

Організаційні аспекти входження вітчизняного інформаційного виробництва в систему міжнародної інформаційної кооперації¹

Друга група проблем міжнародної кооперації при формуванні масивів електронних ресурсів пов'язана з необхідністю забезпечення інформаційно-технологічної сумісності всіх учасників корпоративних проектів. Прихильники «відкритого доступу» доклали значних зусиль, щоб публікації в численних репозитаріях, розкиданих по всьому світу й по-різному організованих їхніми творцями, могли стати для користувачів віртуальною єдністю. По-перше, розроблено єдині принципи опису ресурсів, засновані на використанні метаданих Дублінського ядра. По-друге, розроблено протокол збору цих метаданих з різних архівів – OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting).

Стандарт Дублінського ядра метаданих (Dublin Core Metadata) запропоновано Онлайновим комп'ютерним бібліотечним центром OCLC для опису ресурсів Інтернету². Структура опису документів, на думку розробників, повинна бути простою та замінити складні системи існуючої каталогізації документів. Вона призначена для запису базових структурних значень опису документів на мовах розмітки HTML та XML і містить у собі 15 елементів, семантику яких визначено міжнародними групами професіоналів у галузі бібліотечної справи, обчислювальної техніки, кодування текстів, фахівців музейної справи та інших суміжних галузей наук.

Набір елементів Дублінського ядра метаданих (Dublin Core Metadata Element Set) використовують для опису різноманітних ресурсів, доступних в Інтернеті та інших мережах. Він нескладний за структурою, відносно легкий у застосуванні, інтернаціональний (застосовується в усьому світі). У 2001 р. набір Дублінського ядра метаданих було затверджено в США Американським інститутом національних стандартів як стандарт Z39.85. В Австралії, Великобританії, Данії, Ірландії, Канаді та Фінляндії в цьому ж році формат Дублінського ядра метаданих рекомендовано та прийнято як державний стандарт для онлайнових ресурсів та е-комерції. У 2003 р. Міжнародною організацією стандартів ISO ухвалено фінальний текст документа під назвою ISO 15836, яким набір елементів Дублінського ядра метаданих затверджено як міжнародний стандарт³. У передмові документа зазначено, що ці метадані об'єднують бібліотекарів, дослідників цифрових бібліотек, постачальників

¹ Закінчення. Початок див.: Шляхи розвитку української науки. – 2015. – № 4. – С. 99–105.

² Dublin Core Metadata Initiative Repositories [Electronic resource]. – Mode of access: <http://dublincore.org/>.

³ *Копанєва В.* Бібліотека як центр збереження інформаційних ресурсів Інтернету: [монографія]. – К., 2009. – 198 с.

контенту та експертів з розмітки текстів з метою дослідження стандартів щодо інформаційних ресурсів.

Формат Дублінського ядра метаданих містить 15 елементів для опису електронних ресурсів⁴.

1. Назва (Title): Ім'я цифрового ресурсу, надане йому створювачем чи видавцем цифрового ресурсу.

2. Створювач (Creator): Персональне ім'я чи колективний автор, що несуть первинну інтелектуальну відповідальність за створення інтелектуального змісту ресурсу (наприклад, автор, що створив документ, художник, фотограф, власник оригінального зібрання артефактів, чи ілюстратор візуальних ресурсів).

3. Тема (Subject, предмет): Предметна рубрика цифрового ресурсу. Предметна рубрика може бути виражена ключовим словом чи фразою, що описують тему (теми) ресурсу для позначення зв'язку предмета з людьми, територіями, подіями чи іншою контекстною інформацією.

4. Опис (Description): Текстовий опис змісту ресурсу, у тому числі анотації, або опис змісту стосовно візуальних ресурсів, а також інші описи, не включені в інші поля.

5. Видавець (Publisher): Організація, що створила цифровий ресурс, доступний у його дійсній формі за допомогою телекомунікації (видавництво, музей, університет чи бібліотека).

