

Ключевые слова: надежность, отказ, дискретное устройство, композитный материал, анализ.

Златкин Артур Анатолійович, доктор технічних наук, професор, кафедра інформаційних технологій проектування, Черкаський державний технологічний університет, Україна, e-mail: kafedra.itp@gmail.com.

Кравченко Ольга Віталіївна, старший викладач, кафедра інформаційних технологій проектування, Черкаський державний технологічний університет, Україна, e-mail: kravchenko_ov@ukr.net.

Вовчановський Олександр Сергійович, кафедра інформаційних технологій проектування, Черкаський державний технологічний університет, Україна, e-mail: woolf101@yandex.ru.

Златкин Артур Анатольевич, доктор технических наук, профессор, кафедра информационных технологий проектирования,

Черкасский государственный технологический университет, Украина.

Кравченко Ольга Витальевна, старший преподаватель, кафедра информационных технологий проектирования, Черкасский государственный технологический университет, Украина.

Вовчановский Александр Сергеевич, кафедра информационных технологий проектирования, Черкасский государственный технологический университет, Украина.

Zlatkin Arthur, Cherkasy State Technological University, Ukraine, e-mail: kafedra.itp@gmail.com.

Kravchenko Olga, Cherkasy State Technological University, Ukraine, e-mail: kravchenko_ov@ukr.net.

Vovchanovskyy Alexander, Cherkasy State Technological University, Ukraine, e-mail: woolf101@yandex.ru

УДК 331.45

DOI: 10.15587/2312-8372.2014.27928

**Борисова Н. В.,
Каніщева О. В.,
Кочуєва З. А.**

ВИРІШЕННЯ ЗАДАЧІ АВТОМАТИЗОВАНОГО СТВОРЕННЯ ТЕРМІНОЛОГІЧНОГО СЛОВНИКА ПРЕДМЕТНОЇ ОБЛАСТІ

У статті запропоновано варіант вирішення задачі автоматизованого створення термінологічного словника предметної області за допомогою розробленої авторами інформаційної системи автоматизованого формування лексикографічних ресурсів. Для опису синтаксичних конструкцій, які відповідають контекстам визначень термінів, обрана мова лексико-синтаксичних шаблонів. Представлено оцінку ефективності результатів дослідження.

Ключові слова: інтелектуальні системи, лексико-синтаксичні шаблони, автоматизоване формування термінологічних словників.

1. Вступ

На сучасному етапі розвитку суспільства виникла нагальна потреба автоматизації процесів формування лексикографічних ресурсів. Це зумовлено тим, що зараз швидко з'являються нові предметні області й виникає проблема формалізації та моделювання цих областей знань, а також динамічно розвиваються існуючі предметні області й виникає необхідність видобування нових понять цих предметних областей, встановлення зв'язків між ними та вже відомими поняттями, часткового або повного оновлення змісту відомих понять. Сучасний рівень розвитку інформаційних технологій надає сьогодні можливість тільки частково вирішити цю проблему шляхом розробки та використання спеціалізованих інформаційних систем, в основу яких покладені інтелектуальні підходи та методи, отримані на основі моделювання діяльності людини в процесах розуміння та ідентифікації знань, що видобуваються з текстів. Цим обумовлюється актуальність проведення даних досліджень.

2. Аналіз літературних даних

Оскільки автоматизоване формування термінологічного словника предметної області здійснюється шляхом видобування з повнотекстових електронних до-

