

Міністерство освіти і науки України
Вінницький національний технічний університет

І. А. Кравчук, О. В. Бісікало

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ
ФОРМУВАННЯ МЕТАДАНИХ
ДЛЯ СИСТЕМ АВТОМАТИЗОВАНОГО
ДОКУМЕНТООБІГУ**

Монографія

Вінниця
ВНТУ
2016

УДК 004.822:004.91
ББК 32.97
К78

Рекомендовано до друку Вченою радою Вінницького національного технічного університету Міністерства освіти і науки України (протокол № 5 від 30 листопада 2015 р.)

Рецензенти:

А. Я. Кулик, доктор технічних наук, професор
А. М. Петух, доктор технічних наук, професор

Кравчук, І. А.

К78 Інформаційна технологія формування метаданих для систем автоматизованого документообігу : монографія / І. А. Кравчук, О. В. Бісікало. – Вінниця : ВНТУ, 2016. – 164 с.

ISBN 978-966-641-671-4

В монографії розглянуто задачу формування метаданих для систем автоматизованого документообігу. Запропоновано розвиток асоціативно-статистичної концепції отримання знань шляхом формалізації абстрактної системи обробки текстів та поняття мовного образу (МО). Розроблено метод морфологічного аналізу слів природномовних текстів, що базується на правилах, які визначені шляхом аналізу структури слова флективних мов. Вдосконалено метод визначення семантичної мережі ключових слів тексту. Запропоновано інформаційну технологію формування метаданих, що забезпечує покращення повноти і точності заповнення метаданих у форматі Dublin Core.

УДК 004.822:004.91
ББК 32.97

ISBN 978-966-641-671-4

© І. Кравчук, О. Бісікало, 2016

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ І ПОЗНАЧЕНЬ.....	6
ПЕРЕДМОВА	7
РОЗДІЛ 1 АНАЛІЗ СИСТЕМ АВТОМАТИЗОВАНОГО ДОКУМЕНТООБІГУ, МЕТОДІВ ФОРМУВАННЯ ТА СТАНДАРТІВ МЕТАДАНИХ.....	10
1.1 Функціональні та технологічні характеристики систем автоматизованого документообігу	10
1.2 Концепції, моделі та методи отримання знань з текстової інформації	12
1.2.1 Концептуальні підходи до отримання знань з текстової інформації	12
1.2.2 Методи морфологічного аналізу	14
1.2.3 Методи визначення ключових слів тексту	17
1.3 Міжнародні стандарти та категорії метаданих	19
1.3.1 Класифікація метаданих.....	19
1.3.2 Стандарти метаданих для систем електронного документообігу	20
1.3.3 Стандарти електронного навчання SCORM, IEEE-LOM та IMS	22
1.4 Аналіз технологічних засобів формування метаданих	26
1.5 Критерії інформаційних процесів формування метаданих	26
1.6 Вибір напрямку, мети та постановка завдань дослідження	31
РОЗДІЛ 2 РОЗРОБКА МЕТОДІВ ТА МОДЕЛЕЙ АВТОМАТИЗОВАНОГО ФОРМУВАННЯ МЕТАДАНИХ З ТЕКСТОВОЇ ІНФОРМАЦІЇ.....	33
2.1 Формалізація асоціативно-статистичної концепції отримання знань з текстової інформації та поняття мовного образу 33	
2.1.1 Концептуальна модель природного накопичення асоціацій між мовними образами	33
2.1.2 Модель словника мовних образів Thesaurus	39
2.2 Метод аналізу морфологічної структури слова на основі асоціативно-статистичної концепції	44

2.2.1 Контекстно-вільна граматики для розпізнавання морфем флексивних мов	44
2.2.2 Метод морфологічного аналізу через формування множини морфем.....	45
2.2.3 Аналіз морфологічної структури слова з кількома суфіксами та префіксами.....	48
2.3 Метод визначення семантичної мережі ключових слів тексту	50
2.4 Статистичний аналіз формальних ознак текстових документів.....	55
2.5 Висновки до розділу	63
РОЗДІЛ 3 ТЕХНОЛОГІЧНІ КОМПОНЕНТИ ФОРМУВАННЯ МЕТАДАНИХ ДЛЯ СИСТЕМ АВТОМАТИЗОВАНОГО ДОКУМЕНТООБІГУ	65
3.1 Загальна архітектура та опис інформаційної технології формування метаданих	65
3.2 Підсистема первинної обробки тексту та морфологічного аналізу слів.....	70
3.3 Підсистема визначення семантичної мережі ключових слів тексту	78
3.4 Підсистема накопичення знань та формування додаткових метаданих	83
3.5 Висновки до розділу	86
РОЗДІЛ 4 РЕАЛІЗАЦІЯ ТА ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ ІНФОРМАЦІЙНОЇ ТЕХНОЛОГІЇ ФОРМУВАННЯ МЕТАДАНИХ	87
4.1 Програмне забезпечення інформаційної технології формування метаданих	87
4.2 Тестування коректності формування метаданих документа.....	96
4.3 Тестування результатів визначення метаданих навчального пакета.....	97
4.4 Апробація та експериментальне дослідження ефективності отриманих результатів у порівнянні з існуючими технологіями.....	98

4.5 Висновки до розділу	112
ПІСЛЯМОВА.....	113
ЛІТЕРАТУРА	116
Додаток А Перелік мовних образів, що використовувались для тестування методу морфологічного аналізу.....	132
Додаток Б Результати морфологічного аналізу слів на основі бази морфем, отриманої в результаті обробки 126 мовних образів.....	137
Додаток В Тестовий навчальний матеріал «Структурне програмування»	138
Додаток Д Результати визначення ключових слів російськомовної статті суспільно-політичної тематики альтернативними ресурсами	144
Додаток Е Результати визначення ключових слів україномовної наукової статті лінгвістичної тематики альтернативними ресурсами	147
Додаток Ж Лістинг програмного забезпечення для морфологічного аналізу слів	150
Додаток К Лістинг програмного забезпечення для формування навчальних пакетів.....	155
Додаток Л Лістинг програмного забезпечення для формування метаданих документів.....	161

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ І ПОЗНАЧЕНЬ

ПМ – природномовний
ПМК – природномовна конструкція
САД – система автоматизованого документообігу
AICC – Aviation Industry CBT Committee
CAM – Content Aggregation Model
CMI – Computer Managed Instruction
CSS – Cascading Style Sheets
IEEE – Institute of Electrical and Electronics Engineers
IMS – Instructional Management System
LMS – Learning Management System
LOM – Learning Object Metadata
LTSC – Learning Technology Standards Committee
PIF – Package Interchange File
SCO – Shareable Content Object
SCORM – Shareable Content Object Reference Model
XML – Extensible Markup Language
XSLT – Extensible Stylesheet Language Transformations

ПЕРЕДМОВА

Сьогодні у все більшій кількості великих організацій та установ активно впроваджуються системи автоматизованого документообігу (САД), що надають інструменти для формування електронних документів і контролю версій, мають широкий спектр можливостей щодо поширення документів та інформації в межах всієї організації, засобів управління контентом.

В загальному документообігу текстові документи мають важливе значення завдяки високій інформативності та широкому використанню, наприклад, як об'єкта пошуку. Зрозуміло, що найкращу характеристику змісту текстового документа для формування відповідних метаданих може дати його автор або, принаймні, кваліфікований експерт з потрібної галузі. Але практика експлуатації сучасних глобальних та локальних інформаційних систем, у т. ч. САД, демонструє, що у переважній більшості текстових документів не заповнені саме ті метадані, значення яких є неформальними. Найбільш яскраво такий стан справ проявився внаслідок поширення Інтернет-технологій і необхідності підвищення релевантності інформаційного пошуку – досить часто пошукові системи знаходять інформацію та документи, що не відповідають запиту користувача. Вирішити цю проблему дозволяє ефективне формування метаданих, що характеризують властивості і зміст документів, оскільки в мережі знаходиться велика кількість ресурсів з недостатнім описом їхніх метаданих.

Метадані для опису документів в мережі Інтернет можуть бути подані за допомогою таких мов, як RDF (The Resource Description Framework), OWL (Web Ontology Language), KIF (Knowledge Interchange Format), OCML (Operational Conceptual Modeling Language), F-Logic тощо. Найбільш розповсюдженими сьогодні форматами є RDF та OWL. До інших відомих стандартів формування метаданих відносять Dublin Core та SCORM. Відомі інструменти для формування текстового контенту згідно зі стандартами надають, як правило, платні послуги та не забезпечують автоматизоване формування метаданих. Тому подальший розвиток САД неможливий без використання сучасних методів та інформаційних технологій формування метаданих на основі аналізу текстового змісту документів.

Тематиці формування метаданих для природномовних об'єктів присвячено багато публікацій вчених різних країн: Дж. Грінберг [107, 108], Дж. Гуфріда [105], Ф. С. Пенга [138], Х. Хена [82], С. Марінаї [130], С. Лоуренса [122], К. Сеймора [150] та інших. В цих роботах досліджуються підходи до лінгвістичного аналізу текстового матеріалу, зокрема навчального призначення, вирішуються питання синтезу метаданих природномовних об'єктів, в тому числі у значній кількості робіт описано визначення ключових слів. Значний теоретичний внесок у фундаментальні та прикладні складові комп'ютерної лінгвістики зробили вітчизняні дослідники: В. А. Широков, М. Ф. Бондаренко, Н. В. Шаронова, Д. В. Ланде, А. В. Анісімов, Н. Є. Карпіловська, Н. П. Дарчук, О. Є. Литвиненко, В. В. Пасічник. Завдяки цьому напрацьовано моделі та методи статистичної й лінгвістичної обробки текстів, зокрема, морфологічного, статистичного та, частково, семантичного аналізу української мови та синтезу природномовних об'єктів.

