

МЕТАДАНИ НАУКОВИХ ДОКУМЕНТІВ ЯК СКЛАДОВА СИСТЕМИ ІНФОРМАЦІЙНИХ РЕСУРСІВ «ВІДКРИТОЇ НАУКИ»

Відкрита наука – це рух, мета якого полягає в отриманні більш доступних результатів досліджень, включаючи код, дані та наукові документи. Він охоплює багато різних, але часто пов'язаних аспектів, що впливають на весь життєвий цикл дослідження, включно із відкритим доступом до публікацій, дослідницькими даними, програмним забезпеченням з відкритим кодом і робочими процесами, суспільною наукою, із відкритими освітніми ресурсами та альтернативними методами оцінки досліджень. А також із відкритим рецензуванням, експертними оцінками тощо. Тобто, створюється наука, заснована на спільних принципах, які забезпечують порівняльність, сумісність і користь для всіх досліджень на планеті. Запорукою ефективного застосування та інтеграції ресурсів відкритої науки є їх структурований опис, що базується на принципах повноти та необхідності метаінформації, зручності її використання та інтероперабельності. Одиницями такого опису є метадані.

Дане дослідження не охоплює всього широкого спектру ресурсів відкритої науки. Його метою є структуризація та формалізація лише описів наукових документів як значущої частини наукових знань відкритої науки. Для реалізації поставленої цілі в статті визначено таксономію ресурсів відкритих наукових документів та запропоновано єдину інтегровану систему їхніх метаданих, яка містить як множину загальних характеристик, притаманних усім типам наукових документів, незалежно від їх призначення, форми чи формату представлення, так і ієрархічно поширювану систему специфічних ознак окремих типів ресурсів.

Ключові слова: ресурси відкритої науки, відкриті наукові знання, метадані наукових документів, схеми метаданих, метадані наукових публікацій, метадані освітніх матеріалів, таксономія понять відкритої науки, інфраструктура відкритої науки, тип відкритого ресурсу, метадані дисертацій, метадані монографій, формати представлення наукових ресурсів.

Вступ

Інституціоналізація відкритої науки зародилася кілька десятиріч тому як політика перетворення наукової практики з метою адаптації до змін, викликів, можливостей та ризиків цифрової ери та підсилення впливу науки на суспільство. Для забезпечення колективного розуміння його наслідків в усьому світі, Організація Об'єднаних Націй з питань освіти, науки та культури (ЮНЕСКО) в листопаді 2021 року ухвалила набір рекомендацій для міжнародної основи політики й практики *відкритої науки* із загальними цінностями, принципами та стандартами. ЮНЕСКО визначає відкриту науку як «інклюзивну конструкцію, яка поєднує різні рухи та практики, спрямовані на те, щоб зробити багатомовні наукові знання відкритими, доступними та придатними для повторного використання, розширити наукове співробітництво й обмін інформацією для блага науки й суспільства, а також відкрити процеси створення, оцінки й передачі наукових знань представникам суспільства за межами традиційної наукової спільноти».

Тобто відкрита наука – це прозорі та доступні знання, що розповсюджуються

та розвиваються через спільні мережі. Але слід зазначити, що відкритий підхід не лише надає різні типи знань, що є у вільному доступі для всіх учасників спільноти для повторного використання, а й забезпечує цілісну інфраструктуру сумісних, стандартизованих послуг та апаратних засобів. Це сприяє дефрагментації, підтримці та спрощенню наукових практик і послуг, цілісності досліджень і наукової дипломатії в цілому, а також міжнародному науковому співробітництву. Створюється наука, заснована на спільних принципах, які забезпечують порівняльність, сумісність і користь для всіх досліджень на планеті. Сховища, ресурси, бази даних і сервіси проекту стають актуальними для вчених у географічному масштабі за межами регіону, в якому вони розташовані, що приводить до розвитку потенціалу в глобальному масштабі.

Відкрита наука – це рух, метою якого є отримання більш доступних результатів досліджень, включно із кодом, даними та науковими паперами. Він охоплює багато різних, але часто пов'язаних аспектів,

що впливають на весь життєвий цикл дослідження, включно із відкритим доступом до публікацій, дослідницькими даними, програмним забезпеченням з відкритим кодом і робочими процесами, суспільною наукою, із відкритими освітніми ресурсами та альтернативними методами оцінки досліджень. А також із відкритим рецензуванням, експертними оцінками тощо.

Загальна таксономія понять Відкритої науки [1] наведена на рис.1.

