

УДК 004.91+502+34

М. А. ПОПОВА, Є. С. АНПІЛОВА

*Інститут телекомунікацій і глобального інформаційного простору НАН України***ІНСТРУМЕНТИ ПІДТРИМКИ ДІЯЛЬНОСТІ ЕКСПЕРТА ПРИ КОНЦЕПТ-АНАЛІЗІ ПРИРОДНО-МОВНИХ ТЕКСТІВ (НА ПРИКЛАДІ ПОРІВНЯННЯ ЗАКОНОДАВЧИХ АКТІВ В СФЕРІ ОХОРОНИ ТА РАЦІОНАЛЬНОГО ВИКОРИСТАННЯ ВОДНИХ РЕСУРСІВ УКРАЇНИ І ТУРКМЕНИСТАНУ)**

В статті висвітлено значення води для економічного, екологічного та політичного добробуту населення світу та пов'язана з ним необхідність модернізації законодавства України в сфері охорони та раціонального використання водних ресурсів на основі застосування кращого світового досвіду. Описані інструменти проведення концепт-аналізу природно-мовних текстів на основі онтологічного підходу. Наведено приклад використання інструментів експертом-аналітиком при порівнянні тематичних документів з водоохоронного законодавства для виявлення повноти понятійного опису, суперечностей та невизначеностей.

Ключові слова: *концепт-аналіз, онтологічний підхід, природно-мовний текст, законодавство в сфері охорони та використання водних ресурсів.*

ВСТУП

Вода є основою життя, екосистем і людської діяльності. При раціональному використанні вода - це багаті врожаї, здоров'я, процвітання народів і націй Землі. При нераціональному використанні або за відсутності управління вода приносить хвороби, повені, призводить до ерозії, засолення, заболочування, замулення, деградації навколишнього середовища і конфліктів між людьми.

З усіх природних ресурсів вода є найважливішим. Вона може і повинна сприяти економічному та соціальному прогресу всіх людей планети згідно цілей і принципів ООН, сформульованих у Статуті ООН і Декларації ООН про права людини [1]. Жодна жива істота на Землі не може прожити без води. Саме завдяки воді Земля існує сьогодні з океанами, рослинністю, живими істотами. Запаси прісної води становлять лише 2,5% від загального обсягу води на Землі. Майже вся вода, яку споживає людина, збирається з озер, річок і неглибоких підземних джерел, у той час як основні її запаси містяться в льодовиках (Антарктика, Арктика, Гренландія) і глибоких водоносних горизонтах.

Згідно Дублінської Декларації 1992 та рішення Конференції міністрів і керівників щодо питної води та оздоровлення навколишнього середовища, що відбулася в березні 1994 року в Нордвіку, Нідерланди (за підтримки Комісії зі сталого розвитку та Генеральної Асамблеї ООН), проаналізувавши концепцію Всесвітнього водного форуму, Міжнародна асоціація водних ресурсів, під час засідання в Каїрі в

1994 р, доручила комітету виконати підготовчу роботу з метою створення Всесвітньої водної ради (ВВР). Цей Комітет спочатку зібрався в Монреалі (Канада) в березні 1995 р, потім у Барі (Італія) у вересні 1995 р. Ці дві зустрічі визначили місію та завдання ВВР, офіційно заснованою в Марселі 14 червня 1996 року [2].

На Всесвітньому водному форумі в Марселі експерти ООН проголосили, що світ знаходиться на межі водної катастрофи. Кожен десятий житель Землі відчуває гостру нестачу питної води, а це майже 780 млн. осіб. У той же час за розрахунками французької організації Solidarites International [3] без доступу до чистої питної води у світі в даний час залишаються 1,9 млрд. осіб населення Землі. За прогнозами експертів ООН, до 2050 р. необхідність у воді збільшиться на 20%. Кожна людина щодня потребує від 2 до 4 літрів води, але основна частина питної води витрачається на те, щоб зробити необхідні продукти харчування. Наприклад, для виробництва 1 кг. яловичини потрібно 15 тис. л. води, а для 1 кг. пшениці - 1,5 тис. л.