Проте необхідність істотних витрат часу на ручне заповнення метаданих у діючих САД обумовлена багатьма причинами. Відомі методи визначення ключових слів та анотування тексту не забезпечують достатньої релевантності результатів, а перед статистичними методами навчання на основі розмічених даних (зі вчителем) постають непереборні труднощі при розмаїтті авторів та текстів з різних предметних областей. Використання потужних лінгвістичних пакетів для задачі формування метаданих вимагає значних додаткових обчислювальних ресурсів для і так перевантажених обробкою інформації САД, причому залишається головна проблема аналізу тексту – багатозначність слів. З іншого боку, практично відсутні публікації щодо застосування нестандартних підходів до заповнення метаданих, що не потребують масштабних словників та довідників з експертними мовними знаннями, зокрема моделювання природних шляхів вивчення мови дитиною.

Існуючі проблеми та висока активність досліджень у цій сфері зумовлюють актуальність запропонованої роботи на тему створення інформаційної технології формування метаданих безпосередньо зі змісту природномовних об'єктів на основі асоціативно-статистичної концепції отримання знань, що дозволить покращити, за низкою кількісних та якісних показників, результати обробки текстової інформації системами автоматизованого документообігу.

Монографія, матеріал якої базується на роботах [4, 6–11, 13, 30, 34–40, 55, 64, 86–88, 153], складається з чотирьох розділів. У першому розділі описано основні характеристики САД, категорії метаданих і міжнародних стандартів метаданих, а також аналіз концепцій отримання знань з тексту, методи морфологічного аналізу та методи визначення ключових слів. Визначено критерії ефективності формування метаданих. Визначено проблему, мету і задачі дослідження.

У другому розділі розглянуто асоціативно-статистичну концепцію отримання знань, що базується на природному накопиченні асоціацій між образами та закріпленні рефлексів шляхом повторень. Введено контекстно-вільну граматику, що дозволяє розпізнавати морфеми у словах зі словника мовних образів. На основі введеної формальної теорії визначено комутативну напівгрупу конструкцій мовних образів, яка забезпечує накопичення зв'язків між елементами множини мовних образів, що відповідають ключовим словам тексту. Отримано формальні обмеження складових асоціативно-статистичного підходу у вигляді статистичних оцінок колекції текстових документів певної предметної області.

У третьому розділі запропоновано інформаційну технологію формування метаданих, яка базується на методах морфологічного аналізу слів та визначення семантичної мережі ключових слів тексту. Розроблено схеми, алгоритми та діаграми процесів відповідної інформаційної системи в цілому та окремих її підсистем. Розроблено підсистему накопичення знань на основі побудови словника мовних образів. Формалізовано процедури додаткового формування метаданих.

У четвертому розділі наведено опис розробленого програмного забезпечення складових компонент запропонованої інформаційної технології формування метаданих. Наведено результати проведених експериментальних досліджень інформаційної технології формування метаданих, що показують її переваги.

РОЗДІЛ 1

АНАЛІЗ СИСТЕМ АВТОМАТИЗОВАНОГО ДОКУМЕНТООБІГУ, МЕТОДІВ ФОРМУВАННЯ ТА СТАНДАРТІВ МЕТАДАНИХ

Розглянуто основні характеристики систем автоматизованого документообігу, проаналізовано підходи до отримання знань з природномовних текстів. Зроблено огляд існуючих підходів до морфологічного аналізу та визначення ключових слів тексту. На основі порівняльного огляду найбільш поширених стандартів метаданих обрано Dublin Core, а електронного навчання – SCORM. Критеріями ефективності формування метаданих визначено повноту та точність заповнення метаданих. Сформульовано задачі, які необхідно розв'язати для того, щоб розробити інформаційну технологію формування метаданих на основі асоціативно-статистичного підходу.

1.1 Функціональні та технологічні характеристики систем автоматизованого документообігу

Система автоматизованого документообігу (САД) являє собою організаційно-технічну систему, що забезпечує процес створення, управління доступом і поширення електронних документів в комп'ютерних мережах, контроль над потоками документів в організації.

Під документообігом розуміють рух документів в організації з моменту їхнього створення або отримання до завершення виконання або відправлення.

САД дозволяє компаніям скоротити час та фінансові затрати на оформлення документації та знизити ризики, пов'язані з людським фактором.

До основних принципів електронного документообігу відносять такі:

- однократна реєстрація документа, що дозволяє однозначно ідентифікувати документ в будь-який момент;
- можливість паралельного виконання операцій, що дозволяє скоротити час руху документів і підвищити оперативність їх виконання;

- безперервність руху документа, що дозволяє ідентифікувати відповідального за виконання в кожен момент часу життя документа (процесу);

- єдина (або погоджено розподілена) база документної інформації, що дозволяє унеможливити дублювання документів;

- ефективно організована система пошуку документів, що дозволяє знаходити документ, маючи мінімальну інформацію про нього;

- розвинена система звітності за різними статусами і атрибутами документів, що дозволяє контролювати рух документів у процесах документообігу і приймати управлінські рішення, ґрунтуючись на даних із звітів.

САД поділяються на 3 категорії:

1. Універсальні САД – САД, що мають стандартний набір функцій.

2. Індивідуально розроблені САД – максимально персоніфіковані системи, що повністю відповідають потребам конкретної організації.

3. Комбіновані САД – базова платформа, до якої розробляються необхідні додаткові модулі.

До поширених в Україні САД відносять нижченаведені:

- система електронного документообігу та автоматизації бізнес-процесів Megapolis.DocNet,

- система електронного документообігу «MasterDoc»,

- система електронного документообігу «ДІЛО»,

- програмно-методичний комплекс класу управління документами (Enterprise Content Management) і бізнес-процесами (Business Process & Workflow Management) – ІНТАЛЄВ: Корпоративні документи і процеси,

- система електронного документообігу FossDoc,

- система електронного документообігу АСКОД.

САД реалізує такі функції:

- управління життєвим циклом (можливість управління документами протягом всього їх життєвого циклу – від розробки, редакції, погодження, затвердження до поширення та архівування);

- розподілена робота з документами (САД дозволяє співробітникам, що працюють в територіально віддалених підрозділах, спільно використовувати документи на локальних серверах, зберігаючи при цьому цілісність документів в масштабі всієї розподіленої організації);

- широкий набір засобів управління документами (реєстрація документів, розсилання, взяття на контроль виконання, контроль версій, повнотекстовий пошук, ведення контрольних журналів, робота з шаблонами, повідомлення про зміни тощо);
- редагування/узгодження (САД керує всім інтерактивним циклом редагування та внесення змін до документів аж до остаточного затвердження);
- багаторівневий контроль версій (дозволяє структурувати і зберігати всі чернетки документів);
- захист даних (засоби захисту і шифрування промислового рівня, повномасштабні засоби ідентифікації користувачів);
- підтримка широкого вибору клієнтів;
- масштабованість системи [23, 68].

Для якісної обробки електронної документації в САД використовуються метадані. Причому додавання метаданих – процес інтелектуальний і проводиться, в основному, людиною-оператором, тоді як обробка документів з використанням метаданих виконується системою автоматично.

Метадані можуть використовуватись для забезпечення таких функцій САД:

- пошук (пошук за метаданими, а не за повним змістом тексту документів, дає точніші результати та зменшує час пошуку);
- організація/доступ (метадані допомагають організувати відображення документів) [101, 157].

1.2 Концепції, моделі та методи отримання знань з текстової інформації

1.2.1 Концептуальні підходи до отримання знань з текстової інформації

Задача отримання корисної інформації з природномовного тексту являє собою ідентифікацію в ньому певних елементів. На вході системи отримання інформації маємо текст природною мовою, на виході – певні заповнені структури даних, що дозволяють проводити подальшу автоматизовану або ручну обробку інформації [13]. В залежності від обраного підходу до отримання знань з природномовних джерел у подальшому може бути застосована одна з основних моделей представ-

лення знань: продукційна, фреймова, семантична мережа, формальна логічна модель [17] або популярні останнім часом онтології.

Найбільш поширеними підходами до отримання знань з природномовних текстів є:

1. Синтаксично-семантичний підхід до отримання знань.

Цей підхід опирається на лінгвістичну модель, відповідно до якої основою семантичної структури висловлювання є так званий пропозиційний компонент плану змісту. Основним елементом у структурі пропозиції речення є предикати, що позначають відносини між об'єктами – обов'язковими учасниками ситуації, які називають аргументами або актантами предиката [25]. Такий підхід до отримання знань припускає виділення зі структури фрази її семантичного ядра – об'єктивного опису ситуації, та абстрагування від несуттєвих, суб'єктивних компонентів плану змісту. З цією метою використовується синтаксичний аналізатор тексту, що працює на основі знання загальних правил граматики мови, а також словник моделей керування, що описує для кожного предиката способи вираження в мові його аргументів [87];

2. Підхід до отримання знань, побудований на основі запитів до бази даних.