6. Співавтор (Contributor): Особа чи організація – додаток до тих, які визначено в елементі створювач (Creator), що несе істотну інтелектуальну відповідальність стосовно оригінального ресурсу, проте ця відповідальність є вторинною.

7. Дата (Date): Дата створення чи відкриття доступу до цифрового ресурсу.

8. Формат (Format): Формат даних ресурсу, який описується, та додаткові (необов'язкові) визначення. Використовується для програмного або апаратного забезпечення, необхідного для відображення, або роботи з ресурсом.

9. Тип (Type): Категорія чи жанр ресурсу. Подання: Текст, зображення, звук, набір даних, програма, подія, інтерактивні медіа, фізичний об'єкт.

10. Ідентифікатор (Identifier): Рядок або число, що використовується для унікального ототожнення цифрового ресурсу. Для мережевих ресурсів: URL/URN. Для інших ресурсів – глобальні, унікальні ідентифікатори: ISBN, ISSN, інвентарний номер, або інші формальні імена, що ідентифікують офлайнний (неопублікований в Інтернеті) ресурс.

11. Джерело (Source): Інформація про оригінальний фізичний ресурс, з якого «витягнуто» поточний цифровий ресурс.

⁴ Dublin Core Metadata Initiative Repositories. – Mode of access: <http://dublincore.org/>.

12. Мова (Language): Вказує на мову інтелектуального контексту цифрового ресурсу.

13. Відношення (Relation): Відношення цього ресурсу до іншого ресурсу, плюс ідентифікатор іншого ресурсу.

14. Охоплення (Coverage): описує просторові чи часові характеристики інтелектуального змісту ресурсу.

15. Авторські права (Rights): Зміст цього елемента має позначити авторські права та їх власника.

П'ятнадцять елементів Дублінського ядра метаданих за їх призначенням можна поділити на три основні групи, відповідно до сутності інформації, яку вони у собі вміщують. Це елементи, що мають відношення до:

- змісту ресурсу (заголовки, предмет, опис, джерело, мова, відношення, зона дії);
- інтелектуальної власності (автор, видавець, відомості про відповідальність, правові аспекти);
- поточного стану (властивостей) (дата, тип ресурсу, формат, ідентифікатор).

У стандарті Дублінського ядра метаданих втілено такі характеристики:

- простота створення та підтримки даних – він являє собою мінімально простий набір елементів, що дає змогу створювати прості описові записи для цифрових ресурсів не тільки фахівцями, а й забезпечує ефективний механізм запитів до ресурсів у мережевому оточенні;

- семантика загального вживання – пошук ресурсів серед великої кількості загальної (часто слабо структурованої) інформації у мережі Інтернет є досить складною справою. Дублінське ядро метаданих може допомогти «досліднику цифрових ресурсів» (досліднику-непрофесіоналу) знайти потрібну інформацію через послідовність загальних елементів;

- урахування притаманних різним країнам специфічних лінгвістичних проблем. Розробники Дублінського ядра метаданих – представники практично всіх континентів планети – впевнені, що розвиток цього стандарту вирішить проблеми, пов'язані зі складною багатомовною та мультикультурною природою електронних ресурсів планети;

- розширюваність – у разі досягнення балансу, з одного боку, між простотою опису цифрових ресурсів і необхідністю ретельного опису деталей з іншого, розробники усвідомлюють важливість створення механізму, здатного розширити набір елементів залежно від специфіки ресурсу. Передбачається, що розробники інших подібних стандартів створять додаткові набори елементів для опису ресурсів, які дають можливість інтегрувати їх до набору Дублінського ядра метаданих з метою максимально повного опису ресурсів.

За правилами Дублінського ядра метаданих кожен з 15 елементів не є обов'язковим і може повторюватися. Опис інтернет-ресурсів з використанням

цих метаданих можна, у першому наближенні, розглядати як бібліографічний опис книги чи аналітичний розпис журналу (газети). Цей формат у першому наближенні можна вважати аналогом міждержавного стандарту ДСТУ ГОСТ 7.1:2006⁵.