Багато країн вже досягли граничних можливостей водокористування. В близькому майбутньому проблема нестачі водних ресурсів перетвориться на політичну проблему, на думку експертів ООН. Якщо нічого не робити, то без задовільно очищеної води до 2030 року залишатимуться майже 5 млрд. осіб. Брак води в пустельних і напівпустельних регіонах призведе до інтенсивної міграції населення. За рік понад 20 млн. осіб у світі залишили свої будинки через дефіцит води.

Крім того, за даними ООН, майже 80% складають захворювання, пов'язані з якістю води, через які у країнах, що розвиваються, щороку гине майже 3 млн. осіб. Прісна вода найближчим часом може стати стратегічним ресурсом. Експерти вже зараз говорять про ймовірність водних військових конфліктів.

Влада над глобальними водними ресурсами поділена між народами світу, сотнями тисяч місцевих урядів і численними неурядовими та приватними організаціями, а також великою кількістю міжнародних органів. Законодавство з використання та охорони вод в кожній окремо взятій державі відрізняється залежно від економічних та екологічних умов.

1. Досвід Туркменістану

12 квітня 2015 в м. Тегу (Корея) відбувся 7-й Всесвітній Водний Форум, в якому взяли участь понад 3000 осіб, в тому числі голови шести держав (Угорщина, Корея, Таджикистан, Туркменістан, Ефіопія, Марокко), голови міжнародних організацій і міністри з водних та суміжних питань. Промовці відзначили зростаюче значення водних ресурсів для економічного розвитку країн, підвищення добробуту населення, забезпечення екологічної безпеки та підтримання миру в усьому світі [4].

В якості приймаючої сторони Корея запропонувала активізувати зусилля міжнародного співтовариства щодо застосування інноваційних технологій у поєднанні з традиційними підходами, включивши в якості додаткових процесів науку і технології.

Виступаючи на Всесвітньому водному форумі президент Туркменістану Гурбангули Бердимухамедов, висунув ініціативу з розробки водної стратегії ООН. Тема раціонального водокористування в Центральній Азії є сьогодні однією з актуальних, а Туркменістан в цьому питанні проявляє велику активність. В умовах Туркменістану, 70% території якого займає пустеля Каракуми, значення води особливо важливо [5].

Існує п'ять основних джерел водних ресурсів Туркменістану, а саме: основні транскордонні водотоки (Амудар'я, Сирдар'я, Атрек, Мургаб і Теджен), малі річки північно-західного схилу Копетдага, прісноводні лінзи підземних вод, а також невеликі природні озера. Серед водогосподарських об'єктів можна перерахувати Каракумський канал, штучні водосховища та озера, іригаційні канали, колекторно-дренажна мережа, штучні дренажні озера і озеро Алтин асир (в перекладі «Золоте століття»).

Кодекс Туркменістану «О воде» [6] юридично закріплює структуру державної влади та перерозподілу функцій і повноважень з управління водними

ресурсами. Будучи складовою частиною загальнодержавної законодавчої системи, Водний Кодекс покликаний забезпечити правове регулювання водних відносин (управління, раціональне використання, охорона, та ін.) на засадах сталого розвитку держави, економіки та суспільства [7]. Документа був розроблений на підставі аналізу результатів реформи законодавства, що регулює приватне підприємництво, а також пакету законодавчих актів у галузі охорони навколишнього середовища. З цієї точки зору важливим є те, що Кодекс успадкував кращі риси колишньої правової системи, збагатившись сучасними досягненнями в галузі права і ввібравши в себе національні традиції дбайливого ставлення до води і землі. Водний Кодекс Туркменістану «О воде» підтримує, але не підштовхує позитивні тенденції, створює сприятливі умови для самостійного і відповідального господарювання, сприяє розвитку приватного підприємництва на принципах раціонального використання природних багатств країни.