База даних являє собою сукупність певним чином упорядкованих відомостей про деякі об'єкти. Бази даних використовують для полегшення збору та зберігання інформації для опису елементів словника, що використовується для лінгвістичного аналізу текстової інформації з усіма характеристиками. На сьогодні в комп'ютерній лінгвістиці не існує спеціалізованих лексикографічних баз даних, проте такі бази даних, як ACCESS, FOX-Base, PARADOX, D-Base, можуть бути використані для ведення словникових картотек, а також для їхнього редагування [13, 87];

3. Підхід до отримання знань на основі онтологій.

В рамках комп'ютерної лінгвістики онтологія розуміється як система абстрактних понять, що існують тільки у свідомості людини, яка може бути виражена природною мовою (або засобами іншої системи символів). Лінгвістичний підхід до створення і дослідження онтологій заснований на вивченні природної мови (зокрема, семантики) і побудові онтологій на великих текстових масивах, так званих корпусах. Лінгвістичні онтології охоплюють більшість слів мови і одночасно мають онтологічну структуру, яка виявляється у відношеннях між по-

няттями. Для якісної обробки текстів за допомогою онтологій необхідно мати детальний опис проблемної області з множиною логічних зв'язків, які показують співвідношення між термінами області. Використання онтологій дозволяє представити природномовний текст у вигляді, придатному для автоматичної обробки [50, 87];

4. Асоціативно-статистичний підхід до отримання знань.

Недоліком попередніх підходів до отримання знань є необхідність значної попередньої роботи із залученням висококваліфікованих експертів.

Асоціативно-статистичною обробкою образної інформації варто назвати механізми з фізіологічною основою природного накопичення асоціацій між мовними образами та закріплення рефлексів шляхом повторень. Асоціативно-статистичний підхід до аналізу текстів приводить до побудови асоціативної семантичної мережі, яку складають пов'язані між собою слова тексту, що несуть основне семантичне навантаження. Для кожного поняття формується набір асоціативних зв'язків до/від інших понять, у сполученні з якими воно зустрічалося в тексті [8, 35].

Такий підхід використовується в програмних продуктах RCO і UCO for Oracle та в RCO і UCO Semantic Network (<http://www.rco.ru>, <http://uco.ua>), що використовуються для пошуку для Oracle та для Microsoft [13]. Найбільш зручною моделлю представлення знань для асоціативно-статистичного підходу є спеціалізована семантична мережа або природномовна онтологія.

Отже, асоціативно-статистичний підхід, що моделює відомі фізіологічні механізми, є найбільш близьким до природного способу отримання знань людиною, а тому може розглядатися як перспективний для використання в задачах аналізу текстової інформації.

1.2.2 Методи морфологічного аналізу

Морфологічний аналіз вхідних словоформ є початковим етапом аналізу природномовної конструкції (ПМК). Результати морфологічного аналізу є базовою основою для всіх наступних етапів автоматичного аналізу та синтезу тексту. Існує проблема інтеграції різнорівневих методів аналізу та синтезу ПМК з пріоритетом семантики предметної області з метою розв'язання задач пізнавальної діяльності [9].

В основі існуючих методів морфологічного аналізу лежать: поділ словоформи на основу та афікси з наступною перевіркою; використання інформації, що міститься в кінцевих сполученнях літер; створення універсальних математичних моделей тощо [11]. Класичну схему аналізу шляхом поділу словоформи на основу і афікси з наступною перевіркою на спільність закінчення з основою, що залишається, представляє морфологічний аналіз Г. Г. Белоногова. В основі його моделі [3] лежить флективний аналіз слів, що базується на розбитті слів російської мови на флективні класи.

До інших моделей морфологічного аналізу, які фокусуються на різних граматичних явищах або схемах обробки, належать моделі К. Коскенніємі [125], Ю. П. Шабанова-Кушнарєнко [76], Л. В. Найханової та І. С. Євдокимова [49], А. Гельбуха та Г. Сідорова [103], Р. Хаусера [110], А. Ван Ден Боша та В. Даелеманса [89].

До основних методів морфологічного аналізу відносять:

1. Методи на основі словника (з декларативною орієнтацією).

Для таких методів характерна наявність повного словника всіх можливих словоформ для кожного слова з прописаною морфологічною інформацією. При цьому кожна словоформа забезпечується повною та однозначною морфологічною інформацією, куди входять як постійні, так і змінні морфологічні параметри [19]. Завдання морфологічного аналізу в такому випадку зводиться до пошуку потрібної словоформи в словнику та переписування зі словника морфологічної інформації.

Перевагою таких методів є висока швидкість проведення морфологічного аналізу та точність результатів аналізу відносно всіх словоформ української мови, що внесені до словника, а недоліком – великі затрати пам'яті, пов'язані з тим, що у кожного слова досить велика кількість словоформ. Інформація, що зберігається в словниках, є надлишковою. Ще однією проблемою є морфологічний аналіз слів, що не містяться в словнику словоформ [9].

До таких методів можна віднести підхід з використанням мінімального ациклічного автомата з кінцевим числом станів (Б. В. Ватсон, С. Міхов) [114], підхід з використанням дворівневого компілятора правил (Л. Картунен, К. Р. Біслі) [123], підхід з використанням принципів побудови зваженого перетворювача з кінцевим числом станів (М. Мохрі, Ф. Перейра, М. Рілі) [135];

2. Методи на основі правил (з процедурною орієнтацією).

Для таких методів характерним є те, що модуль морфологічного аналізу містить набір правил морфологічних перетворень. Кожне слово розділяється на основу і афікс (закінчення і, можливо, суфікс); словник містить тільки основи слів разом з посиланнями на відповідні рядки в таблиці можливих афіксів. Основний критерій при розбитті слова на основу і афікс – основа повинна залишатися незмінною в усіх можливих словоформах цього слова [20].

Перевагою процедурних методів є зменшення обсягу інформації, яку необхідно зберігати для виконання морфологічного розбору слова. Велика кількість слів української мови має одні й ті ж афікси, які не потрібно дублювати при зберіганні словника афіксів. Саме тому обсяг словника, що використовується при процедурних методах, є значно меншим, ніж обсяг повного словника словоформ, що використовується в декларативних методах. Але при цьому процедурний метод передбачає попередню систематизацію морфологічних знань про природну мову і розробку алгоритмів присвоєння морфологічної інформації окремій словоформі [9].

До процедурної групи можна віднести моделі І. Ножова [51], К. Офлайзера, С. Ніренбурга та М. Мак-Шейн [137], П. Шона та Д. Джурафського [147] тощо;

3. Комбіновані методи.

Такі методи використовують словник і набір правил для обробки слів, що не містяться в словнику. За цим методом на першому етапі здійснюється пошук у словнику словоформ, як при декларативному методі, і у випадку успішного пошуку аналіз на цьому завершується. В протилежному випадку використовується словник основ і процедурний метод аналізу [49].

Такий підхід до морфологічного аналізу використовує, наприклад, пошукова система Yandex для системи пошуку в мережі Інтернет з відгадуванням невідомих слів за допомогою використання словника [149]. Цей підхід, що його запропонував І. Сегаловіч, використовує за словник набори можливих коренів та набори можливих суфіксів. Границя між коренем та суфіксом береться з вхідного файлу, якщо слово там наведене, або обчислюється автоматично як довжина всіх словоформ заданої парадигми.

Зрозуміло, що метод, який не потребує для виконання морфологічного аналізу словника з заданими наперед словоформами, буде мати перевагу над існуючими методами з точки зору трудомісткості процесів його розробки та взаємодії з іншими задачами аналізу тексту.

Отже, існує потреба у розробці методів морфологічного аналізу, що були б найбільш близькими до природного способу отримання знань людиною з огляду на розуміння змісту ПМК і не потребували масштабних початкових знань експертів-лінгвістів з обраної мови [9]. Потенційно асоціативно-статистичний підхід до отримання знань з природномовних текстів може базуватися лише на конструктивних правилах формування слів флективної мови.

1.2.3 Методи визначення ключових слів тексту

Мета визначення ключових слів полягає у знаходженні невеликого набору слів, які прийнято називати ключовими (словами, словосполученнями, фразами) та які можуть охарактеризувати зміст документа [112].

Підходи до визначення ключових слів тексту діляться на 3 категорії:

- статистичні,
- лінгвістичні,
- змішані.

Статистичні методи, зазвичай, зосереджені на таких нелінгвістичних характеристиках тексту, як частота повторення терміна, зворотна частота документа, а також позиції ключового слова. Статистична інформація про слова може бути використана для ідентифікації ключових слів у документі. Дж. Д. Коен використовує статистичну інформацію з N-грам [91]. Інші статистичні методи використовують частоту вживання слів, TF-IDF (TF – term frequency, IDF – inverse document frequency), спільну появу слів [131] тощо. До таких методів також відносять [95, 131, 145]. Перевагами чисто статистичних методів є їх простота у використанні, обмежені вимоги до обчислень і, як правило, хороші результати.

Лінгвістичні методи звертають увагу на такі лінгвістичні особливості, як частина мови, синтаксична структура та семантичні властивості, а іноді відіграють функцію фільтрів для «поганих ключових слів» [112].

Л. Плас та ін. [139] використовують для оцінювання два лексичні ресурси: електронний словник EDR [133] і WordNet Прінстонського університету, у вільному доступі [163]. Обидва надають масштабні словники, що містять семантичні відношення і зв'язки, такі як IS-A та PART-OF відношення і концепції багатозначності. А. Халт [112] розглядає кілька різних методів використання лінгвістики у визначенні ключових слів. Слова, що вважаються ключовими, будуть перевірені за трьома ознаками: частота документа, частота колекції та відносна положення його першої появи в документі. До таких підходів також відносять [102, 83].