кументів предметної області термінів та їх визначень, було проведено огляд найпоширеніших та найпопулярніших систем видобування термінів, а саме Simple Concordance Program, MonoConc Pro, Concordancer for Windows, MultiTerm Extract, PROMT Terminology Manager (PROMT TerM), Word Tabulator. Порівняльний аналіз ефективності роботи розглянутих вище програм видобування термінів здійснювався на прикладі обробки текстів двох частин навчального посібника Некоса В. Ю. «Основи загальної екології та неоекології» (мова видання російська) [1, 2]. На основі отриманих даних, можна зробити висновок, що лідерами серед програм автоматизованого видобування термінів можна вважати системи PROMT Terminology Manager (PROMT TerM) та MultiTerm Extract, які підтримують велику кількість функцій, що значно підвищує ефективність роботи та скорочує час на постредагування отриманих даних. Далі можна виділити систему Simple Concordance Program — просту у використанні та безкоштовну, яка має достатню кількість налаштувань для пошуку та демонструє хороші результати обробки текстів. Але жодна з проаналізованих програм не видобуває терміни разом з їх визначеннями та не підтримує українську мову, взагалі ефективних програм для видобування термінів з текстів українською мовою на ринку програмних засобів на даний час не представлено. Зроблені висновки дозволили сформулювати мету та завдання дослідження.

3. Формування мети і задач

Метою проведеного дослідження було вирішення задачі автоматизованого формування термінологічного словника предметної області.

Для досягнення поставленої мети необхідно було вирішити наступні задачі:

1. Визначити функціональні вимоги до інформаційної системи автоматизованого формування лексикографічних ресурсів.

2. Розробити математичне, алгоритмічне, програмне забезпечення задачі автоматизованого формування термінологічного словника предметної області.

3. Здійснити оцінку ефективності роботи інформаційної системи щодо вирішення задачі автоматизованого формування термінологічного словника предметної області.

4. Вирішення задачі автоматизованого створення термінологічного словника предметної області

Інформаційна система (ІС) автоматизованого формування лексикографічних ресурсів — це інформаційна система, яка оперує як даними (у традиційному розумінні), так і знаннями, представленими у належній формальній формі [3]. У нашому дослідженні використовуємо визначення поняття «знання», представлене у [4]: знання — це сукупність фактів, видобутих з даного тексту і представлених у формальному вигляді, яка відображає смисл, закладений у вихідному тексті.

ІС автоматизованого формування лексикографічних ресурсів дозволяє вирішити такі задачі: автоматизоване створення термінологічного словника предметної області; видобування термінів та ключових слів з текстів предметної області [5] з метою подальшого їх використання, наприклад, при індексуванні документів для покращення повнотекстового пошуку [6] тощо; автоматизоване створення словника синонімів [7]; автоматизоване створення словників сполучуваності слів предметної області [8].

У межах даної статті буде представлено вирішення задачі автоматизованого створення термінологічного словника предметної області. Алгоритм цього процесу представлено на рис. 1.

Спочатку на множині науково-технічних текстів було виокремлено дискурсивні маркери дискурсивної операції «визначення» [9], до яких відносяться такі конструкції: *будемо називати; розуміється; так званий* та ін. Далі було визначено контексти визначень термінів, що дозволило створити набір правил розташування дискурсивних маркерів по відношенню до терміну та його визначення. Потім контексти були поділені на групи. На даний момент виділено близько 20 груп контекстів, наприклад, група «*називатися*», група «*бути + називати*», група «*так + званий*», група «*розуміти*», група «*бути + розуміти*», група «*->*» та ін.