Для забезпечення доступу до наукових знань вводиться поняття «інфраструктура відкритої науки». Мається на увазі віртуальна або фізична дослідницька інфраструктура загального користування включно із основним науковим обладнанням та набором інструментів, таких як колекції, журнали та платформи відкритого доступу до публікацій, сховища, архіви та наукові дані. А також сучасні інформаційні та відкриті наукометричні системи, що використовуються

для оцінки та аналізу різних галузей науки. Крім того відкрита інфраструктура для обчислень та обробки даних дозволяє здійснювати колективний і мультидисциплінарний аналіз даних, вивчати об'єкти цифрової інфраструктури, які необхідні для підтримки відкритої науки і задоволення потреб різних товариств. Різні сховища даних мають бути адаптовані до конкретних характеристик об'єктів, розміщених в них (публікацій, даних або кодів), до місцевих умов, потреб користувачів та вимог дослідницьких спільнот, але водночас вони повинні застосовувати стандарти функціональної сумісності та передові методи роботи для забезпечення належної перевірки тих матеріалів, що розміщуються в сховищах та для полегшення їх пошуку й повторного використання. Але, одразу зазначимо, що відкрите апаратне забезпечення та інфраструктура, яка забезпечує доступ до відкритих знань, виходять за межі цього дослідження.



Рис.1. Таксономія Відкритої науки

Фокусом даного дослідження є лише структуризація та формалізація описів наукових документів як значущої частини наукових знань (ресурсів) відкритої науки за допомогою створення системи їхніх метаданих.

Таксономія ресурсів «відкриті знання»

Насамперед, слід визначитися, що ж саме ми будемо розуміти під науковими знаннями. Тобто, які типи ресурсів відкритої науки створюють «відкриті наукові знання».

Загалом, під науковим знанням розуміють набір надійної інформації про Світ, отриманої в процесі збору даних, експериментів та аналізу, досягнутих за допомогою наукових досліджень. На макро рівні всі типи наукових знань можна розділити на кілька великих груп: наукові документи, науковці (фізичні та юридичні особи), експонати (макети, моделі), наукові програми досліджень та наукові заходи. Аналіз джерел наукових знань дозволив детальніше класифікувати основні типи ресурсів відкритої науки у такий спосіб:

– Наукові документи

- Наукові публікації
 - Монографії
 - Дисертації
 - Препринти
 - Автореферати
 - Наукові реферативні огляди
 - Тези
 - Бакалаврські
 - Магістерські
 - Кандидатські
 - Докторські
 - Інтернет-ресурси
 - Наукові звіти
 - Аналітичні огляди
 - Наукові видання (журнали, збірники тощо)
 - Статті
 - З журналів
 - Зі збірників
 - З газет
- Матеріали форумів, конференцій, семінарів тощо

- Наукові доповіді
- Тези доповідей
- Презентації
- Матеріали стендів
- Освітні матеріали
 - Матеріали курсів
 - Підручники
 - Лекції
 - Методичні та дидактичні матеріали
 - Презентації
 - Реферати
- Патенти
- Програмне забезпечення (ПЗ)
 - Проекти, схеми, моделі даних, процесів тощо
 - Програмні застосунки та бібліотеки
 - Відкритий вихідний код ПЗ
 - Ліцензії
 - Набори даних
 - Івенти у програмних застосунках
- Експертні оцінки
- Рецензії
- Свідоцтва про реєстрацію авторських прав
- Інтерактивні ресурси
- Мапи
- Музичні нотатки
- Аудіо матеріали
- Зображення
 - Статичні рисунки
 - Відеоматеріали
- Протоколи випробувань
- Статистична інформація різних типів
 - Міжнародна музейна статистика
- Закони, нормативні акти тощо
- Експонати, моделі, макети, історичні і культурні пам'ятки (цифрові описи з фотоматеріалами), матеріали музейних і бібліотечних архівів
- Експозиції, виставки
- Наукові програми

- Освітні
- Дослідницькі
- Експериментальні
- Наукові заходи
 - Конференції
 - Симпозіуми
 - Семінари
 - Наради
 - Випробування
 - Виставки
- Наукові дослідники (фізичні та юридичні особи)
 - Науковці/дослідники (фізичні особи)
 - Наукові установи
 - Науково – дослідницькі інститути
 - Учбові центри
 - Конструкторські бюро
 - Університети
 - Музеї
 - Виставкові центри
 - Архіви
 - Наукові бібліотеки
 - Інші дослідницькі організації