Змішані методи поєднують в собі деякі лінгвістичні методи з такими загальними статистичними методами, як наприклад, частота появи терміна. Ці підходи можуть також використовувати такі евристичні знання, як позиція та довжина слова, HTML теги навколо слів тощо [119].

Іноді виділяють ще підходи, що пов'язані з машинним навчанням. Визначення ключових слів може розглядатися як контрольоване навчання за прикладами. Механізм машинного навчання працює таким чином. По-перше, система забезпечується набором навчальних документів, кожен з яких має набір визначених людиною ключових слів. Тоді отримані знання застосовуються для визначення ключових слів з нових документів. Алгоритм визначення ключових фраз (KEA) [124] використовує технологію машинного навчання та формулу Байеса для отримання технічних фраз.

Й. Сузукі та ін. [154] використовують технології обробки мови (голосу) для визначення ключових слів з радіоновин за допомогою енциклопедії та газетних статей як вказівки відповідності.

Основним недоліком існуючих методів автоматизованого визначення ключових слів є суттєво нижчий рівень релевантності отриманих ними наборів слів у порівнянні з авторськими або експертними варіантами. Перспективним шляхом до усунення або, принаймні, зменшення впливу цього недоліку є застосування асоціативно-статистичного підходу до отримання знань з природномовних текстів, який дозволяє отримати додаткову чисельну інформацію щодо ваги кожного слова з семантичної мережі ключових слів.

1.3 Міжнародні стандарти та категорії метаданих

1.3.1 Класифікація метаданих

Метадані – це структуровані дані, що є характеристиками певних описаних сутностей для цілей їхньої ідентифікації, пошуку, оцінювання, управління ними [156].

Роль метаданих, які орієнтовані, насамперед, на користувачів і авторів електронних інформаційних ресурсів в розподілених інформаційних системах, полягає в:

- управлінні життєвим циклом інформаційних ресурсів (процесами створення, використання і зберігання цифрових документів);
- наданні можливостей більш точного і повного виявлення необхідних ресурсів;
- забезпеченні гнучких і різноманітних механізмів відбору та подання інформації відповідно до вимог користувача (пошукові запити) [46].

У наш час існують різні класифікації метаданих, які відрізняються між собою, головним чином, ступенем деталізації.

За класифікацією, поданою в [74], метадані діляться на дві великі групи:

- адміністративні метадані,
- метадані опису змісту.

Адміністративні метадані складають різні групи і характеризуються значним різноманіттям. Основна група адміністративних метаданих призначена для того, щоб забезпечувати чітку інформацію про обраний об'єкт, що містить авторизацію, аутентифікацію, управління авторськими правами, доступом, а також служить для ідентифікації та категоризації інформаційних об'єктів у межах спеціальної колекції або організації. Друга група може являти собою схеми зберігання даних у базах даних, схеми розподілених баз даних тощо. Третя група адміністративних метаданих може використовуватися для позиціонування заданого інформаційного ресурсу в контексті групи подібних документів, інформаційно-пошукової системи, предметної області тощо.

Метадані опису змісту охоплюють опис всіх аспектів заданого інформаційного ресурсу (об'єкта) як окремої сутності. Іноді їх додатково поділяють на структурні та описові [46, 74].

Метадані складаються з елементів, об'єднаних в набори. Такі набори розробляються з різними цілями (наприклад, для опису різних інформаційних об'єктів) різними організаціями. У тому випадку, якщо набір елементів метаданих розглядається і приймається відповідною уповноваженою організацією, наприклад, ISO (International Standard Organisation), то він стає офіційним стандартом метаданих.

Формат метаданих – стандарт, призначений для формального опису деякої категорії ресурсів (об'єктів, сутностей тощо). Такий стандарт, зазвичай, містить в собі набір полів (атрибутів, властивостей, елементів метаданих), що дозволяють характеризувати цей об'єкт.

Формати можна класифікувати за:

- охопленням типів описуваних ресурсів;
- деталізацією опису ресурсів і потужності структури елементів метаданих;
- предметними областями;
- цілями розробки та використання формату метаданих.

Формати метаданих нерідко розробляються міжнародними організаціями або консорціумами, що охоплюють зацікавлені в появі стандарту державні організації та приватні компанії. Розроблений формат часто закріплюється як стандарт в одній або декількох організаціях, що займаються розробкою і прийняттям стандартів (наприклад W3C, ISO, ANSI тощо).

1.3.2 Стандарти метаданих для систем електронного документообігу

Метадані САД, як правило, зберігаються для кожного документа. САД можуть отримувати метадані з документа автоматично або звертатися до користувача з пропозицією додати метадані.

Метадані САД описуються у двох стандартах:

- 1) ISO 15836:2009 (Dublin Core);
- 2) ISO 23081.

Стандарт ISO 15836:2009 (Dublin Core)

Простий набір елементів метаданих стандарту Dublin Core складається з 15 елементів для опису матеріалів:

- 1) назва (назва ресурсів, яка створена автором або видавцем);
- 2) автор (люди чи організації, що сприяли створенню ресурсу);

ЛІТЕРАТУРА

1. Бардачов Ю. М. Дискретна математика : підручник / Ю. М. Бардачов, Н. А. Соколова, В. Є. Ходаков ; за ред. В. Є. Ходакова. – К. : Вища школа, 2002. – 287 с. – ISBN 966-642-090-2.
2. Бахвалов Н. С. Численные методы / Н. С. Бахвалов, Н. П. Жидков, Г. Г. Кобельков. – М. : Лаборатория Базовых Знаний, 2000. – 632 с.
3. Белоногов Г. Г. Компьютерная лингвистика и перспективные информационные технологии. – М. : Русский мир, 2004. – 248 с.
4. Бисикало О. В. Модель формальной теории в виде коммутативной полугруппы образных конструкций / О. В. Бисикало, И. А. Кравчук, А. А. Кириленко // Проблемы машиностроения. – Харьков, 2013. – Вип. 6 (16). – С. 24–30.
5. Бісікало О. В. Алгебраїчна модель лінгвістичного процесора / О. В. Бісікало // Информационные технологии в управлении сложными системами : сб. докладов и тезисов междунар. научн.-практ. конф., (Днепропетровск, 22–23 мая 2008 г.). – Днепропетровск : ИТМ НАНУ и НКАУ, 2008. – С. 23–24. – ISBN 978-966-02-4754-3.
6. Бісікало О. В. Аналіз морфологічної структури слова на основі асоціативно-статистичного підходу / О. В. Бісікало, І. А. Кравчук // Вісник Вінницького політехнічного інституту. – 2011. – № 4. – С. 134–136.
7. Бісікало О. В. Аналіз морфологічної структури слова на основі асоціативно-статистичного підходу [Електронний ресурс] / О. В. Бісікало, І. А. Кравчук // Контроль і управління в складних системах: Х Міжнар. наук. конф., 19–21 жовтня 2010 р. : тези доповідей. – Вінниця, 2010. – Режим доступу : http://mccs.vntu.edu.ua/mccs2010/materials/subsection_1.2.pdf.
8. Бісікало О. В. Асоціативно-статистичний метод морфологічного аналізу слів / О. В. Бісікало, І. А. Кравчук // Інформаційні технології та комп'ютерна інженерія: І Міжнар. наук.-практ. конф., 19–21 травня 2010 р. : тези доповідей. – Вінниця, 2010. – С. 110–111.
9. Бісікало О. В. Використання морфологічного аналізу в задачах дистанційної освіти / О. В. Бісікало, І. А. Кравчук // Системні технології. – 2011. – № 4. – С. 41–47.
10. Бісікало О. В. Концептуальні алгоритми виокремлення морфем для реалізації інформаційної технології обробки природномовних тек-

стів / О. В. Бісікало, І. А. Кравчук // Інформаційні технології та комп'ютерна інженерія: II Міжнар. наук.-практ. конф., 26–27 травня 2011 р. : тези доповідей. – Харків, 2011. – С. 11–12.

11. Бісікало О. В. Концептуальні алгоритми виокремлення морфем для реалізації інформаційної технології обробки природномовних текстів / О. В. Бісікало, І. А. Кравчук // Системи обробки інформації. – 2011. – № 3. – С. 7–9.

12. Бісікало О. В. Концептуальні основи моделювання образного мислення людини / О. В. Бісікало. – Вінниця : ПП Балюк І. Б., ВДАУ, 2009. – 163 с.

13. Бісікало О. В. Методи отримання знань з текстів для подальшого використання в задачах лінгвістичної обробки / О. В. Бісікало, І. А. Кравчук // Інтернет–Освіта–Наука–2012: VIII Міжнар. наук.-практ. конф., 1–5 жовтня 2012 р. : збірник праць. – Вінниця, 2012. – С. 84.

14. Бісікало О. В. Формальні методи образного аналізу та синтезу природно-мовних конструкцій : монографія / О. В. Бісікало. – Вінниця : ВНТУ, 2013. – 324 с.

15. Бочаров П. П. Теория вероятностей. Математическая статистика / П. П. Бочаров, А. В. Печинкин. – М. : Наука, 1998. – 325 с.