2. Ситуація в Україні

В Україні більше 63 тис. річок і струмків, загальна протяжність яких перевищує 206 тис. км. З них близько 60 тис. (93%) дуже малих (тобто довжина яких складає менше 10 км) сумарною довжиною - 112 тис. км, середня довжина такого водотоку - 1,9 км; малих річок, що мають довжину більше 10 км, налічується 3212, а їх загальна довжина складає близько 74 тис. км.

Переважає більшість малих річок довжиною менше 10 км має площу водозабору від 20,1 до 500 км² (87% загальної кількості і 72% загальної довжини малих річок, що протікають територією України).

Малих річок з площею водозабору від 50,1 до 100 км² налічується 890 (28% загальної кількості), а 797 річок (25%) мають площу водозабору 20,1-50 км².

До великих річок належать Дунай, Тиса, Дністер, Південний Буг, Дніпро, Прип'ять, Десна, Сіверський Донець, Західний Буг.

Водозабори великих і багатьох середніх річок розташовані в кількох географічних зонах та геоморфологічних областях; малі річки течуть в межах однієї геоморфологічної області [8]. Сумарний обсяг водних ресурсів становить 209,8 км³ на рік, у тому числі 52,4 км. на рік стік місцевого формування. Маємо характерно нерівномірний територіальний розподіл водних ресурсів. Питома водозабезпеченість на одного жителя мала - 1,7 тис. м³ на рік, а на 1 км² території - 86,8 тис. м³ на рік. Краще за інших забезпечені водними ресурсами західні райони України, найменш забезпечені її південні райони і

Донбас. Україна, незважаючи на велику кількість річок, відноситься до одного з найбільш маловодних регіонів Європи.

За даними Українського державного науково-дослідного інституту «УкрВОДГЕО», 69% всієї питної води, яка поставляється в оселі українців, не відповідає встановленим санітарним нормам. Держава не здатна вирішити цю проблему, і в найближчі десятиліття ця проблема буде наростати.

Єдиним правильним та ефективним кроком для вирішення цих проблем постає модернізація законодавства України у сфері використання та охорони водних ресурсів. Аналіз та відбір кращого досвіду в сфері світового водного законодавства дозволить адаптувати нормативно-правову базу України у сфері використання та охорони водних ресурсів до сучасних еколого-економічних реалій країни.

У даній статті розглядається онтологічний підхід і сучасний інструментарій підтримки аналітичної діяльності експерта при концепт-аналізі нормативно-правової документації на прикладі порівняння текстів Водного Кодексу України та Кодексу Туркменістану «Про воду» з метою виявлення повноти опису понятійних категорій законодавства, виявлення двоякого тлумачення понять тощо.

3. Інструменти концепт-аналізу

Найбільш ергономічним типом уявлення неструктурованого інформаційного масиву є модель, що відображає всі концепти-поняття та концептуальні зв'язки між ними [9]. Така модель - комп'ютерна онтологія - здатна якісно забезпечити процес аналізу експерта за рахунок відображення у вигляді таксономії - структури, ієрархічного упорядкування, класифікації, узагальненої логічної ознакової моделі класів концептів-понять. Структурування інформації забезпечує ефект швидкого пошуку необхідної інформації у відносно повільній людській пам'яті [10]. Процес послідовного прочитання тексту не дає можливості отримати повну інформацію стосовно поняття миттєво, одноразово (наприклад, поняття зустрічається в різних розділах). Тому процес аналізу комп'ютерної онтології тексту повинен допускати можливість «дозування», аналізу окремих частин моделей класів концептів. Оскільки в різних процесах обробки інформації поняття представляються одним з двох способів: ім'ям (згорнуте, конверговане подання) або у вигляді набору значень ознак (розгорнуте уявлення), таксономія повинна забезпечувати зручний перехід від одного подання до іншого.