Слід додати, що існують програми-конвертори з формату Дублінського ядра метаданих у MARC-формати. Дотримання Дублінського ядра метаданих дає змогу досягти інформаційної сумісності учасників міжнародних кооперативних проектів. Для забезпечення «телекомунікаційної сумісності» розроблено протокол збору метаданих з різних відкритих архівів – OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting)⁶. Кожен відкритий архів (репозитарій) має бути доступним для авторизованих збирачів метаданих, що функціонують відповідно до згаданого протоколу. Для цього він повинен мати механізм доступу, здатний виконувати за запитами збирачів метаданих набір операцій, який забезпечується протоколом OAI-PMH.

Завдяки стандартизації подання метаданих у репозиторіях і стандартизації протоколу OAI-PMH, що визначає взаємодії відкритих архівів, забезпечується їх інтеоперабельність. Відповідно до протоколу OAI-PMH, архів, визначений центральним, може здійснювати збір метаданих з інших, зареєстрованих у ньому відкритих архівів відповідно до встановленого регламенту.

У результаті такий центральний архів виступає як своєрідний інформаційний хаб – система інтеграції даних, що базується на матеріалізованій інтеграції метаданих із зареєстрованих у ньому джерел – відкритих архівів. Метадані для опису інформаційних об'єктів – публікацій – не порушуються. Вони піддаються віртуальній інтеграції. В інтегрованому репозитарії метаданих забезпечується їх унікальна ідентифікація. Повні тексти зберігаються в місцях постійного зберігання та стають доступними за допомогою сервісу для користувачів інтегрованого відкритого архіву, який розглядається як посилання на нього в картці метаданих, що описує цей об'єкт. Пошук здійснюється на центральному сервері, а за повним текстом користувач звертається до відповідного архіву. Серед переваг такого підходу можна виділити простоту, сучасність, забезпечення високої якості сервісів, можливості розвитку, масштабованість, можливість інтегрувати ресурси з багатьма іншими відкритими ресурсами.

В Україні за такою схемою функціонує сервер oai.org.ua, що був розроблений фахівцями Інституту програмних систем НАН України та Житомирського державного університету ім. І. Франка. Він забезпечує

⁵ *Копанєва В.* Бібліотека як центр збереження інформаційних ресурсів Інтернету: [монографія]. – К., 2009. – 198 с.

⁶ Open Archives Initiative – Protocol for Metadata Harvesting. – Mode of access: <http://www.openarchives.org/pmh/>.

централізований пошук у 48 публічних архівах України, створених бібліотеками університетів і академічних інститутів ⁷.

Вищевикладене говорить про наявність певних досягнень у сфері розв'язання проблем дотримання авторських прав і забезпечення інформаційно-технологічної сумісності електронних ресурсів при реалізації кооперативних проектів їх формування в рамках як національних, так і міжнародних проектів.

Значно менш проробленою є організаційно-управлінська проблематика. Вона легко розв'язується за умови належного фінансування відповідних програм з визначенням головної організації та розподілу завдань між співвиконавцями. Однак, враховуючи економічні негаразди сьогодення, слід орієнтуватися на синергетичні підходи та механізми організації кооперації.

Цілісну концепцію такої міжнародної кооперації для створення інтегрованого славистичного ресурсу запропонував О. Онищенко ⁸. Він зазначив, що ідеальним варіантом була б організація доступу через одного інтегратора інформації (або декількох регіональних, тематичних) до всіх існуючих у світі фондів слов'янської та слов'янознавчої літератури. Хоча всі вони розпорошені по мережі Інтернет, однак певні впорядковані масиви інформації про слов'янство можна знайти вже сьогодні в міжнародних бібліотечних каталогах.