16. Воронина И. Е. Программные средства выявления семантического поля слов / И. Е. Воронина, А. А. Кретов, О. С. Титова // Вестн. Воронеж. гос. ун-та. Серия Системный анализ и информационные технологии. – 2008. – № 2. – С. 111–122.

17. Гаврилова Т. А. Базы знаний интеллектуальных систем // Т. А. Гаврилова, В. Ф. Хорошевский. – СПб. : Питер, 2001. – 384 с.

18. Граматичний словник української літературної мови. Словозміна / [Відп. ред. Н. Ф. Клименко]. – К. : Видавничий Дім Дмитра Бураго, 2011. – 760 с.

19. Губин М. В. Влияние морфологического анализа на качество информационного поиска / М. В. Губин, А. Б. Морозов // Труды RCDL–2006. – 2006. – С. 224–228.

20. Дорохина Г. В. Модуль морфологического анализа слов русского языка / Г. В. Дорохина, А. П. Павлюкова // Искусственный интеллект. – 2004. – № 3. – С. 636–642.

21. Дубовой В. М. Моделювання систем контролю та керування : навч. посіб. / В. М. Дубовой. – Вінниця : ВНТУ, 2005. – 175 с.

22. Жуков Виктор. Создание объектов контента SCORM [Электронный ресурс] / В. Жуков. – Режим доступа : http://kibi.ru/_media/notes/scorm.ppt

23. Золотарьова І. О. Автоматизація документообігу : навч. посіб. / І. О. Золотарьова, Р. К. Бутова. – Харків : Вид. ХНЕУ, 2008. – 156 с.

24. Евромайдан как спецоперация. Кто инициировал беспорядки на Украине? [Электронный ресурс] // АиФ. – 2013. – № 49, 04/12/2013. – Режим доступа : <http://www.aif.ru/euromaidan/prediction/1036380>.

25. Ермаков А. Е. Неполный синтаксический анализ текста в информационно-поисковых системах / А. Е. Ермаков // Компьютерная лингвистика и интеллектуальные технологии : труды Международного семинара Диалог'2002. – М. : Наука. – 2002. – Т. 2. – С. 180–185.

26. Каневский Е. А. Семантико-синтаксический анализатор SemSin [Электронный ресурс] / Е. А. Каневский, К. К. Боярский // Международная конференция по компьютерной лингвистике «Диалог–2012». – Режим доступа : <http://www.dialog-21.ru/digests/dialog2012/materials/pdf/Kanevsky.pdf>.

27. Карпіловська Є. А. Вступ до прикладної лінгвістики: комп'ютерна лінгвістика / Є. А. Карпіловська. – Донецьк : Юго-Восток, 2006. – 188 с.

28. Карпіловська Є. А. Стратифікація лексики в тлумачному словнику // Критика у слов'янській лексикографії. Паламарчуку Леоніду Сидоровичу / Ін-т укр. Мови НАН України: упорядн. : І. С. Гнатюк. – К. : КММ, 2012. – С.133–139.

29. Кветний Р. Н. Методи комп'ютерних обчислень : навч. посіб. / Р. Н. Кветний. – Вінниця : ВДТУ, 2001. – 148 с.

30. Кветний Р. Н. Морфологічний аналіз слова на основі асоціативно-статистичного підходу / Р. Н. Кветний, О. В. Бісікало, І. А. Кравчук // Вісник Черкаського державного технологічного університету. – 2010. – № 3. – С. 132–135.

31. Келеберда И. Н. Использование мультиагентного онтологического подхода к созданию распределенных систем дистанционного обучения / И. Н. Келеберда, Н. С. Лесная, В. Б. Репка // Educational Technology & Society, 2004. – № 7(2). – С. 190–205.

32. Клименко Н. Ф. Словотвірна морфеміка сучасної української літературної мови : монографія / Н. Ф. Клименко, Є. А. Карпіловська – К., 1998. – 161 с.

33. Колмогоров А. Н. Основные понятия теории вероятностей / А. Н. Колмогоров. – 2-е изд. – М. : Наука, 1974. – 120 с.

34. Кравчук І. А. Алгоритм вирішення проблеми багатосуфіксальності та багатопрефіксальності при морфологічному аналізі слів / І. А. Кравчук // Контроль і управління в складних системах (КУСС–2014) : міжнар. конф., 14–16 жовтня 2014 р. : тези доповідей. – Вінниця : ВНТУ, 2014. – С. 98.

35. Кравчук І. А. Асоціативно-статистичний підхід до отримання знань [Електронний ресурс] / І. А. Кравчук // XXXIX Науково-технічна конференція професорсько-викладацького складу, співробітників та студентів ВНТУ з участю працівників науково-дослідних організацій та інженерно-технічних працівників підприємств м. Вінниця та області, 9–12 березня 2010 р. : тези доповідей. – Вінниця, 2010. – Режим доступу : <http://conf.vntu.edu.ua/allvntu/2010/inaeksu/txt/kravchuk.pdf>.

36. Кравчук І. А. Використання морфологічного аналізу для покращення подання навчального матеріалу для дистанційних курсів / І. А. Кравчук // Електронні ресурси та технології: створення, використання, доступ : міжнар. наук.-практ. Інтернет-конференція, 10–17 травня 2011 р. : збірник матеріалів конф. – Вінниця, 2011. – С. 105–108.

37. Кравчук І. А. Порівняльний огляд стандартів електронного навчання / І. А. Кравчук, О. В. Бісікало // Інформаційні технології в металургії та машинобудуванні (ІТММ–2013) : наук.-техн. конф., 26–28 березня 2013 р. : матеріали конф. – Дніпропетровськ : НМетАУ, 2013. – С. 131–132.

38. Кравчук І. А. Порівняльний огляд стандартів електронного навчання / І. А. Кравчук, О. В. Бісікало // Системні технології. – Дніпропетровськ, 2013. – Вип. 4 (87). – С. 146–156.

39. Кравчук І. А. Створення «розумного» контенту електронного навчання / І. А. Кравчук // Технологический аудит и резервы производства. – 2012. – № 5/2 (7). – С. 9–10.

40. Кравчук І. А. Формальний аналіз морфемних складових тезауруса мовних образів [Електронний ресурс] / І. А. Кравчук // XL регіональна науково-технічна конференція професорсько-викладацького складу, співробітників та студентів ВНТУ з участю працівників науково-дослідних організацій та інженерно-технічних працівників підприємств м. Вінниця та області, 9–11 березня 2011 р. : тези доповідей. –

Вінниця, 2011. – Режим доступу : <http://conf.vntu.edu.ua/allvntu/2011/inaeksu/txt/kravchuk.pdf>.

41. Кригін М. Ю. Дослідження інформаційно-статистичних властивостей українського тексту / М. Ю. Кригін, В. А. Широков // Математичні машини і системи. – 2000. – № 1. – С. 120–127.

42. Крылов С. А. Некоторые уточнения к определениям понятий словоформы и лексемы / С. А. Крылов // Семиотика и информатика. – 1982. – Вып. 19. – С. 118–136.

43. Кузьмин И. В. Основы теории информации и кодирования : учеб. для вузов по спец. «Автоматика и телемеханика» и «Прикл. Математика» / И. В. Кузьмин, В. А. Кедрус. – 2-е изд., перераб. и доп. – К. : Вища школа, 1986. – 238 с.

44. Литневская Е. И. Русский язык: краткий теоретический курс для школьников [Электронный ресурс] / Е. И. Литневская. – Режим доступу : <http://www.gramota.ru/book/litnevskaya.php?part2.htm>.

45. Лінгвістичний портал Mova.info [Електронний ресурс] / Дарчук Н. П. та ін. // Корпус текстів української мови. – Режим доступу : http://www.mova.info/corpus_sint.aspx?id=23346.

46. Мазов Н. А. Метаданные и их роль в распределённых информационных системах на основе использования протокола Z39.50 / Н. А. Мазов О. Л. Жижимов // Библиосфера. – 2006. – № 2. – С. 51–59.

47. Манако А. Ф. Моделі агрегування об'єктів навчального контенту на базі систем інформаційних технологій / А. Ф. Манако // Проблеми програмування. – 2004. – № 2, 3. – С. 587–594.

48. Методологія моделювання процесів образного мислення у відповідності до задач пізнавальної та мовленнєвої діяльності людини: звіт про НДР (заключний) / Вінницький національний технічний університет; керівн. Р. Н. Кветний; викон. О. В. Бісікало, С. Г. Кривогубченко, І. А. Кравчук, Г. О. Кириленко. – Вінниця, 2013. – 201 с. – № д. р. 0111U001115. – Інв. № 0214U002854.

49. Найханова Л. В. Методы и алгоритмы трансляции естественно-языковых запросов к базе данных в SQL-запросы / Л. В. Найханова, И. С. Евдокимова. – Улан-Удэ : Изд-во ВСГТУ, 2004. – 148 с.

50. Никоненко А. А. Обзор баз знаний онтологического типа / А. А. Никоненко // Искусственный интеллект. – 2009. – № 4. – С. 208–219.

51. Ножов И. Процессор автоматизированного морфологического анализа без словаря. Деревья и корреляция / И. Ножов // Диалог'2000 : труды конференции, 2000. – Протвино, 2000. – Т. 2. – С. 284–290.

52. Обработка знаний: технологии анализа и поиска текстовой информации [Электронный ресурс] / Компания Гарант-Парк-Интернет. – Режим доступа : <http://www.olap.ru/basic/knowledge.asp>.