Наукові документи є відкритими, до них забезпечується безкоштовний доступ, вони можуть багаторазово використовуватися, зберігатися та повторно поширюватись будь-якою особою за умови зазначення джерела. Для програмного забезпечення відкритий вихідний код в людино- і машиночитаному форматі надається користувачам у зручній для них формі відповідно до відкритої ліцензії. Це дає третім особам право доступу, зміни, доповнення, вивчення, використання та/або розповсюдження програмного забезпечення і його вихідного коду, дизайну або концепції, а також створення на їхній основі похідних робіт. Найважливіші аспекти відкритих даних пов'язані з їхнім змістом та внутрішньою організацією. В деяких ситуаціях також важливу роль відіграє третій аспект відкритих даних – їх оточення. Це найскладніший, але дуже важливий аспект для стратегічного й тематичного розвитку аналітики й пошуку знань, особливо з підключенням суміжної проблематики. В прос-

торі відкритих даних контекстом для заданого набору є всі інші дані, з якими їх зможуть коректно зв'язати аналітики з тих чи інших підстав. Правильно визначити контекст можна лише, якщо для основних даних правильно задана предметна область (ПО) та їхнє призначення. Це досягається через опис даних набором спеціальних характеристик – метаданих.

Якість відкритих даних починається з якості їхніх метаданих. Зазвичай схеми метаданих визначаються, виходячи, перш за все, з таких аспектів даних, як сенс, структура та формат.

Метою даного дослідження є визначення схем метаданих для ресурсів «відкритих наукових документів» як складової інфраструктури відкритих ресурсів на основі існуючих схем та стандартів метаданих. Для кожної з виділених груп можна виокремити загальні метадані, які можуть бути використані для опису ресурсів усіх типів групи, та більш специфічні, що враховують особливості окремих типів.

Різновиди і типи метаданих

Загалом, у поняття метаданих вкладається додатковий опис якогось – об'єкту або процесу. В ракурсі даних досліджень такими об'єктами/процесами є різноманітні компоненти відкритих наукових знань. Сам об'єкт метаданих може мати різні форми і класифікуватись за певними ознаками. Наприклад, зміст – опис шуканого об'єкта у вигляді розміру й типу файлу, інформації про вміст ресурсу, ставлення до ресурсу або його компонентів – базова інформація про об'єкт тощо.

Будь-яка інформаційна система зазвичай передбачає класифікацію метаданих на три великі головні групи: внутрішні – опис явних ознак самого об'єкту (розмір або тип файлу); адміністративні – інформація про об'єкт (автор, виконавець тощо); описові – інформація про природу об'єкту, його особливі ознаки, посилання на інші об'єкти, пов'язані з шуканим [2].

Формат метаданих являє собою уніфіковану форму опису властивостей об'єкта, що дозволяє отримати більш повне уявлення про цей об'єкт. На сьогодні розроблено велику кількість стандартів мета-

даних для об'єктів різних типів, призначення та форматів представлення. Найпоширенішими є:

- MARC та його різновиди – стандарт, в основному застосовується для книг і бібліографічних ресурсів із зазначенням назви, автора, року написання або виходу.
- DCMІ – стандарт, прийнятий для опису інтернет-об'єктів, електронних документів, ресурсів тощо.
- FOAF і vCard – опис персоніфікованих даних людей і організацій (у форматі vCard при експорті з мобільних пристроїв зберігаються списки контактів).
- CDWA – стандарт для опису історичних або музейних цінностей.
- ONIX і PRISM – інформація про видавництво.
- CIF – кристалографія;
- VICAR – обробка зображень, отриманих із супутників.
- NewsXML – новинні метадані тощо.

Цей список можна продовжувати до нескінченності, оскільки для будь-якого аспекту людської діяльності сьогодні можна знайти єдиний підхід в описі. Тому мета даного дослідження – на основі запропонованої вище (з можливим розширенням) класифікації ресурсів відкритої науки та на підставі аналізу існуючих схем метаданих, пошукових систем, стандартів опису, так чи інакше пов'язаних з тим чи іншим об'єктом наукових знань, сформулювати єдині системи метаданих для кожної з груп наукових документів як ресурсів відкритої науки.

Базові аспекти побудови схеми метаданих наукових документів

Наукові публікації не лише складають більшу частину контенту відкритих документів просто за його обсягом, а й посідають центральне місце серед ресурсів даного типу. Адже зазвичай містять опис результатів наукових досліджень, а, отже, й нові наукові знання. Метадані наукових публікацій також є важливою частиною наукового дослідження, оскільки від них залежить показник цитування публікації. Українські та міжнародні журна-

ли існують з відкритою, закритою та змішаною політикою доступу. В наукометричних базах даних, зокрема, Scopus та Web of Science, незалежно від рівня доступу, назва, анотація та ключові слова завжди відображаються вичерпно. Вони повинні виконувати функцію реклами, оскільки із загальної кількості публікацій, представлених у реферативних базах даних, дослідник обере саме інформативну. Більшість наукових публікацій являють собою різновиди бібліографічних ресурсів, для опису яких найпоширенішими стандартами, на сьогодні, лишається MARC [14] та його різновиди. Також, як і для решти типів контентів та інтернет-ресурсів, значущими лишаються загальні характеристики, визначені DCMІ (Dublin Core Metadata Initiative) стандартом.