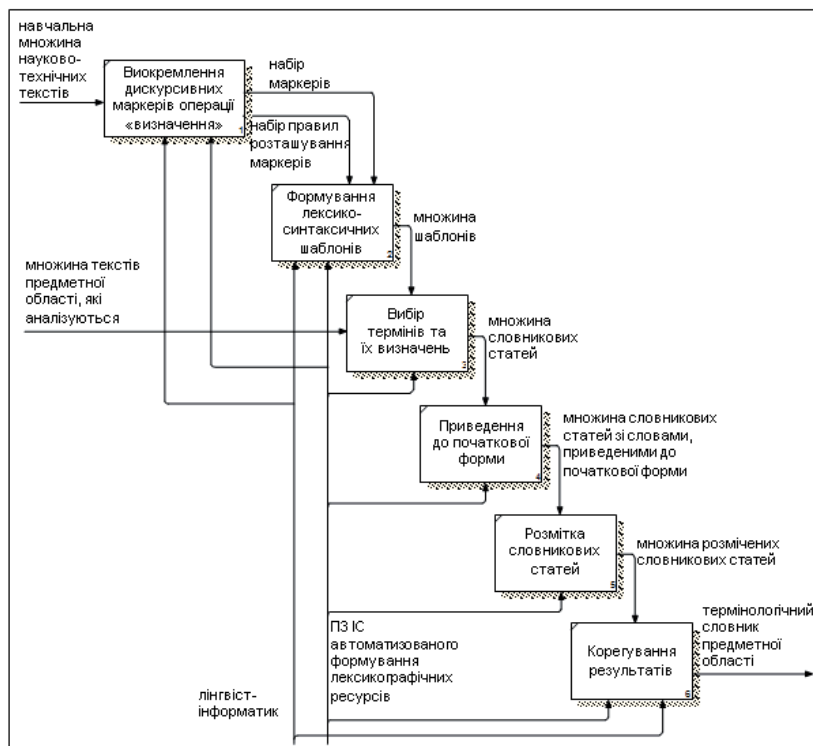


Рис. 1. Алгоритм автоматизованого створення термінологічного словника

Далі синтаксичні конструкції, які відповідають контекстам визначень термінів, було формалізовано у вигляді лексико-синтаксичних шаблонів [10, 11], які крім граматичної інформації, враховують також і правила розташування дискурсивних маркерів по відношенню до терміну та його визначення.

До складу шаблонів входять наступні елементи [10, 11]:

— літерали, тобто конкретні лексеми (наприклад, «розуміється», «називати» та ін.), а також скорочення («т.з.») й розділові знаки;

— символічні позначення слів певної частини мови та граматичної форми, які можуть заповнювати слоти (вільні місця) шаблону (наприклад, **N** — іменник, **A** — прикметник, **V** — дієслово, **Ab** — прислівник, **Nu** — числівник, **Pr** — займенник, **APa** — дієприслівник, **Pa** — дієприкметник, **P** — прийменник, **C** — сполучник);

— символічні позначення певних граматичних конструкцій (наприклад, **T** — термін-кандидат, виражений простою або складною іменною групою, причому складна іменна група може містити сполучники, групи з прийменниками, конструкції з дужками, підрядні речення, дієприкметникові звороти, **TA** — частина терміну-кандидату, виражена одним або декількома прикметниками, з'єднаними за допомогою сполучника, **TN** — частина терміну-кандидату, виражена простою або складною іменною групою, **Ng** — іменна група);

— умови, які уточнюють граматичні характеристики розглянутих елементів. При записі умов використовуються символічні позначення граматичних характеристик: часу (**tense**), числа (**number**), особи (**person**), роду (**gender**), відмінку (**case**). Ці граматичні характеристики можуть мати конкретні значення. Наприклад, граматична характеристика **tense** може приймати значення «pres», «pas», «fut»,

що відповідають теперішньому, минулому та майбутньому часові; граматична характеристика **number** може приймати значення «sing» та «plur» (відповідно однина і множина) (наприклад, запис **<Ng.number=V.number>** означає, що число іменної групи **Ng** та дієслова **V** співпадають, а умова **<person=3>** фіксує використання третьої особи).

У табл. 1 представлено приклади шаблонів, які відповідають групам контекстів визначень термінів.