53. Освітній портал™ – освіта в Україні, освіта за кордоном [Електронний ресурс] – Режим доступа : <http://www.osvita.org.ua>.

54. Основы дискретной математики : учебник / [Ю. В. Капитонова, С. Л. Кривий, О. А. Летичевський та ін.]. – К. : Наукова думка, 2002. – 579 с. – ISBN 966-00-0622-5.

55. Патент України на корисну модель № 75250, МПК (2006.01) G06F 17/27. Спосіб морфологічного аналізу на основі асоціативно-статистичного підходу до отримання знань / О. В. Бісікало, І. А. Кравчук; заявник та власник патенту Вінницький національний технічний університет. – № u201205834 : заявл. 14.05.2012 ; опубл. 26.11.2012, Бюл. № 22. – 4 с.

56. Песин Л. Гибкая поддержка метаданных учебных объектов, опыт системы ALE / Л. Песин, М. Шпехт // Educational Technology & Society. – Vol. 6(4). – 2003. – Pp. 186–193.

57. Побудова методології моделювання функцій образного мислення людини : звіт про НДР «Методологія моделювання процесів образного мислення у відповідності до задач пізнавальної та мовленнєвої діяльності людини» (проміжний) / Вінницький національний технічний університет; керівн. Р. Н. Кветний; викон. О. В. Бісікало, І. А. Кравчук, Г. О. Кириленко. – Вінниця, 2012. – 96 с. – № д. р. 0111U001115. – Інв. № 0213U002624.

58. Побудова теорії розуміння сенсу образних конструкцій природномовного контенту : звіт про НДР «Методологія моделювання процесів образного мислення у відповідності до задач пізнавальної та мовленнєвої діяльності людини» (проміжний) / Вінницький національний технічний університет; керівн. Р. Н. Кветний; викон. О. В. Бісікало, І. А. Кравчук, Є. М. Бочков, І. П. Борщова, Г. О. Кириленко. – Вінниця, 2011. – 85 с. – № д. р. 0111U001115.

59. Подгорнов А. Object Orchestrator – приложение для создания и поддержки учебного материала [Электронный ресурс] / А. Подгорнов, К. Сеница // Educational Technology & Society. – 2005. – № 8(3). –

С. 319–324. – Режим доступу : http://ifets.ieee.org/russian/depositary/v8_i3/html/s0.html.

60. Реформатский А. А. Введение в языковедение / А. А. Реформатский. – М. : Аспект Пресс, 1996. – 536 с.

61. Розенберг Г. С. Поль Жаккар и сходство экологических объектов / Г.С. Розенберг // Самарская Лука : проблемы региональной и глобальной экологии. – 2012. – Вып. № 1. – Т. 21. – С. 190–202.

62. Розробка програмної оболонки для створення електронних підручників з розширеними дидактичними можливостями : звіт про НДР (закл.) / Вінницький національний технічний університет; керівн. Р. Н. Кветний; викон. О. В. Бісікало, І. В. Богач, І. А. Кравчук. – Вінниця, 2010. – 40 с. – № д. р. 0110U004301. – Інв. № 0210U007657.

63. Саак А. Э. Информационные технологии управления : учебник для вузов / А. Э. Саак. – СПб. : Питер, 2009. – 320 с.

64. Свідोцтво про реєстрацію авторського права на твір № 57089. Комп'ютерна програма для створення навчального пакету згідно стандарту SCORM з автоматизованим визначенням метаданих / І. А. Кравчук. – № 57475 : заявл. 03.09.2014 ; опубл. 20.10.2014.

65. Словник афіксальних морфем української мови / [Н. Ф. Клименко, Є. А. Карпіловська, В. С. Карпіловський, Т. І. Недозим]. – К., 1998. – 441 с.

66. Столл Р. Множества. Логика. Аксиоматические теории / Р. Столл. – М. : Просвещение, 1968. – 231 с.

67. Структурне програмування [Електронний ресурс] / Підручник «Програмна інженерія» // Programs Factory Site, 2014. – Режим доступу : <http://www.programsfactory.univ.kiev.ua/content/books/2/62>.

68. Сысоева Л. В. Разработка концепции общекорпоративной системы управления документами / Л. В. Сысоева // Секретарское дело. – 2006. – № 7. – С. 30–35.

69. Тельнов Ю. Ф. Разработка инновационных образовательных технологий на основе модели с использованием SCORM-спецификаций / Ю. Ф. Тельнов, О. В. Рогозин // Открытое образование. – М., 2009. – № 4. – С. 37–46.

70. Теорія образного аналізу тексту : звіт про НДР «Інтелектуальна інформаційна технологія образного аналізу тексту та синтезу інтегрованої бази знань природномовного контенту» (проміжний) / Вінницький національний технічний університет; керівн. Р. Н. Кветний; ви-

кон. С. Г. Кривогубченко, О. В. Бісікало, І. А. Кравчук. – Вінниця, 2014. – 86 с. – № д. р. 0114U003462.

71. Фрумкина Р. М. Статистические методы изучения лексики / Р. М. Фрумкина. – М. : Наука, 1964. – 115 с.

72. Хомский Н. Аспекты теории синтаксиса / Н. Хомский ; пер. с англ. под ред. В. А. Звегинцева. – М. : Издательство Московского университета, 1972. – 258 с.

73. Хомский Н. Язык и мышление / Н. Хомский ; пер. с англ. Б. Ю. Городецкого. – М. : Издательство Московского университета, 1972. – 121 с.

74. Хохлов Ю. Е. Обзор форматов метаданных [Электронный ресурс] / Ю. Е. Хохлов, С. А. Арнаутов // Институт развития информационного общества. Портал «Российские электронные библиотеки». – Режим доступа : http://www.elbib.ru/index.phtml?env_page=methodology/metadata/md_review/md_review.html.

75. Частотні словники [Електронний ресурс] / Лінгвістичний портал Mova.info. – Режим доступу : <http://www.mova.info/Page.aspx?l1=57>.

76. Шабанов-Кушнарєнко Ю. П. Теорія інтелекта. Математическіє средства / Ю. П. Шабанов-Кушнарєнко. – Харків : Изд-во Харківського університета, 1984. – 144 с.

77. Широков В. А. Информационный подход в моделировании лексикографических систем [Электронный ресурс] / В. А. Широков // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества : материалы 4-й Междунар. конф. «Крым-97». – 1997. – Т. 2. – Режим доступа : <http://www.nbu.gov.ua/articles/crimea/1997/DOC/sem11/s11doc8.html>.

78. A multi-layer, metadata driven approach enhancing SCORM to support adaptive and mobile learning objects [Rim Drira, Mona Laroussi, Alain Derycke, Henda Ben Ghezala] // Proceedings of the 3rd WSEAS/IASME International Conference on Engineering Education, July 11–13, 2006. – Vouliagmeni, Greece, 2006. – P. 245–250.

79. Allert H. Situated Models and Metadata for Learning Management [Електронний ресурс] / Heidrun Allert, Christoph Richter, Wolfgang Nejdl. – Proceedings of I-KNOW '03, July 2–4, 2003. – Graz, Austria, 2003. – Режим доступа : http://i-know.tugraz.at/wp-content/uploads/2008/11/31_situated-models.pdf.

80. Andrade M. A. Automatic extraction of keywords from scientific text: application to the knowledge domain of protein families / M. A. Andrade, A. Valencia // *Bioinformatics*. – 1998. – V. 14, Issue 7. – P. 600–607.

81. Articulate [Электронный ресурс] // Articulate Systems Official Site, 2014. – Режим доступа : <https://www.articulate.com/>.

82. Automatic document metadata extraction using support vector machines / [H. Han, L. C. Giles, E. Manovogly and others] // *Proceedings of 2003 Joint Conference on Digital Libraries*, 27–31 May 2003. – Houston, 2003. – P. 37–48.

83. Automatic Keyword Extraction from SpokenText / [L. Van der Plas, V. Pallotta, M. Rajman, H. Ghorbel] // *In Proceeding of LREC*, 2004. – P. 2205–2208.

84. Automatic metadata extraction and indexing for reusing e-learning multimedia objects / [Paolo Bolettieri, Fabrizio Falchi, Claudio Gennaro, Fausto Rabitti] // *Proceeding of MS '07: Workshop on multimedia information retrieval on The many faces of multimedia semantics*. – ACM New York, NY, USA, 2007. – P. 21–28.

85. Bianco A. M. Standards for e-learning [Электронный ресурс] / A. M. Bianco, M. De. Marsico, M. Temperini. – Режим доступа : http://www2.tisip.no/quis/public_files/wp5-standards-for-elearning.pdf.

86. Bisikalo O. Automation of course content construction / O. Bisikalo, I. Kravchuk // *Сборник научных трудов SWorld*. – Иваново : МАРКОВА АД, 2013. – Вып. 3. – Т. 6. – С. 33–38.

87. Bisikalo Oleg. Methods of obtaining knowledge from natural language texts / Oleg Bisikalo, Irina Kravchuk // *Perspektywiczne opracowania są nauką i technikami* – 2012: Materiały VIII Międzynarodowej naukowo-praktycznej konferencji, 07–15 listopada 2012 roku, Vol. 19. – Przemyśl, 2012. – Str. 3–5.

88. Bisikalo Oleg. Support for Interactive Features of E-learning Content Based on the Formal Theory / O. Bisikalo, I. Kravchuk // *Journal of Applied Computer Science & Mathematics*. – Romania, Suceava, 2014. – No. 16 (8). – P. 9–15.