Але, зрозуміло, що цілісність наукових знань не можлива без усього широкого спектру наукових документів. Так аналітичні звіти та огляди не є опублікованими матеріалами, але зазвичай містять більш ємні та детальні описи результатів, ніж опубліковані за темою статті, препринти, монографії тощо. Дисертаційні роботи також мають особливості, як підсумок дослідження, проведеного з метою здобуття наукового ступеню автором. Зокрема, їх опис має враховувати вимоги та стандартні протоколи, застосовані до робіт даного типу. Недарма, в цифрових наукових бібліотеках дисертаційні роботи та автореферати до них виділяються в окремі репозиторії та пов'язані з певними критеріями пошуку у сховищі.

На відміну від дисертації або наукового звіту, які є суто науковою публікацією, монографії [21] належать до науково-популярної або навчально-методичної літератури. Розділи монографії можуть містити наукові положення, як-от, дисертації, але, водночас, монографія адресована ширшій аудиторії, а не лише фахівцям з теми. Тож має бути адаптована для читача і, за можливості, викладена простішою та зрозумілішою мовою.

Препринти щодо інших наукових, науково-популярних та навчально-методичних публікацій (статей, тез, монографій тощо), фактично є попереднім варіантом наукового видання, яке публікують ще до

випуску офіційного кінцевого варіанту матеріалу. Це роботи без рецензії та публікації в журналі. Їх ніяк не враховують у звітах, статистиці, вони не впливають на наукометричні показники вченого. Електронні препринти розміщені в Інтернеті, називаються "Е-принтом". Із 2017 року препринти введені до профілю Scopus [24]. На відміну від книжок (монографій) для препринту обов'язковими характеристиками є код та установа, де відбуваються відповідні дослідження.

Однак більшість наукових публікацій все ж вимагають попереднього рецензування науковими експертами. Такі рецензії виявляють актуальність досліджень, напрямки розвитку й зауваження до них, і самі є науково змістовними та у поєднанні з об'єктом рецензування значущими елементами наукових знань. Їхні описові характеристики мають формуватися, виходячи з вимог щодо наукового рецензування, та враховувати загальноприйнятну структуру побудови наукової рецензії.

Особливої уваги заслуговують також архівні документи, що складають чималу частину наукових знань та є підґрунтям сучасних знань. Основний міжнародний стандарт архівного опису – ISAD(G) [12]. Даний стандарт вводить правила багаторівневого опису архівних документів, визначає базову термінологію опису та пропонує систему характеристик, організованих у сім груп описової інформації (одиноцею опису в даному випадку є документ будь-якого типу чи група документів, що розглядається як єдине ціле). Слід зазначити, що стандарт спрямований саме на опис архівних документів, тому містить також характеристики, які визначають фонди, архівну інформацію тощо.

Окрему групу наукових публікацій складають освітні матеріали, до яких належать: курси лекцій, матеріали навчальних курсів, підручники, е-підручники, методичні та дидактичні матеріали, презентації, реферати тощо. Частина з перелічених типів ресурсів, таких, як підручники, методичні та дидактичні матеріали тощо, передбачають наявність паперового варіанту. Решта зазвичай існує лише в цифровому представленні. Для освітніх матеріалів, що мають

паперове та цифрове представлення, застосовуватиметься більшість метаданих, що є дійсними для монографій та статей. А для існуючих лише у цифровому представленні, – ті, що визначаються для інтернет-ресурсів, зокрема, е-підручників [19, 20].

Окрім специфіки, пов'язаної з типом відкритого ресурсу, певні особливості можуть бути обумовлені форматом його представлення. Головні три форми представлення наукових документів – це текст, зображення та презентації. Так, зокрема, діаграмам, схемам, фотоматеріалам, рисункам, що можуть бути як окремими повноцінними об'єктами наукових документів, так і приєднаним вкладенням або частиною більш складного об'єкта (наукової роботи), притаманні метадані зображень.

Усі ці особливості обумовлюють визначення специфічних метаданих для різних типів ресурсів у межах групи «науковий документ», разом із загальними описовими характеристиками, притаманними всьому спектру відкритих ресурсів цієї групи.