Таблиця 1

Приклади шаблонів, що відповідають групам контекстів визначень термінів

Опис шаблону	Приклади речень
Група «називатися»	
T<:case=«ins»> V<:називатися>:tense=«pres» person=«3» Ng<:case=«nom»> <T.number=V.number=Ng.number>	Атмосферу називається зовнішня газова оболонка Землі, що сягає від її поверхні в космічний простір приблизно на 3000 км
T<:case=«nom»> V<:називатися>:tense=«pres» person=«3» TA<:case=«ins»> <TN.gender=TA.gender TN.number=V.number=TA.number>	Фактори називаються лімітуючими, якщо їх дії недостатньо для нормального життя будь-якого організму
Група «називати»	
T<:case=«ins»> [«ми»] V<:називати>:tense=«pres»> Ng<:case=«nom»> <TN.gender=TA.gender TN.number=V.number=TA.number>	Екологічною валентністю ми називаємо здатність організму витримувати певну амплітуду коливання фактора
Група «бути + називати»	
T<:case=«ins»> [«ми»] V<:бути>:tense=«fut» person=«1» number=«plur»> «називати» Ng<:case=«nom»> <T.number=Ng.number>	Угрупуванням тут будемо називати досить чітко окреслений та територіально сильно обмежений рівень живого
Група «так + званий»	
«так» Pa<:званий> T <Pa.case=T.case Pa.gender=T.gender Pa.number=T.number:T>	Небажаним є досягнення такого рівня взаємодії суспільства та природи, перевищення якого буде мати катастрофічні наслідки для людства, так званого «екологічного імперативу»
Група «розуміти»	
«під» T<:case=«ins»> [«ми»] V<:розуміти>:tense=«pres»> Ng<:case=«acc»>	Під біоценозом екологи розуміють історично сформовану сукупність рослин, тварин та мікроорганізмів, що населяє біотоп
Група «бути + розуміти»	
«під» T<:case=«ins»> [«тут»] [«ми»] V<:бути>:tense=«fut» person=«1» number=«plur»> «розуміти» Ng<:case=«acc»>	Під системою тут будемо розуміти цілісний комплекс взаємозв'язаних елементів, але значно складніший, ніж просто сума елементів
Група «←»	
T<:case=«nom»> «←» [«це»] Ng<:case=«nom»> <T.number=Ng.number>	Ланцюги живлення — це ряди взаємопов'язаних видів, у яких кожний попередній є об'єктом живлення наступного

Далі на вхід системи подаються тексти предметної області, з яких за допомогою заданих шаблонів виокремлюються терміни та їх визначення, причому визначення термінів видобуваються з текстів, враховуючи правила їх розташування по відношенню до самих термінів.

Обрані таким чином терміни та їх визначення піддаються процедурі нормалізації та за допомогою заданих правил розмітки перетворюються на словникові статті, з яких формується термінологічний словник. Розмітка здійснюється з використанням мови розмітки XML.

Результат роботи програми представлено на рис. 2. Готові словникові статті видаються в окремому вікні, у якому можна здійснити їх корегування та за необхідністю форматування тексту як у звичайному текстовому редакторі.

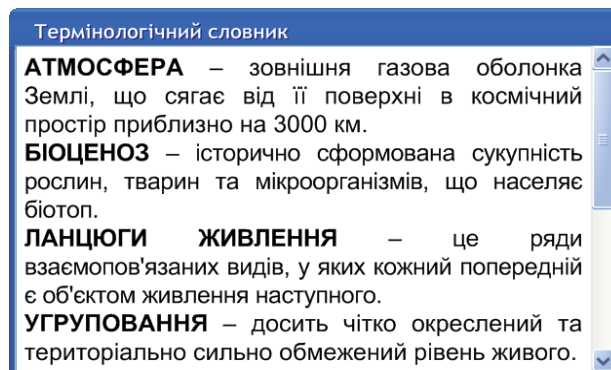


Рис. 2. Приклад роботи програми

Для оцінки ефективності були використані такі коефіцієнти: коефіцієнт точності (*precision*), коефіцієнт повноти (*recall*), помилка видобування (*error*). Обрані коефіцієнти розраховуються на основі таких показників: *a* – кількість видобутих термінів-кандидатів, які є термінами; *b* – кількість видобутих термінів-кандидатів, які не є термінами; *c* – кількість не видобутих термінів-кандидатів, які є термінами; *d* – кількість не видобутих термінів-кандидатів, які не є термінами. Було проаналізовано близько 400 документів предметної області «екологія». Були отримані такі результати: *Recall* = 0,67; *Precision* = 0,75; *Error* = 0,05.