Для лінгвістичного аналізу текстів кодексів використовується модуль «КОНСПЕКТ» [11] (рис. 1) системи ТОДОС [12].

Формалізований опис відображення вхідного

природно-мовного тексту (ПМТ) з відтворенням повної акторної складової семантики можна записати у вигляді:

$$G : C_O(\text{ПМТ}) \rightarrow C_A(\text{ПМТ}), \quad (1)$$

де $C_O(\text{ПМТ}) = \bigcup_{i=1}^N C_{Лі}$, $C_{Лі}$ - формалізований опис

i -ої компоненти об'єктної складової семантики слів повнозначних частин мови, які входять у вихідний ПМТ, $i = \overline{1, N}$ - кількість вказаних слів;

$C_A(\text{ПМТ}) = F_A(C_O)$ - формалізований опис акторної складової семантики вихідного ПМТ.

В якості початкового етапу лінгвістичного аналізу виконується графемно-морфологічний аналіз, призначенням якого є побудова моделі графемної структури вхідного тексту, в якій виділено і зв'язано відношеннями (де це можливо) такі змістовні одиниці тексту, як фрагмент, речення та лексема. Результатом першого етапу є впорядкований список складників природно-мовного тексту - концептів-понять x предметної області X , $x \in X$, із зазначенням для кожного з них чи є він поняттям-об'єктом або поняттям-класом об'єктів, можливих суттєвих відношень з іншими поняттями зі списку та істотних властивостей понять. На основі отриманих результатів формується таблиця класів понять аналізованого тексту, об'єднаних за властивостями на основі множини семантичних відповідностей між ними [13].

Генерація таксономії на основі класифікації концептів відбувається методом індуктивного формування понять в середовищі модулю «Graph Editor (ЕДИТОР)» [14, 15] (рис. 2, 3) системи ТОДОС [12]. Семантично значущі відношення між концептами-поняттями x , відібраними на попередньому етапі аналізу тексту, генеруються шляхом встановлення між елементами таксономії $x_i \in X$ семантичних k -арних відношень $R_k \subseteq X^k$, які представляють множини ребер, що зв'язують вузли орієнтованого графа $G = (X, R)$. Сформовані класи співвідносяться з результатами попереднього етапу, зокрема, уточнюються залежності для конкретних пар концептів-понять (x_i, x_j) .

Вершини, що не мають вхідних ребер, відповідають концептам-поняттям, решта вершин - класам концептів-понять. Вершини графа, що входять у вершину класу концептів-понять, утворюють її субмножину. Множина вершин, до яких є шляхи від вершини, називається її супермножиною. У субмножині і супермножині вершин виділяються 0-субмножина і 0-супермножина, що складаються з тих вершин, які пов'язані з нею безпосередньо.