Отож, система метаданих «відкритих наукових документів» буде базуватися на:

- стандартах загального опису документів, головним з яких досі є DCMII (Dublin Core Metadata Initiative) [3]
- аналізі систем описових характеристик існуючих цифрових бібліотек, зокрема, описів дисертаційних робіт та авторефератів, електронних підручників, монографій тощо, та наборах критеріїв пошуку в них [15-17, 20]
- стандартах із бібліотечної справи та суміжних областей [23],
- загальному аналітичному аналізу існуючих міжнародних дескриптивних стандартів [13],
- аналізі систем індексації публікацій наукометричними базами даних Scopus [18], Web of Science тощо,
- міжнародному стандарті архівного опису – ISAD(G),
- стандартах і схемах метаданих зображень [4], таких як EXIF [5], DCF [6], IPTC, XMP [7 – 9], ISO 16684-2:2014,
- вивченні особливостей системи індексації монографій [22].

Необхідно усвідомлювати, що перелічені типи ресурсів тісно пов'язані з

іншими різновидами відкритих ресурсів, які також можна розглядати як різновиди відкритих наукових знань та з їх описами. Такими ресурсами, наприклад, є наукові видання, фізичні та юридичні особи – науковці тощо. Із побудовою єдиної системи метаданих відкритої науки такі взаємозв'язки дозволять повніше та багатогранніше здійснити опис відкритого наукового ресурсу. Так, наприклад, посилання на автора в метаописі наукової публікації дозволить отримати доступ до вичерпного метаопису даного науковця. Розгорнута система характеристик видання, де опубліковано наукову роботу, також може свідчити щонайменше про її важливість. Це павутиння розширюється зв'язками з відкритими ресурсами інших типів, як-от, наукові заходи, апаратне обладнання тощо.

Метадані відкритих наукових документів

Базуючись на вище викладених аспектах, існуючих стандартах та цифрових системах роботи з інформацією (пошукових системах, цифрових бібліотеках тощо), пропонується схема метаданих наукових документів, класифікованих за спільними та особливими ознаками типів ресурсів, поєднаних поняттям «відкритий науковий документ». Загальні метадані притаманні всім типам інформаційних ресурсів, що розглядаються.

Загальні метадані відкритих наукових документів:

- ідентифікатор інформаційного ресурсу (посилання на ресурс);
- назва ресурсу;
- тип ресурсу (наукова публікація/освітні матеріали/діаграма/модель тощо);
- дата останнього оновлення метаданих інформаційного ресурсу;
- дата реєстрації інформаційного ресурсу;
- служба, що реєструє інформаційний ресурс;
- інформація про власника ресурсу;
 - ПІБ рідною мовою;
 - ім'я та прізвище англійською мовою;
- консультант – особа, до якої можна звертатися за додатковою інформацією;

- інформація про авторів (може співпадати з власником)/співавторів
 - ПІБ рідною мовою;
 - ім'я та прізвище англійською мовою;
 - посада;
 - звання;
 - наукова ступінь;
 - посада, звання, науковий ступінь авторів англійською мовою (переклад згідно з міжнародними стандартами);
 - міжнародні ідентифікатори (коди) вчених:
 - ORCID ID;
 - Scopus ID;
 - Researcher ID;
 - h-індекс;
- додаткові відомості про авторів/співавторів:
 - місце роботи:
 - назва установи;
 - адреса;
 - контактні телефони;
 - електронна адреса;
 - кількість публікацій:
 - в українських виданнях;
 - в зарубіжних виданнях;
- коди рубрикатора – тематика ресурсу, що виражена кодами стандартного переліку тематичних рубрик;
- ключові слова;
- стислий опис (зміст ресурсу, влючно із анотацією чи рефератом, або опис контенту візуальних, аудіо– або мультимедійних ресурсів);
- фінансування – форма фінансування під час створення та введення ресурсу;
- дата створення/подання або останнього оновлення контенту;
- період оновлення;
- чи існує паперовий варіант;
- інформація про інтернет –ресурс, де опубліковано (цифрова бібліотека/електронний журнал тощо);
- інформація про файл ресурсу:
 - розмір;
 - формат;
 - місце зберігання у файловому сховищі;
- посилання на ресурс в мережі;

- метричні показники документу:
 - розмір сторінки;
 - шаблон;
 - кількість сторінок у документі.