89. Bosch A. van den. Memory-Based Morphological Analysis [Электронный ресурс] / A. van den Bosch, W. Daelemans // *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on*

Computational Linguistics. – Режим доступа : <http://acl.ldc.upenn.edu/P/P99/P99-1037.pdf>.

90. Chomsky N. Syntactic Structures / N. Chomsky. – The Hague : Mouton, 1957. – 117 p.

91. Cohen J. D. Language and domain-independent automatic indexing terms for abstracting / J. D. Cohen // Journal of the American Society for Information Science, 1995. – P. 162–174.

92. Collier G. E-Learning Interoperability Standards [Электронный ресурс] / G. Collier, R. Robson. – USA: Sun Microsystems White Paper, 2002. – 25 p. – Режим доступа : http://www.eduworks.com/Documents/eLearning_Interoperability_Standards_wp.pdf.

93. Delivery of accessible and SCORM-based courses in dotLRN [Электронный ресурс] / [Olga C. Santos, Carmen Barrera, Emmanuelle Gutiérrez y Restrepo et al.]. Режим доступа : <http://openacs.org/file/988243/Delivery+of+accessible+and+SCORM-based+courses+in+dotLRN.pdf>.

94. Deza E. Encyclopedia of Distances / Elena Deza, Michel Marie Deza. – Springer, 2009. – P. 94.

95. Drouin, P. Term Extraction Using Non-technical Corpora as a Point of Leverage / P. Drouin // Terminology. – 2003. – № 9(1). – P. 99–115.

96. Edward R. Jones. Implications of SCORM and Emerging E-learning Standards On Engineering Education [Электронный ресурс] / Edward R. Jones // Proceedings of the 2002 ASEE Gulf-Southwest Annual Conference, March 20–22, 2002. – LA, Lafayette, 2002. – Режим доступа : <http://sorubank.ege.edu.tr/~e190411147/scorm/scorm5.pdf>.

97. e-Learning Software [Электронный ресурс] // e-Learning Consulting, 2014. – Режим доступа : <http://www.e-learningconsulting.com/products/index.html>.

98. Enhancing SCORM metadata for assessment authoring in e-Learning [Wen-Chih Chang, Hui-Huang Hsu, Timothy K. Smith, Chun-Chia Wang] // Journal of Computer Assisted Learning. – 2004. – Vol. 20, Issue 4. – P. 305–316.

99. eXe – the eLearning XHTML editor [Электронный ресурс] // eXe eXeLearning, 2014. – Режим доступа : <http://exelearning.org>.

100. Extended Relation Metadata for SCORM-based Learning Content Management Systems [Eric Jui-Lin Lu, Gwoboa Horng, Chia-Ssu Yu, Ling-Ying Chou] // Educational Technology & Society. – 2010. – Vol. 13 (1). – P. 220–235.

101. Franks Pat Why METADATA Matters [Электронный ресурс] / Pat Franks, Nancy Kunde. – Режим доступа : <http://www.arma.org/bookstore/files/Franks-Kunde1.pdf>.
102. Fung, P. Extracting key terms from Chinese and Japanese texts / P. Fung // The International Journal on Computer Processing of Oriental Language (IJCPOL). – 1998. – № 12(1). – Pp. 99–121.
103. Gelbukh A. Approach to construction of automatic morphological analysis systems for inflective languages with little effort / A. Gelbukh, G. Sidorov // Lecture Notes in Computer Science, 2003. – No. 2588. – P. 215–220.
104. Ghali Fawaz. Augmenting e-learning standards with adaptation [Электронный ресурс] / Fawaz Ghali, Alexandra I. Cristea, Maurice Hendrix. – Режим доступа : <http://prolearn.dcs.warwick.ac.uk/bibtexbrowser/files/fg-aic-mh-cate-2008.pdf>.
105. Giuffrida G. Knowledge-based metadata extraction from PostScript files / Giovanni Giuffrida, Eddie C. Shek, and Jihoon Yang // Proceedings of the fifth ACM conference on Digital libraries. – ACM New York, NY, USA, 2000. – P. 77–84.
106. Gonzalez-Barbonea V. Creating the first SCORM object / Victor Gonzalez-Barbonea, Luis Anido-Rifonb // Computers & Education, 2008. – № 51. – P. 1634–1647.
107. Greenberg J. Final report for the amega (automatic metadata generation applications) project / J. Greenberg., K. Spurgin, A. Crystal. – School of Information and Library Science, University of North Carolina at Chapel Hill. – 2005. – 14 p.
108. Greenberg J. Metadata extraction and harvesting: A comparison of two automatic metadata generation applications / J. Greenberg // Journal of Internet Cataloging. – Vol. 6, Issue 4. – 2004. – P. 59–82.
109. HaCohen-Kerner Y. Automatic Extraction of Keywords from Abstracts / Yaakov HaCohen-Kerner // Knowledge-Based Intelligent Information and Engineering Systems. – 2003. – Vol. 2773. – P. 843–849.
110. Hausser R. Three principled methods of automatic word form recognition / R. Hausser // Proc. of VEXTAL: Venecia per il Trattamento Automatico delle Lingue, Sept. 1999. – Venice, Italy, 1999. – P. 91–100.
111. Hetzner E. A simple method for citation metadata extraction using hidden markov models / Erik Hetzner // Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries. – ACM New York, NY, USA, 2008. – P. 280–284.

112. Hulth A. Improved automatic keyword extraction given more linguistic knowledge / A. Hulth // Proceedings of the 2003 conference on Empirical methods in natural language processing. – 2003. – P. 216–223.
113. IEEE Standard for Learning Object Metadata: IEEE Standard 1484.12.1-2002. – USA, NY, 2002. – 44 p.
114. Incremental Construction of Minimal Acyclic Finite-State Automata / [J. Daciuk, B. W. Watson, S. Mihov, R. E. Watson] // Computational Linguistics. – 2000. – № 1. – P. 3–16.
115. ISO 23081-1:2006 Information and documentation – Records management processes – Metadata for records – Part 1: Principles.
116. iSpring Products [Электронный ресурс] // iSpring Official Site, 2014. – Режим доступа : <http://www.ispringsolutions.com/products.html>.
117. Issues, Tasks and Program Structures to Roadmap Research in Question & Answering (Q&A) [Электронный ресурс] / [John Burger, Claire Cardie, Vinay Chaudhri et al.]. – Режим доступа : http://www.inf.ed.ac.uk/teaching/courses/tts/papers/qa_roadmap.pdf.
118. Kechedzhy K. E. Rank distributions of words in additive many-step Markov chains and the Zipf law [Электронный вариант] / K. E. Kechedzhy, O. V. Usatenko, V. A. Yampol'skii. – Режим доступа : <http://arxiv.org/pdf/physics/0406099.pdf>.
119. Keith Humphreys J. B. Phraserate: An HTML keyphrase extractor. Technical Report [Электронный ресурс] / J. B. Keith Humphreys. – 2002. – Режим доступа : <http://infomine.ucr.edu/projects/publications/Humphreys-2002-PhraseRate.pdf>.
120. Krause Eugene F. Taxicab Geometry: An Adventure in Non-Euclidean Geometry / Eugene F. Krause. – Dover, 1987. – 96 p.
121. Laura McGrath. Developing eLearning Policies at the Department Level / Laura McGrath // MERLOT Journal of Online Learning and Teaching. – 2006. – Vol. 2, No. 3. – P. 177–186.
122. Lawrence S. Indexing and Retrieval of Scientific Literature / S. Lawrence, K. Bollaker, L. Giles // Proceedings of the eighth international conference on Information and knowledge management. – New York, USA, 1999. – P. 139–146.
123. Karttunen L. Two-Level Rule Compiler [Электронный ресурс] / L. Karttunen, K. R. Beesley // Technical Report. XEROX Palo Alto Research Center / Palo Alto, CA, USA. – 1992. – Режим доступа : <http://www.stanford.edu/~laurik/.book2software/twolc.pdf>.

124. KEA: practical automatic keyphrase extraction / [I. Witten, G. Payne, E. Frank and others] // In Proceedings of the 4th ACM Conference on Digital Library, 1999. – New York, NY, USA, 1999. – P. 254–255.
125. Koskenniemi K. Two-level Morphology: A General Computational Model for Word-Form Recognition and Production [Электронный ресурс] / К. Koskenniemi. – Режим доступа : <http://www.ling.helsinki.fi/~koskenni/doc/Two-LevelMorphology.pdf>.
126. Lectora Inspire [Электронный ресурс] // Lectora Official Site, 2014. – Режим доступа : <http://lectora.com/e-learning-software/>.
127. Litvak M. Graph-based keyword extraction for single-document summarization / Marina Litvak, Mark Last // Proceedings of the Workshop on Multi-source Multilingual Information Extraction and Summarization. – Association for Computational Linguistics Stroudsburg, PA, USA, 2008. – P. 17–24.
128. Mao S. A dynamic feature generation system for automated metadata extraction in preservation of digital materials / S. Mao, J. W. Kim, G. R. Thoma // Proceedings. First International Workshop on Document Image Analysis for Libraries. – 2004. – P. 225–232.
129. Maratea A. Automatic Generation of SCORM Compliant Metadata for Portable Document Format Files / Antonio Maratea, Alfredo Petrosino, Mario Manzo // CompSysTech'12: International Conference on Computer Systems and Technologies, 2012. – Ruse, Bulgaria, 2012. – P. 360–367.
130. Marinai S. Metadata Extraction from PDF Documents for Digital Library Ingest / S. Marinai // 10th International Conference on Document Analysis and Recognition, IEEE Computer Society Press. – Washington, DC, 2009. – P. 251–255.
131. Matsuo, Y. Keyword Extraction from a Single Document using Word Co-occurrence Statistical Information / Y. Matsuo, M. Ishizuka // International Journal on Artificial Intelligence Tools (IJAIT). – 2004. – № 13(1). – Pp. 157–169.
132. Matthew Fox. Learning design and e-learning [Электронный ресурс] / Matthew Fox. – Epic Group plc., 2003. – 44 p. – Режим доступа : http://www.dokeos.com/doc/thirdparty/Epic_Whtp_learningdesign.pdf.