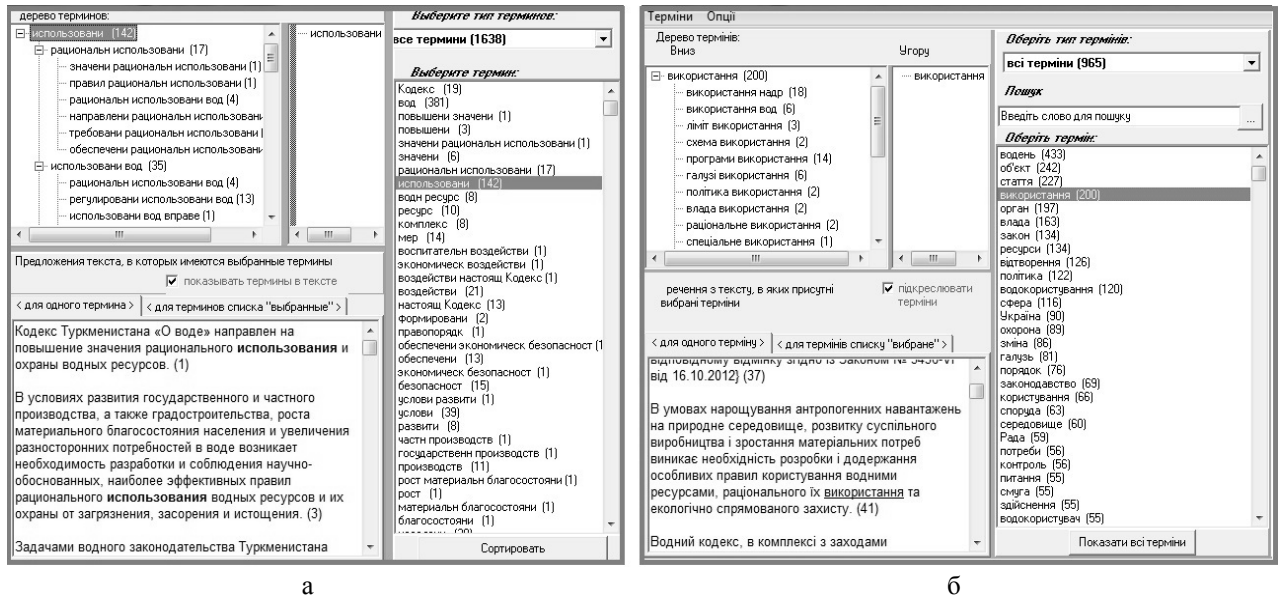


Рис.1. Фрагмент інтерфейсу користувача модулю «КОНСПЕКТ» з проаналізованим текстом і виділеними концептами-поняттями з текстів:
 а – Кодексу Туркменістану «О воде»; б – Водного Кодексу України