Метадані текстових ресурсів:

- мова (для текстового ресурсу);
- УДК (код класифікатора);
- анотація мовою оригіналу;
- анотація англійською мовою;
- ключові слова рідною та англійською мовами;
- список використаних джерел (references):
 - посилання;
 - статус (закрите, відкрите, обмежене);
- відсоток відкритих посилань (open references), тобто доступних усім користувачам усіх API та сервісів Crossref;
- Funder Registry ID (назва та ідентифікатор організації, яка фінансує дослідження та є у спеціальному реєстрі Crossref);
- Funding Award Numbers (інформація щодо фінансування дослідження, наприклад, номеру грантів або премій, пов'язаних з науковим дослідженням);
- чи застосовується сервіс, що відслідковує зміни статті після публікації;
- повнотекстове посилання для сервісу Crossref Similarity Check (перевірка подібності, перевірка на плагіат);
- повнотекстове посилання для антивірусної бази iThenticate;
- посилання у списках літератури (Citation);
- додаткові метричні показники:
 - номери сторінок у журналі/збірнику;
 - кількість використаних джерел;
 - обсяг приєднаного матеріалу:
 - графічного (кількість рисунків);
 - таблиць;
 - формул;
 - схем;
- структура матеріалу (зміст):
 - глави/розділи/лекції;
 - додатки.

Додаткові метадані, специфічні для тез/дисертацій/авторефератів/дипломних робіт:

- на здобуття якого ступеня (кандидатський, докторський, магістерський тощо);
- код спеціальності ВАК;
- керівник досліджень;
- перелік друкованих робіт за темою досліджень;
- інформація про джерела фінансування дослідження (грант, стипендія тощо);
- дата захисту (для дисертації);
- місце захисту (для дисертації);
- дата присвоєння наукового ступеня (для дисертації);
- інформація про практичне впровадження;
- список використаних джерел англійською мовою (references);
- політика доступу;
- посилання на автореферат (дисертації).

Метадані, специфічні для звітів про науково-дослідну роботу:

- тип звіту (анотований, остаточний, річний тощо);
- період виконання досліджень;
- назва проєкту, в межах якого виконуються дослідження;
- кількість наукових робіт, опублікованих у рамках звітнього періоду.

Додаткові метадані для наукових робіт, що мають паперовий варіант:

- видавництво, що опублікувало статтю (не визначається для інтернет-ресурсів);
- інформація про видання/журнал/газету (де опубліковано);
- дата публікації.

Специфічні метадані матеріалів конференцій/семинарів/форумів:

- тип події (конференція, семінар, форум тощо);
- назва події;
- назва матеріалів (збірника матеріалів);
- регулярність проведення;
- дата проведення події;
- місце проведення події:
 - країна;
 - місто;
 - установа;
 - адреса;
- загальна наукова спрямованість;
- інформація про наукові розділи:

- назва розділу;
- кількість доповідей в розділі;
- інформація про доповіді:
 - назва;
 - інформація про доповідача;
 - відомості про співавторів;
 - тип доповіді (стендова, повідомлення, повноцінна тощо);
 - анотація;
 - тези доповіді;
 - посилання на повний текст доповіді;
 - посидання на аудіозапис/відеозапис доповіді;
 - приєднані матеріали:
 - наявність;
 - тип (презентації, додатки, роздаткові матеріали, стендові матеріали тощо);
 - назва;
 - посилання на приєднані матеріали.

Специфічні метадані для публікацій (монографій/підручників/препринтів):

- рецензенти:
 - посада;
 - вчене звання;
 - науковий ступінь;
 - місце роботи;
 - ПІБ;
- рекомендація до видання;
- знак охорони авторського права;
- унікальний ідентифікатор DOI (роботи в цілому або розділів монографії).

Метадані, специфічні лише для препринтів:

- код препринту;
- установа.

Додаткові метадані, притаманні для освітніх матеріалів:

- тип матеріалу (підручник, методичні вказівки тощо);
- тип представлення (текст/ презентація/відео тощо);
- міністерство (чи інша керуюча установа);
- назва навчального закладу;
- до якого освітнього курсу належить;
- вихідні відомості.

Метадані, специфічні для рецензій:

- назва рецензованої роботи;

- посилання на науковий матеріал, що рецензується;
- дата рецензії;
- дані про автора рецензії:
 - посада;
 - вчене звання;
 - науковий ступінь;
 - місце роботи;
 - ПІБ;
- висновок (позитивний/негативний);
- посилання на повний текст рецензії;
- стислий зміст:
 - переваги рецензованої роботи;
 - зауваження до роботи.

