- Режим індивідуального доповнення, розвитку та інтеграції. Передбачає роботу в системі з доповненням існуючої бази знань і створенням власної бази. Надалі передбачається інтеграція баз знань, створених окремими користувачами, в єдину систему, з перевіркою можливих суперечностей у тлумаченнях понять.

• Навчання та оцінювання знань. У цьому режимі створюється система оцінювання на основі послідовностей взаємодії понять і формування онтології «питання - відповідь».

• Онтологічний аналіз. В основі онтологічного аналізу лежить опис предметної області в термінах понять, їх взаємозв'язків, а також перетворення понять в процесі прийняття рішень.

Система розробляється таким чином, що можлива робота користувача (експерта) з кожним реченням висхідних текстів послідовно. Тобто послідовність роботи буде наступною:

- користувач читає чергове речення і усвідомлює його зміст;
- виділяє поняття, які згадуються в реченні;
- визначає зв'язки між групами понять;
- встановлює додаткові зв'язки, які уточнюють зміст основних зв'язків або описують деяким чином складні поняття, у тому числі пов'язує їх з законодавчими визначеннями.

Далі йде технічний етап внесення знань у базу даних.

Основною підсистемою, що визначає межі використання системи, є аплікація користувача. Глобальними завданнями аплікації користувача є: 1) навігація в онтології; 2) пошук фрагментів текстів - джерел, що відповідають елементам онтології; 3) перегляд текстів - джерел у повному обсязі, у вигляді тексту, поділеного по розділам, маркованого тексту.

Кожне із зазначених завдань припускає реалізацію певних сценаріїв роботи з аплікацією.

Система JurOnt в даний час прийнята в дослідну експлуатацію в локальній мережі Національного юридичного університету імені Ярослава Мудрого. Іде процес наповнення бази знань інформацією з області кримінального права України. Експерти та користувачі системи будують онтологію на основі змісту підручника з кримінального права, який є базовим для вивчення цієї дисципліни в університеті і з будь-яких інших джерел. Надалі в базу знань буде включена інформація з суміжних галузей права.

Слід наголосити на тому, що процес наповнення онтології буде виконуватись великою кількістю користувачів. Це, з одного боку, дає можливість швидко просуватися у наповненні онтології, а з іншого – створює проблеми, пов'язані із узгодженням фрагментів онтології, створених різними користувачами і включенням їх до основної онтології. Поки що це відбувається за рахунок розмежування гілок онтології для заповнення окремими користувачами. У наступному планується побудова автоматизованої (а у перспективі – і автоматичної) процедури злиття фрагментів

онтології. Тобто передбачається використання принципів краудсорсінгу при наповненні онтології.

Висновок

У результаті дослідження, аналізу переваг і недоліків різних систем подання знань обрана онтологічна модель знання. Сформована структура бази даних, яка реалізує базу знань; розроблені програмні додатки, які орієнтовані на технологію "клієнт-сервер" і забезпечують побудову семантичної мережі на сервері з можливістю багатокористувальницької роботи. Система прийнята в дослідну експлуатацію в Центрі інформаційних технологій Національного юридичного університету.

Перспективними для подальших досліджень є такі напрямки:

- розробка засобів автоматичного порівняння онтологій, це важливий інструмент для оцінки повноти, непротиворіччя онтологій, створених різними користувачами, а також для порівняння онтологій різних напрямків юридичної діяльності або різних країн, за рахунок розробки метрики близькості;

- розробка інтерфейсу системи на природній мові, це важлива частина створення інтелектуальних систем, яка дозволить залучати до роботи непрофесійних користувачів, або використовувати систему для автоматичного наповнення баз знань довільних текстів;

- створення інтернет-порталу для надання консультаційних юридичних послуг без участі фахівця при використанні запропонованої системи в юридичній клініці;

- дослідження впливу принципів самоорганізації на якість створюваної множиною користувачів онтології в предметній області правознавства.

Література

1. *Breuker J., Hoekstra R.* Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law // Proceedings of the EKAW04 Workshop on Core Ontologies in Ontology Engineering, Northamptonshire, UK. - p. 15 – 27. Режим доступу: <http://dare.uva.nl/document/8751>.
2. *Gangemi A., Prisco A., Sagri M.T., Steve G., Tiscornia D.* Some ontological tools to support legal regulatory compliance, with a case study // Workshop WORM Core, LNCS: Springer Verlag. – 2003. - p. 607–620. Режим доступу: <http://www.loa.istc.cnr.it/Papers/WORM-CORE.pdf>

3. *Sagri M. T., Tiscornia D., Bertagna F.* Jur-WorNet // Second International Wordnet Conference - GWC - 2004. Brno: Masaryk University. - p. 305-310.
4. *Henderson J., Bench-Capon T.* (2001) Dynamic arguments in a case law domain. // ICAIL '01: proceedings of the 8th international conference on artificial intelligence and law. New York: ACM Press. - 2001. - p 60–69. DOI:10.1145/383535.383542
5. *Zeng Y., Wang R., Zeleznikow J., Kemp E.A.* (2005) Knowledge representation for the intelligent legal case retrieval. // KES (Part 1). Lecture Notes in Computer Science, Springer. – 2005. – Vol. 3681. – p. 339 – 345. DOI: 10.1007/11552413_49
6. *Poblet M., Casanovas P. López-Cobo J-M. Castellás N.* ODR, Ontologies, and Web 2.0. // Jurnal of Universal Computer Science. – 2011. – Vol. 17(4). – p. 618-634. Режим доступа: http://www.researchgate.net/publication/258046605_ODR_Ontologies_and_Web_2.0
7. *Хала Е.* О построении онтологии для правовой области с применением технологии METHONTOLOGY / Е.А. Хала // Збірник наукових праць. К.: Інститут проблем моделювання в енергетиці ім. Г.Є.Пухова. – 2012. – Випуск 64. – с. 64 – 71.
8. *Tatsyi V., Getman A., Ivanov S., Karasiuk V., Lugoviy O., Sokolov O.* Semantic network of knowledge in science of law // Automation, Control, and Information Technology (ACIT 2010): Proceedings of the IASTED International Conference on Automation, Control, and Information Technology, Novosibirsk, Russia / The International Association of Science and Technology for Development. – Anaheim, USA, Zurich, Switzerland: ACTA Press. - 2010. p. 218 – 222.
9. *Noy N., McGuinness D.* Ontology development 101. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05. – 2001. Режим доступа: <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf>
10. *Bench-Capon T., Visser P.* (1997) Ontologies in Legal Information Systems; The Need for Explicit Specifications of Domain Conceptualisations // Proc. of the 6th Int. Conf. on Artificial Intelligence and Law. Melbourne, Australia. – 1997. – p. 132-141. Режим доступа: <http://www.egov.ufsc.br/portal/sites/default/files/anexos/5718-5710-1-PB.pdf>
11. *Visser P., Bench-Capon T.* A Comparison of Four Ontologies for the Design of Legal Knowledge Systems // Artificial Intelligence and Law. – 1998. – № 6. – p. 27-57.