4. Висновки

У результаті проведеного дослідження були отримані наступні результати:

1. Визначені функціональні вимоги до інформаційної системи автоматизованого формування лексикографічних ресурсів.
2. Розроблено математичне, алгоритмічне, програмне забезпечення задачі автоматизованого формування термінологічного словника предметної області.
3. Експериментально доведено ефективність роботи інформаційної системи щодо вирішення означеної задачі.

Література

1. Некос, В. Е. Основы общей экологии и неоекологии [Текст]. Часть 1. Основы общей и глобальной традиционной экологии: учебное пособие в 2-х частях / В. Е. Некос. – 2-е изд., доп. и перераб. – Харьков: Торнадо, 1999. – 192 с.
2. Некос, В. Е. Основы общей экологии и неоекологии [Текст]. Часть 2. Основы общей и глобальной неоекологии: учебное пособие в 2-х частях / В. Е. Некос. – 2-е изд., доп. и перераб. – Харьков: Прапор, 2001. – 287 с.
3. Борисова, Н. В. Информационная система автоматизованого формування лексикографічних ресурсів [Текст] / Н. В. Бори-

- сова, І. С. Ямшанов // Проблеми інформаційних технологій. — 2014. — № 01(015). — С. 193–199.
4. Петренко, М. Г. Методи та засоби побудови знання-орієнтованих комп'ютерних систем з онтолого-керованою архітектурою [Текст]: автореф. дис. ... д-ра техн. наук: 05.13.05 / М. Г. Петренко. — Київ: Інститут кібернетики ім. В. М. Глушкова, 2014. — 40 с.
 5. Борисова, Н. В. Автоматизоване видобування термінологічних одиниць з науково-технічних текстів [Текст] / Н. В. Борисова, С. С. Решетило // Матеріали III Всеукраїнської науково-практичної конференції «Інтелектуальні системи та прикладна лінгвістика», м. Харків, 17 квітня 2014 р. — Харків: НТУ «ХПІ», 2014. — С. 43.
 6. Борисова, Н. В. Индексирование полнотекстовых документов для задачи интеллектуального поиска информации по ключевым словам [Текст] / Н. В. Борисова, Э. А. Кочуева // Восточно-Европейский журнал передовых технологий. — 2014. — № 1/2(67). — С. 4–8. — Режим доступа: \www/URL: <http://journals.uran.ua/eejet/article/view/20332>.
 7. Борисова, Н. В. Автоматизированное формирование словаря синонимов [Текст] / Н. В. Борисова, О. В. Канищева, Е. Н. Юрченко // Вестник ХНТУ. — 2013. — № 1(44). — С. 91–95.
 8. Борисова, Н. В. Формирование словаря сочетаемости терминов предметной области [Текст] / Н. В. Борисова, О. В. Канищева // Восточно-Европейский журнал передовых технологий. — 2013. — № 5/3(65). — С. 16–19. — Режим доступа: \www/URL: <http://journals.uran.ua/eejet/article/view/18469>.
 9. Ефремова, Н. Э. Терминологический анализ текста на основе лексико-синтаксических шаблонов [Текст] / Н. Э. Ефремова, Е. И. Большакова, А. А. Носков, В. Ю. Антонов // Компьютерная лингвистика и интеллектуальные технологии. — 2010. — Вып. 9(16). — С. 124–129.
 10. Большакова, Е. И. Лексико-синтаксические шаблоны для автоматического анализа научно-технических текстов [Электронный ресурс] / Е. И. Большакова, Н. Э. Васильева, С. С. Морозов. — Режим доступа: \www/URL: <http://masters.donntu.edu.ua/2012/iii/sarry/library/bolshakova/Bolshakova.htm>.
 11. Васильева, Н. Э. Шаблоны употреблений терминов и их использование при автоматической обработке научно-технических текстов [Электронный ресурс] / Н. Э. Васильева. — Режим доступа: \www/URL: <http://www.dialog-21.ru/Archive/2004/Vasiljeva.htm>.