Метадані, специфічні для архівних документів:

- характеристики ідентифікації:
 - рівень структури документа (наприклад, збірка робіт, матеріали конференції чи стаття, опис колекції чи конкретного експонату);
 - носій, де зберігається документ;
- характеристики контексту:
 - біографічні дані автора;
 - архівна історія документу (для архівних документів);
 - безпосереднє джерело комплектування або переклад (джерело, звідки було отримано документ, або дата та/або спосіб комплектування, якщо ця інформація частково або повністю конфіденційна);
- характеристики змісту та структури:
 - рамки (період часу, географія тощо) та зміст (форма документа, тема тощо) документа відповідно до рівня опису;
 - інформація щодо оцінювання, знищення документу або планування відповідних дій;
 - подальші поповнення (доповнення) одиниці опису;
- характеристики доступу та використання:
 - правила доступу (правовий статус або інші правила, що регламентують доступ до документу);
 - умови відтворення документу;
 - фізичні характеристики та технічні умови (апаратні та програми, необхідні для доступу до документу, вимоги до зберігання тощо);

- науково-довідковий апарат (посилання на наявні довідкові засоби, які містять інформацію із контексту та змісту документу);
- характеристики взаємопов'язаних матеріалів:
 - наявність та місце розміщення оригіналу документу;
 - наявність та місцезнаходження копій;
 - пов'язані документи;
 - примітки про наявні публікації (публікації, що присвячені даній одиниці опису або засновані на її використанні);
- примітки (інформація про документ, що не може бути подана в інших характеристиках);
- характеристики контролю опису:
 - примітки архівіста (як і ким було створено опис документу);
 - правила, за якими створений опис;
 - дати створення опису.

Метадані зображень:

- розмір зображення в пікселях;
- тип зображення (цифрова фотографія, рисунок, схема, діаграма, мапа);
- опис зображення;
- коментарі користувача;
- права доступу до інформаційного об'єкту;
- приєднане аудіо (якщо є, ідентифікатор, назва та посилання);
- приєднана анотація (якщо є, ідентифікатор, назва та посилання);
- кольорове/чорно-біле;
- характеристики, притаманні лише цифровим фотографіям [10]:
 - параметри фотокамери, з якої було зроблено фотографію (з повним переліком можна ознайомитися в описі EXIF-схеми):
 - модель;
 - виробник;
 - діафрагма;
 - витримка;
 - швидкість;
 - експокорекція;
 - фокусна відстань;
 - світлосила;
 - експозамір;

- відстань до об'єкта;
- спалах режим;
- спалах енергія;
- тощо;
- геолокація, де була зроблена фотографія;
- семантичні текстові анотації:
 - текстові мітки, приєднані до зображення;
 - назва рисунку, що виводиться на самому зображенні чи приєднана до нього;
 - розширений опис;
 - опис основних елементів зображення;
- характеристики візуалізації:
 - набір візуальних слів [11].

Метадані, що специфічні для презентацій:

- кількість слайдів;
- шаблон презентації, що використовується;
- розширений опис;
- пов'язаний файл доповіді;
- подія, до якої стосується презентація;
- спецефекти, що використовуються для слайд-шоу в цілому;
- метадані слайдів:
 - номер слайду;
 - тема слайду;
 - текстові мітки, приєднані до слайду;
 - рисунки/діаграми/таблиці, вставлені в слайд;
 - використані спецефекти.

Висновки

Мета дослідження полягає у розробці єдиної схеми метаданих ресурсів відкритої науки. Зокрема, фокусом даної роботи є визначення системи характеристик, які описують загальні та специфічні ознаки різних типів наукових документів як значущої частини наукових знань відкритої науки.

В результаті досліджень:

- 1) здійснено аналіз головних понять, визначень та потреб відкритої науки;
- 2) на базі цього аналізу виокремлені ресурси відкритої науки, що фор-

мують клас «відкриті наукові знання», та запропоновано таксономію цих ресурсів;

- 3) здійснено аналіз ресурсів у межах визначеної ієрархії «відкриті наукові знання» та базі цього аналізу запропоновано систему схем метаданих таких ресурсів.

Напрямки подальших досліджень полягають у вдосконаленні запропонованих схем метаданих і проведенні всебічного аналізу програмних засобів створення, управління та інтеграції метаданих для ресурсів відкритої науки та визначення можливості й ефективності їх застосування для тих чи інших груп ресурсів.