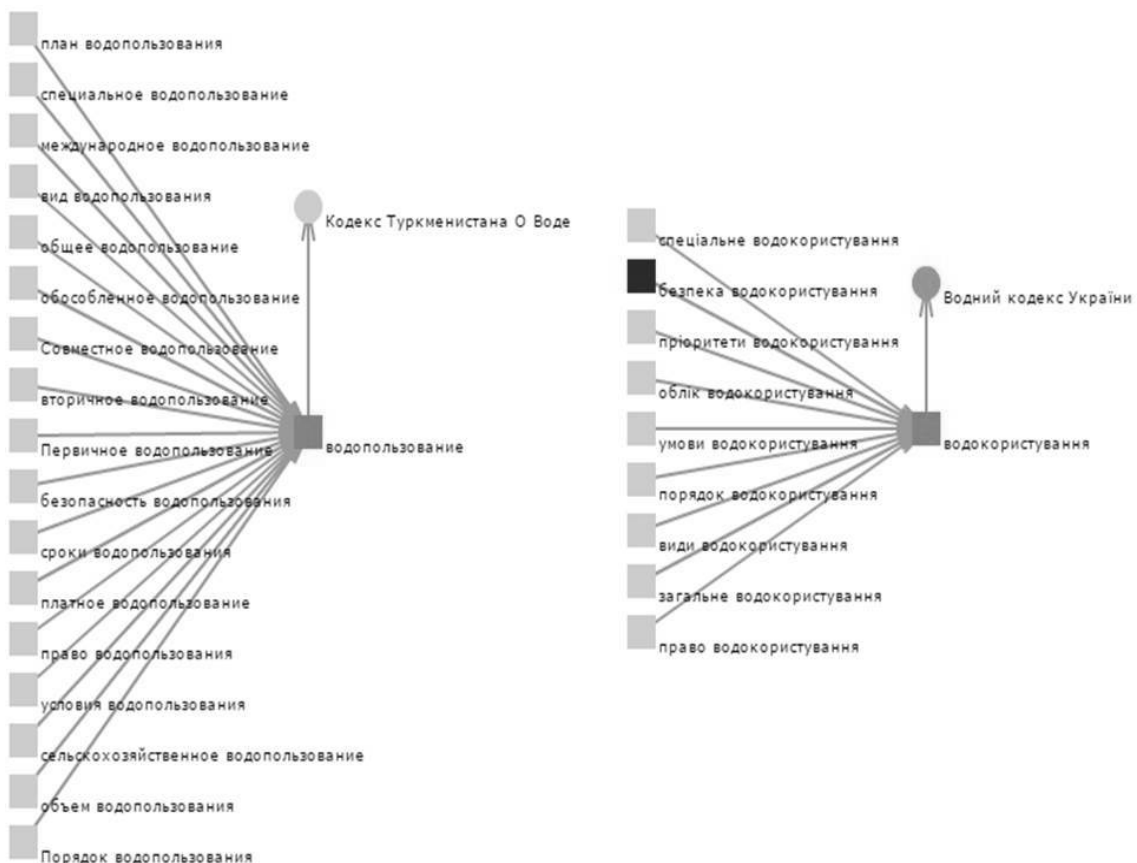


Рис. 2. Фрагмент таксономій Кодексу Туркменістану «О воде» та Водного Кодексу України.
 Субмножина поняття «водокористування»

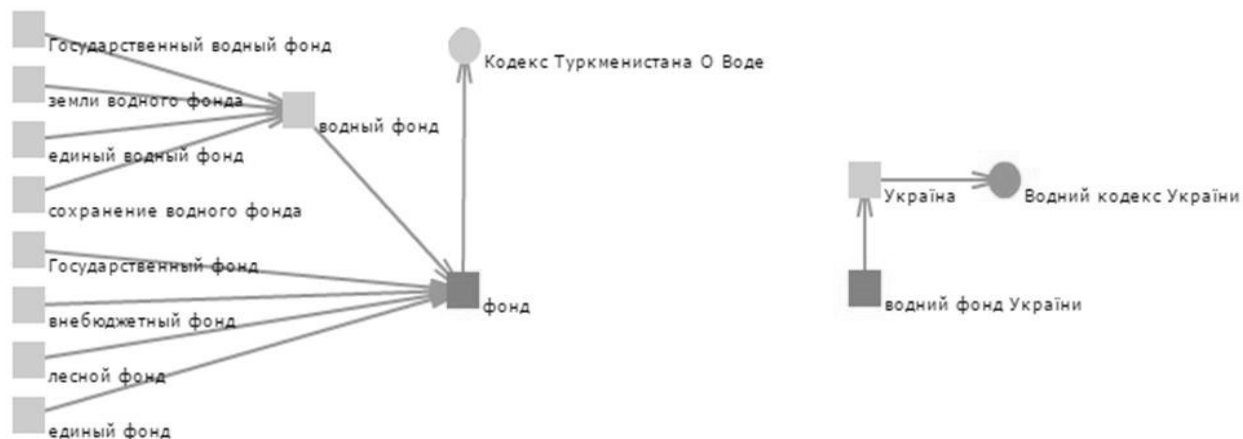


Рис. 3. Фрагмент таксономії Кодексу Туркменістану «О воде» та Водного Кодексу України. Поняттєне оточення концепту «фонд»

Функціонал модуля «Graph Editor (ЕДІТОР)», допомагає експерту з водоохоронного законодавства аналізувати характер повноти подання в тексті кодексу того чи іншого поняття, наявність можливого двоякого тлумачення понять, неточності, невизначеності та взаємовиключні поняття. Такого роду аналіз дозволяє провести якісне порівняння вітчизняної нормативної документації з іноземними аналогами.

На відміну від звичайного, суб'єктивного підходу при проведенні контент-аналізу різноманітних документів, системно-онтологічний підхід припускає строгу (наскільки це можливо на даному етапі розвитку науки) структуризацію термінів і понять предметної області. Основна технологія аналітики - встановлення причинно-наслідкових зв'язків між різного роду даними та їх дослідження під різними кутами зору. Побудова подібних ланцюжків дозволяє оброблені дані перетворити на інформацію й, зробивши висновки в предметній області, синтезувати відповідні рекомендації для прийняття оптимального рішення.