133. Miike S. The Structure and Function of the EDR Concept Dictionary / S. Miike, S. Amano, H. Uchida, T. Yokoi // Terminology and Knowledge Engineering. – 1990. – Vol. 1. – P. 31–40.
134. Milošević D. Adaptive Learning by Using SCOs Metadata / Danijela Milošević, Mirjana Brković // Interdisciplinary Journal of Knowledge and Learning Objects. – 2007. – Vol. 3. – P. 163–174.
135. Mohri Mehryar The Design Principles of a Weighted Finite-State Transducer Library / Mehryar Mohri, Fernando Pereira, Michael Riley // Theoretical Computer Science. Special issue on implementing automata. – 2000. – Vol. 231. – Issue 1. – P. 17–32.
136. MOS Solo [Электронный ресурс] // MOS: MindOnSite, 2014. – Режим доступа : <http://www.mindonsite.com/en/produits/mos-solo/>.
137. Oflazer K. Bootstrapping Morphological Analyzers by Combining Human Elicitation and Machine Learning / K. Oflazer, S. Nirenburg, M. McShane // Computational Linguistics, 2001. – No. 1. – P. 59–85.
138. Peng F. C. Accurate information extraction from research documents using conditional random fields // F. C. Peng, A. McCallum // Proc. of the Human Language Technology Conf. and North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2004). – New York: ACM Press, 2004. – P. 329–336.
139. Plas L. Automatic keyword extraction from spoken text. A comparison of two lexical resources: the EDR and WordNet // L. Plas, V. Pallotta, M. Rajman, H. Ghorbel // Proceedings of the 4th International Language Resources and Evaluation, European Language Resource Association, 26–28 May 2004. – Portugal, Lisbon, 2004. – P. 2205–220.
140. Plodzien J. An approach to the quality and reusability of metadata specifications for e-learning objects / Jacek Plodzien, Ewa Stemposz, Alina Stasiecka // Online Information Review. – 2006. – Vol. 30, No. 3. – P. 238–251.
141. Qu Ch. Searching SCORM Metadata in a RDF-based E-Learning P2P Network Using XQuery and Query by Example / Changtao Qu, Wolfgang Nejdl // Proceedings of the 3rd IEEE International Conference on Advanced Learning Technologies, 9–11 July 2003. – Athens, Greece, 2003. – P. 81–85.
142. Reference metadata extraction using a hierarchical knowledge representation framework / [Min-Yuh Daya, Richard Tzong-Han Tsai, Cheng-Lung Sung and others] // Decision Support Systems. – Vol. 43, Issue 1. – 2007. – P. 152–167.

143. RELOAD Project [Електронний ресурс] // RELOAD Official Site, 2014. – Режим доступу : <http://www.reload.ac.uk/>.
144. Rodgers J. L. Thirteen ways to look at the correlation coefficient / J. L. Rodgers, W. A. Nicewander // The American Statistician. – 1988. – Vol. 42(1). – P. 59–66.
145. Salton, G. Automatic text processing: the transformation, analysis and retrieval of information by computer / G. Salton. – Reading, MA : Addison Wesley, 1998.
146. Şah M. Automatic metadata extraction from multilingual enterprise content / Melike Şah, Vincent Wade // Proceedings of the 19th ACM international conference on Information and knowledge management. – ACM New York, NY, USA, 2010. – P. 1665–1668.
147. Schone P. Knowledge-free induction of inflectional morphologies / P. Schone, D. Jurafsky // NAACL'01: Second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies, 2001. – Morristown, NJ, USA, 2001. – P. 1–9.
148. SCORM Users Guide for Instructional Designers: Scorm 2004 4th. Edition [Електронний ресурс]. – September 15, 2011. – Version 8. – Режим доступу : http://www.adlnet.gov/wp-content/uploads/2011/12/SCORM_Users_Guide_for_ISDs.pdf.
149. Segalovic Ilya. A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search Engine [Електронний ресурс] / Ilya Segalovic // Proceedings of the International Conference on Machine Learning; Models, Technologies and Applications / Las Vegas, Nevada, USA. – 2003. – Режим доступу : <http://download.yandex.ru/company/iseg-las-vegas.pdf>.
150. Seymore K. Learning for Information Extraction Learning Hidden Markov Model Structure for Information Extraction / K. Seymore, A. McCallum, R. Rosenfeld // In Working Notes of the AAAI Workshop on Machine Learning for Information Extraction. – 1999. – P. 37–42.
151. Sheth A. Managing semantic content for the Web / [A. Sheth, C. Bertram, D. Avant and others] // Internet Computing, IEEE. – Vol. 6, Issue 4. – P. 80–87.
152. SoftChalk [Електронний ресурс] // SoftChalk Official Site, 2014. – Режим доступу : <http://softchalk.com/products>.
153. Stepanova I. S. Creating learning material according to SCORM standard / I. S. Stepanova, I. A. Kravchuk // Мова, культура та освіта:

наук.-практ. конф., 18 квітня 2013 р: тези доповідей. – Вінниця : ВНАУ, 2013. – С. 277–279.

154. Suzuki Y. Keyword extraction of radio news using term weighting with an encyclopedia and newspaper articles / Y. Suzuki, F. Fukumoto, Y. Sekiguchi // SIGIR, 1998.

155. TableSeer: automatic table metadata extraction and searching in digital libraries / [Ying Liu, Kun Bai, Prasenjit Mitra, C. Lee Giles] // Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries. – ACM New York, NY, USA, 2007. – P. 91–100.

156. Task Force on Metadata. Summary Report [Електронний ресурс] // American Library Association. – 1999. – Режим доступу : <http://www.libraries.psu.edu/tas/jca/ccda/tf-meta3.html>.

157. The Importance of Metadata to Electronic Document Management [Електронний ресурс]. – Режим доступу : <http://www.doculex.com/the-importance-of-metadata-to-electronic-document-management-2/>.

158. The Xerte Project [Електронний ресурс] // The University of Nottingham Official Site, 2014. – Режим доступу : <http://www.nottingham.ac.uk/xerte/index.aspx>.

159. ToolBook [Електронний ресурс] // SumTotal Systems Official Site, 2014. – Режим доступу : <http://tb.sumtotalsystems.com/index.html>.

160. Trace A. Urdan. Corporate E-learning: Exploring a New Frontier / Trace A. Urdan, Cornelia C. Weggen. – W. R. Hambrecht, 2000. – 92 p.

161. Turing A. Computing Machinery and Intelligence» / A. Turing // Mind, 1950. – Vol. LIX. – No. 236. – P. 433–460.

162. Uduu Course Authoring Tool [Електронний ресурс] // Uduu Official Site, 2014. – Режим доступу : http://www.uduu.com/solutions_uduu.html.

163. WordNet: A lexical database for English [Електронний ресурс] / Princeton University. – Режим доступу : <http://wordnet.princeton.edu/wordnet/>.

164. Zhang Ch. Automatic Keyword Extraction from Documents Using Conditional Random Fields / Zhang, Chengzhi // Journal of Computational Information Systems. – 2008. – Vol. 4, No. 3. – P. 1169–1180.

165. Zhang K. Keyword Extraction Using Support Vector Machine / Kuo Zhang, Hui Xu, Jie Tang, Juanzi Li // Advances in Web-Age Information Management. – 2006. – Vol. 4016. – P. 85–96.

Наукове видання

**Кравчук Ірина Анатоліївна
Бісікало Олег Володимирович**

**ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ФОРМУВАННЯ
МЕТАДАНИХ ДЛЯ СИСТЕМ
АВТОМАТИЗОВАНОГО ДОКУМЕНТООБІГУ**

Монографія

Редактор Н. Мазур
Оригінал-макет підготовлено І. Кравчук

Підписано до друку 17.11.2016 р.
Формат 29,7×42¼. Папір офсетний.
Гарнітура Times New Roman.
Друк різнографічний. Ум. др. арк. 9,47.
Наклад 300 (1-й запуск 1–75) пр. Зам № В2016-16

Вінницький національний технічний університет,
КІВЦ ВНТУ,
21021, м. Вінниця, Хмельницьке шосе, 95,
ВНТУ, ГНК, к. 114.
Тел. (0432) 59-85-32.

publish.vntu.edu.ua; email: kivc.vntu@gmail.com.

Свідоцтво суб'єкта видавничої справи
серія ДК № 3516 від 01.07.2009 р.

Віддруковано ФОП Барановська Т. П.
21021, м. Вінниця, вул. Порики, 7.
Свідоцтво суб'єкта видавничої справи
серія ДК № 4377 від 31.07.2012 р.