Найбільшим з них є OCLC (Online Computer Library Center – Онлайновий комп'ютерний бібліотечний центр). OCLC спеціалізується на сприянні розширенню доступу до світової інформації. Це містить у собі і проблему накопичення славистичної інформації та розширення доступу до неї. Сьогодні OCLC відображає інформаційні ресурси понад 36 тис. бібліотек із 79 країн, у т. ч. слов'янських. У зведеному електронному каталозі OCLC є понад 42 млн записів на 400 мовах (у т. ч. і багатьох слов'янських). Убачається доцільність створення в системі OCLC тематичного інтегратора славистичних ресурсів з деталями свого пошукового апарату до них. Це вписується у стратегію розвитку OCLC, що розгортає новий проект – CORC (OCLC Cooperative Online Resource Catalog). Він дає змогу створювати тематичні покажчики (у т. ч. і славистичного характеру). Однак OCLC – комерційна структура, що припускає відомі обмеження.

⁷ Резніченко В. А., Новицький О. В., Проскудіна Г. Ю. Інтеграція наукових електронних бібліотек на основі протоколу OAI-PMH // Проблеми програмування. – 2007. – № 2. – С. 97–112.

⁸ Онищенко А. С. Концептуальные и практические проблемы создания интегрированного электронного ресурса славянских библиотечных фондов // Сучасна славистика: ключові проблеми та тенденції розвитку: тези Міжнар. наук. конф. до Дня слов'янської писемності і культури, 23–24 трав. 2013 р. / НАН України, Укр. ком. славистів, Нац. б-ка України ім. В. І. Вернадського. – К., 2013. – С. 3–4.

Тому паралельно з OCLC слід шукати платформи для інтегрування славістичних інформаційних ресурсів у відкритих інтернет-системах. Такими найбільш відомими на сьогодні є Slavistik-portal Віртуальної бібліотеки Східної Європи та база Ethnodoc на сайті ViFa Ost. Остання містить матеріали з дослідження національних меншин Східно-Центральної та Південно-Східної Європи (на сьогодні переважно останньої). Присутня тут інформація стосується 170 національних меншин з 20 країн і регіонів. Вона структурована за етнічними, релігійними та мовними групами.

Майже 10 років ведеться реєстр сховищ відкритого доступу до наукових публікацій – ROAR (Registry for Open Access Repositories). Станом на середину 2014 р. у ньому зафіксовано близько 3,8 тис. відкритих архівів, у т. ч. 113 – у Польщі, 70 – в Україні, 49 – у Росії, 11 – у Білорусі⁹. Слов'янська частина ROAR, як бачимо, чимала. І це говорить про можливість створення в його межах об'єднаного славістичного архіву.

Більшість із перерахованих вище та їм подібних інтеграторів електронних ресурсів мають універсальний характер, де переважає природничо-наукова, технічна, технологічна інформація. Соціогуманітарний матеріал займає там менше місця. Але досвід і можливості цих інтеграторів інформації з успіхом можуть бути використані для посилення інформаційного забезпечення славістичної справи.

Ближче до завдань створення об'єднаного електронного ресурсу славістичної інформації підійшли інтегратори соціогуманітарних і профільних слов'янознавчих баз даних. Прикладом спеціалізованого інтегратора соціогуманітарних матеріалів може бути система «Соціонет»¹⁰. Вона є варіантом реалізації ідеї відкритих архівів. Інститути Відділення суспільних наук Російської академії наук вирішили всі свої дослідження обов'язково депонувати в електронній формі в архівах своїх установ. Дані про ці інформаційні об'єкти передаються в систему «Соціонет», де накопичуються метадані та встановлюються гіперактивні зв'язки між ними.

У всьому світі, і у слов'янських країнах зокрема, іде оцифрування рукописної, архівної та книжкової спадщини. Це дає можливість поставити питання про організацію Всесвітньої цифрової слов'янської бібліотеки. Якщо виникла ідея (і реалізується) Світової цифрової бібліотеки універсального характеру, то вбачається доцільним започаткувати створення спеціалізованої слов'янської бібліотеки. Для початку вона мала б націлилася на збір максимально повної бібліографічної інформації про наявність слов'янських документів у світі та паралельно розглянути можливості формування

⁹ Registry of Open Access Repositories. – Mode of access: <http://roar.eprints.org/>.

¹⁰ Онлайновая научная инфраструктура «Соционет». – Режим доступа: <https://socionet.ru/>.