Література

1. <https://www.fosteropenscience.eu/content/open-science-training-handbook>
2. <https://hi-news.pp.ua/tehnka-tehnologyi/7373-metadan-ce-viznachennya-vidi-ta-varantivikoristannya-metadanih-u-prikladnomu-programuvann.html>
3. <https://www.dublincore.org/>
4. <https://habr.com/ru/post/93119/>
5. <https://www.exif.org/category/specifications>
6. <http://exif.org/dcf.PDF>
7. <https://helpx.adobe.com/after-effects/using/xmp-metadata.html>
8. ISO 16684-1:2012, Graphic technology – Extensible metadata platform (XMP) specification – Part 1: Data model, serialization and core properties
9. <https://www.adobe.com/devnet/xmp.html>
10. <https://www.ixbt.com/digimage/metadxph.shtml>
11. J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In Proc. of 9th IEEE Int'l Conf. on Computer Vision, Vol. 2, 2003.
12. 1999_ISAD_G.pdf (archives.gov.ua)
13. <https://undiasd.archives.gov.ua/doc/aodeskreptyvni%20standarty.pdf>
14. <https://www.loc.gov/marc/>
15. <https://nim.media/articles/shcho-take-metadani-statti-yak-voni-vplivayut-na-yiyi-indeksatsiyu>
16. <http://openscience.in.ua/key-metadata-elements.html>
17. <http://www.4author.com/ru/blog/story/metadannye-nauchnoi-stati-osnova-raboty-kotoruiubudut-tcitirovat>
18. http://www.library.univ.kiev.ua/ukr/res/scopus_author.pdf
19. <https://dspace.uzhnu.edu.ua/jspui/handle/lib/38156>
20. Каталоги – НБУВ Національна бібліотека України імені В. І. Вернадського (irbis-nbuv.gov.ua)
21. https://ukrlogos.in.ua/ua_monographie.php
22. <https://isg-konf.com/uk/indeksatsiya-monografij-ukr/>
23. http://www.nbuv.gov.ua/sites/default/files/basicpage_files/202107_basicpage_files_mat/perelik_standartiv-2021.07.01.pdf
24. <https://nim.media/articles/nova-funktsiya-v-scopus---preprinti>

References

1. <https://www.fosteropenscience.eu/content/open-science-training-handbook>
2. <https://hi-news.pp.ua/tehnka-tehnologyi/7373-metadan-ce-viznachennya-vidi-ta-varantivikoristannya-metadanih-u-prikladnomu-programuvann.html>
3. <https://www.dublincore.org/>
4. <https://habr.com/ru/post/93119/>
5. <https://www.exif.org/category/specifications>
6. <http://exif.org/dcf.PDF>
7. <https://helpx.adobe.com/after-effects/using/xmp-metadata.html>
8. ISO 16684-1:2012, Graphic technology – Extensible metadata platform (XMP) specification – Part 1: Data model, serialization and core properties
9. <https://www.adobe.com/devnet/xmp.html>
10. <https://www.ixbt.com/digimage/metadxph.shtml>
11. J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In Proc. of 9th IEEE Int'l Conf. on Computer Vision, Vol. 2, 2003.
12. 1999_ISAD_G.pdf (archives.gov.ua)
13. <https://undiasd.archives.gov.ua/doc/aodeskreptyvni%20standarty.pdf>
14. <https://www.loc.gov/marc/>
15. <https://nim.media/articles/shcho-take-metadani-statti-yak-voni-vplivayut-na-yiyi-indeksatsiyu>
16. <http://openscience.in.ua/key-metadata-elements.html>
17. <http://www.4author.com/ru/blog/story/metadannye-nauchnoi-stati-osnova-raboty-kotoruiubudut-tcitirovat>
18. http://www.library.univ.kiev.ua/ukr/res/scopus_author.pdf
19. <https://dspace.uzhnu.edu.ua/jspui/handle/lib/38156>

20. Catalogs – NBUV National Library of Ukraine named after V. I. Vernadskyi (irbis-nbuv.gov.ua)
21. https://ukrlogos.in.ua/ua_monographie.php
22. <https://isg-konf.com/uk/indeksatsiya-monografij-ukr/>
23. http://www.nbuv.gov.ua/sites/default/files/basicpage_files/202107_basicpage_files_mat/perelik_standartiv-2021.07.01.pdf
24. <https://nim.media/articles/nova-funktsiya-v-scopus---preprinti>

Одержано: 24.10.2023

Про автора:

Захарова Ольга Вікторівна,
кандидат технічних наук,
старший науковий співробітник.
Кількість наукових публікацій
в українських виданнях – 35.
<http://orcid.org/0000-0002-9579-2973>.

Місце роботи автора:

Інститут програмних систем
НАН України,
проспект Академіка Глушкова, 40.
Тел.: 526 5139.
E-mail: ozakharova68@gmail.com.
Моб.тел.: +38(068)594756