Отже, використання онтологічних інструментів класифікації, систематизації та семантичного аналізу дає можливість експерту виявляти принципово нові, раніше невідомі, взаємозв'язки, сприяє зміщенню акцентів із пасивних методів пошуку, орієнтованих на передачу інформації, до ширшого застосування активних методів аналізу проблем і пошуку рішень, співпраці користувачів тощо.

Висновки

Описані інструменти підтримки процесів аналітичної діяльності експерта водоохоронного законодавства дозволяють проводити концепт-аналіз текстів на основі онтологічного підходу. Наведений приклад використання інструментів експерт-аналітиком при порівнянні текстів Водного Кодексу

України та Кодексу Туркменістану «О воде» ілюструє необхідність переосмислення та модернізації сучасного законодавства з огляду на екологічні, соціально-економічні та політичні реалії. Запропонований інструментарій можна застосовувати для якісного аналізу нормативної документації з різних сфер законодавства з метою виявлення суперечностей і взаємовиключних понять.

Література

1. Офіційний сайт *Charter of the United Nation* [Електронний ресурс]. – Режим доступу: <http://www.un.org/ru/documents/charter/> (дата звернення 02.07.2015). Назва з екрану.
2. Офіційний сайт *World Water Council* [Електронний ресурс]. – Режим доступу: http://www.cawater-info.net/int_org/www/ (дата звернення 02.07.2015). Назва з екрану.
3. Офіційний сайт *Solidarites International* [Електронний ресурс]. – Режим доступу: <http://www.solidarites.org/en/> (дата звернення 02.07.2015). Назва з екрану.
4. Офіційний сайт *7th World Water Forum* [Електронний ресурс]. – Режим доступу: <http://eng.worldwaterforum7.org/main/> (дата звернення 02.07.2015). Назва з екрану.
5. Офіційний сайт *Turkmenistan Live* [Електронний ресурс]. – Режим доступу: <http://turkmenistanlive.com/2015/04/13/> (дата звернення 02.07.2015). Назва з екрану.
6. Кодекс Туркменістану «О воде» [Електронний ресурс]. – Режим доступу: <http://www.cawater-info.net/library/eco-tm.htm> (дата звернення 02.07.2015). Назва з екрану.
7. Бердыев, А. Национальное водное право Туркменистана и его увязка с международным водным правом. Приоритеты и проблемы. Рекомендации по улучшению [Електронний ресурс] / А. Бердыев. – Режим доступу: http://www.cawater-info.net/review/legal_tm.htm. (дата звернення 02.07.2015). Назва з екрану.

8. Паламарчук, М. М. Водний фонд України [Текст] : Довідниковий посібник / М. М. Паламарчук, Н. Б. Загорчевна ; за ред. В. М. Хорєва, К. А. Алієва. – К. : Ніка-Центр. – 2002. – 392 с.

9. Sowa, J. F. Knowledge Representation: Logical, Philosophical, and Computational Foundations [Text] / J. F. Sowa. – Pacific Grove, CA : Brooks Cole Publishing Co., 2000. – 594 p.

10. Гладун, В. П. Тематичний аналіз природно мовних текстів [Текст] / В. П. Гладун, В. Ю. Величко, Л. А. Святогор // Комп'ютерна лінгвістика та інтелектуальні технології : праці міжнародної конференції «Діалог 2006» (Бекасово, 31 травня - 4 червня 2006). – М. : Изд-во РДГУ. - 2006. - С. 115-118.

11. Величко, В. Автоматизоване створення тезауруса термінів предметної області для локальних пошукових систем [Текст] / В. Величко, П. Волошин, С. Світла // «Knowledge - Dialogue - Solution» International Book Series «Information Science & Computing». - FOI ITHEA Sofia, Bulgaria., 2009. - № 15. - P. 24-31.