повнотекстові бази. Очевидно, у найближчому майбутньому славістичний ресурс існуватиме в розподіленому вигляді. А шлях розширення віддаленого доступу лежатиме через прискорення оцифрування конкретних фондів по всьому світу. Інтеграція цифрового матеріалу відбуватиметься в національному, регіональному та міжнародному масштабах. При цьому будуть наповнюватися існуючі та створюватися нові сховища слов'янської електронної інформації.

Виправдали б себе і міжнародні центри-інтегратори ресурсів конкретних напрямів славістики. Вони фактично сформувалися на базі провідних бібліотек кожної зі слов'янських країн. Тому реально ставити завдання, щоб, скажімо, усі електронні матеріали щодо богемістики з усього світу збиралися, оброблялися та пропонувалися в науковий та навчальний обіг у одній з бібліотек Чехії, з полоніки – у Польщі, з україніки – в Україні тощо.

Найважливіше та найскладніше – формування інтегрованих повнотекстових славістичних баз і забезпечення вільного доступу до них. Тут звичайно виникають питання авторського, майнового права, а також технічні, технологічні, мовні проблеми. Але все ж об'єднаний славістичний ресурс світу можливий. Його створення має багатопланове історико-культурне значення. По-перше, у світовій інформаційний обіг увійшли б зафіксовані в документах багатовікові досягнення в науці, культурі, освіті, громадському облаштуванні слов'янства. По-друге, сама наука славістика отримала б потужну інформаційну платформу для осмислення історичного минулого, сучасних тенденцій та перспектив слов'янських народів. Навіть уже опрацьовані національні та міжнародні частини майбутнього інтегрованого слов'янського інформаційного ресурсу говорять про вагоме місце слов'янства у світі. Широке використання такого історичного, соціокультурного, соціогуманітарного надбання дасть змогу слов'янським народам швидше й легше вписатися в інформаційне суспільство, що формується, увійти в нього зі збереженням і збагаченням кращих рис своєї самобутності, національної історико-культурної спадщини. Примножиться науково-інформаційна база для розвитку міжслов'янської співпраці, діалогу з іншими народами, творчого внеску в побудову нової цивілізаційної картини світу.

Викладене дає змогу зробити такі висновки й пропозиції щодо розвитку міжнародної кооперації в інформаційному виробництві.

1. Проблема розвитку правового підґрунтя міжнародної кооперації при формуванні інформаційних ресурсів має вирішуватися шляхом переходу від заборонної парадигми класичного авторського права до дозвільної системи розповсюдження документованої інформації з урахуванням прав інтелектуальних власників. Таку можливість забезпечують, зокрема, публічні ліцензії та ліцензії Creative Commons. Вони надають інтелектуальним

власникам гнучкі можливості обмежити використання своїх напрацювань у тих рамках, які вони вважають потрібними.

2. Проблему інформаційно-технологічної сумісності в кооперативних проєктах формування електронних ресурсів необхідно розв'язувати дотриманням міжнародних стандартів і вимог. У телекомунікаційному аспекті слід активізувати використання протоколу обміну метаданими ініціативи відкритих архівів ОАІ-РМН (Open Archives Initiative – Protocol for Metadata Harvesting), а в змістовному – здійснювати опис мережевих ресурсів на основі Дублінського ядра метаданих.

3. Значна частина проєктів міжнародної кооперації в інформаційному виробництві здійснюється на синергетичних засадах. Безперечно вбачається доцільність участі в них українських фахівців. Водночас має сенс ініціювати ряд міжнародних, регіональних і міждержавних проєктів інтеграції славістичних інформаційних ресурсів. Він дасть змогу Україні разом з іншими слов'янськими народами увійти в інформаційне суспільство, що формується, зі збереженням своєї самобутності (*Національні інформаційні ресурси як інтегративний чинник вітчизняного соціокультурного середовища: [монографія] / [О. С. Онищенко, В. М. Горовий, В. І. Попик та ін.]; НАН України, Нац. б-ка України ім. В. І. Вернадського. – К., 2014. – С. 208–216*).