РЕШЕНИЕ ЗАДАЧИ АВТОМАТИЗИРОВАННОГО СОЗДАНИЯ ТЕРМИНОЛОГИЧЕСКОГО СЛОВАРЯ ПРЕДМЕТНОЙ ОБЛАСТИ

В статье предложен вариант решения задачи автоматизированного создания терминологического словаря предметной области с использованием разработанной авторами информационной системы автоматизированного формирования лексикографических ресурсов. Для описания синтаксических конструкций, отвечающих контекстам определений терминов, выбран язык лексико-синтаксических шаблонов. Представлена оценка эффективности результатов исследования.

Ключевые слова: интеллектуальные системы, лексико-синтаксические шаблоны, автоматизированное формирование терминологических словарей.

Борисова Наталья Владимировна, старший викладач, кафедра інтелектуальних комп'ютерних систем, Національний технічний університет «Харківський політехнічний інститут», Україна, e-mail: borisova_nv@mail.ru.

Канищева Ольга Валеріївна, кандидат технічних наук, доцент, кафедра інтелектуальних комп'ютерних систем, Національний технічний університет «Харківський політехнічний інститут», Україна, e-mail: olya-kanisheva@rambler.ru.

Кочуева Зоя Анатоліївна, старший викладач, кафедра інтелектуальних комп'ютерних систем, Національний технічний університет «Харківський політехнічний інститут», Україна, e-mail: kochueva@kochuev.com.

Борисова Наталья Владимировна, старший преподаватель, кафедра интеллектуальных компьютерных систем, Национальный технический университет «Харьковский политехнический институт», Украина.

Канищева Ольга Валерьевна, кандидат технических наук, доцент, кафедра интеллектуальных компьютерных систем, Национальный технический университет «Харьковский политехнический институт», Украина.

Кочуева Зоя Анатольевна, старший преподаватель, кафедра интеллектуальных компьютерных систем, Национальный технический университет «Харьковский политехнический институт», Украина.

Borysova Natalia, National Technical University «Kharkiv Polytechnic Institute», Ukraine, e-mail: borisova_nv@mail.ru.

Kanishecheva Olga, National Technical University «Kharkiv Polytechnic Institute», Ukraine, e-mail: olya-kanisheva@rambler.ru.

Kochueva Zoia, National Technical University «Kharkiv Polytechnic Institute», Ukraine, e-mail: kochueva@kochuev.com

УДК 621.317

DOI: 10.15587/2312-8372.2014.27926

Івіцька Д. К.

РОЗРОБКА ЕЛЕКТРОСТАТИЧНОГО МЕТОДУ НЕРУЙНІВНОГО КОНТРОЛЮ З ПІДВИЩЕНОЮ ЗАВАДОСТІЙКІСТЮ

В статті запропоновано новий електростатичний метод неруйнівного контролю з підвищеною завадостійкістю для широкого спектру матеріалів та конструкцій, діелектриків та провідних матеріалів, без використання контактної рідини. Запропоновано структурну схему приладу для використання даного методу в області неруйнівного контролю.

Ключові слова: електростатичний метод, контроль діелектриків, двохосьове сканування, підвищення завадостійкості.

1. Вступ

Останнім часом значно підвищився інтерес до створення методів неруйнівного контролю об'єкту, які не

потребують використання контактної рідини з ним. Найчастіше використовують вихрострумний метод контролю, який передбачає наведення вихрових струмів на поверхні об'єкта контролю (ОК), а результат