12. Стрижак, О. Є. Засоби онтологічної інтеграції і супроводу розподілених просторових та семантичних інформаційних ресурсів [Текст] /

О. Є. Стрижак // Екологічна безпека та природокористування : Збірник наукових праць / М-во освіти і науки України, Київ. Нац. ун-т буд-ва і архіт., НАН України, Ін-т телекомунікацій і глобал. інформ. Простору ; редкол. : О. С. Волошкіна, О. М. Трофимчук [та ін.]. - К., 2013. - Вип. 12. - С. 166-178.

13. Попова, М. А. Модель онтологического интерфейса агрегации информационных ресурсов и средств ГИС [Текст] / М. А. Попова // International Journal "Information Technologies and Knowledge". – 2013. – Vol. 7, № 4. – P. 362–370.

14. Приходнюк, В. В. Графічне представлення та візуалізація зростаючих пірамідальних мереж і онтології предметних областей [Текст] / В. В. Приходнюк // Матеріали конференції «Управление знаниями и конкурентная разведка» в рамках Международного молодежного форума «Радиоэлектроника и молодежь в XXI веке». – Харьков : ХНУРЭ, 2012. – Т. 8. – С. 120-121.

15. Офіційний сайт Graph Editor Service [Електронний ресурс]. – Режим доступу: <http://editor3.inhost.com.ua> (дата звернення 02.07.2015). Назва з екрану.

Надійшла до редакції 2.07.2015, розглянута на редколегії 11.09.2015

ИНСТРУМЕНТЫ ПОДДЕРЖКИ ДЕЯТЕЛЬНОСТИ ЭКСПЕРТА ПРИ КОНЦЕПТ-АНАЛИЗЕ ЕСТЕСТВЕННО-ЯЗЫКОВЫХ ТЕКСТОВ (НА ПРИМЕРЕ СРАВНЕНИЯ ЗАКОНОДАТЕЛЬНЫХ АКТОВ В СФЕРЕ ОХРАНЫ И РАЦИОНАЛЬНОГО ИСПОЛЬЗОВАНИЯ ВОДНЫХ РЕСУРСОВ УКРАИНЫ И ТУРКМЕНИСТАНА)

М. А. Попова, Е. С. Анпилова

В статье описано значение водных ресурсов для экономического, экологического и политического благосостояния населения мира и необходимость модернизации законодательства Украины в сфере охраны и рационального использования водных ресурсов на основе применения лучшего мирового опыта. Описаны инструменты проведения концепт-анализа естественно-языковых текстов на основе онтологического подхода. Приведен пример использования инструментов экспертом-аналитиком при сравнении тематических документов водоохранного законодательства для выявления полноты понятийного описания, противоречий и неопределенностей.

Ключевые слова: концепт-анализ, онтологический подход, естественно-языковой текст, законодательство в сфере охраны и использования водных ресурсов.

INSTRUMENTS TO SUPPORT OF EXPERT'S ACTIVITY IN NATURAL LANGUAGE TEXTS CONCEPT-ANALYSIS (BY COMPARING LEGISLATIVE ACTS IN THE SPHERE OF PROTECTION AND WATER CONSERVATION OF UKRAINE AND TURKMENISTAN)

М. А. Popova, Y. S. Anpilova

This paper considers the important role of water resources for economic, environmental and political well-being of the world population. Demonstrate necessitate for modernization of Ukrainian legislation in the sphere of water protection and conservation sphere with the application of actual international practices. The instruments of natural language texts concept-analysis based on the ontological approach were described. An example of the use of instruments by expert in comparing thematic documents protection legislation to identify the completeness of the conceptual description, contradictions and uncertainties was given.

Keywords: concept-analysis, ontological approach, natural language texts, legislation in the sphere of protection and water conservation.

Попова Марина Андріївна – канд. техн. наук, науковий співробітник, Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна, e-mail: pma1701@gmail.com.

Анпилова Євгенія Сергіївна – канд. техн. наук, старший науковий співробітник, Інститут телекомунікацій і глобального інформаційного простору НАН України, Київ, Україна, e-mail: anpilova@ukr.net.