

Г.Ю. Проскудіна, К.О. Кудім

ГЛОБАЛЬНІ СЛУЖБИ АГРЕГАЦІЙ РЕСУРСІВ ВІДКРИТОГО ДОСТУПУ ТА ЇХ ВИМОГИ ДО ПОСТАЧАЛЬНИКІВ ДАНИХ

У роботі розглянуті сучасні глобальні агрегатори документів відкритого доступу BASE, OpenAIRE та CORE. Проаналізовані їхні кількісні характеристики, такі як кількість зібраних описів документів та повних текстів, кількість постачальників даних, наявність інтерфейсу прикладного програмування для отримання даних. Розглянуто склад і види їх постачальників даних, такі як інституційні репозитарії, відкриті журнали, видавництва, наукові репозитарії препринтів, тематичні електронні бібліотеки, а також системи, які в свою чергу теж є агрегаторами. Досліджено також, яку саме інформацію про документи збирають ці агрегатори, як вона представлена в інтерфейсі користувача, а також яка інформація збирається про постачальників даних, і теж, яким чином вона представлена у інтерфейсі користувача. Як відбувається взаємодія агрегатора з постачальниками даних, які протоколи обміну даних підтримуються, з якою частотою відбувається оновлення зібраних даних. Також сучасні агрегатори на базі зібраних корпусів даних, використовуючи методи машинного навчання, методи бібліометрії, вебометрики, альтиметрії, семантометрії надають цілий ряд корисних сервісів науковцям. Ми як розробники низки наукових електронних бібліотек з відкритим доступом вже зареєстровані як провайдери даних у цих системах. Тому знайомі з їхніми вимогами у практичній площині. В цій роботі ми спробували дещо узагальнити ці вимоги, а саме – підтримка стандартних протоколів передачі даних, підтримка стандартних наборів метаданих, підтримка стандартних кодувань символів для вмісту та інші.

Ключові слова: відкритий доступ, провайдер сервісів, провайдер даних, протокол OAI-PMH.

G.Yu. Proskudina, K.O. Kudim

GLOBAL OPEN ACCESS RESOURCE AGGREGATION SERVICES AND THEIR REQUIREMENTS FOR DATA PROVIDERS

The paper looks at the modern global aggregators of open access documents BASE, OpenAIRE and CORE. Their quantitative characteristics are analysed, such as the number of collected document descriptions and full texts, the number of data providers, and the availability of an application programming interface for data retrieval. The composition and types of their data providers, such as institutional repositories, open journals, publishing houses, scientific repositories of preprints, thematic digital libraries, as well as systems that are also aggregators, are considered. We also investigate what kind of information about documents these aggregators collect and how it is presented in the user interface, as well as what information is collected about data providers and how it is presented in the user interface. How the aggregator interacts with data providers, what data communication protocols are supported, and how often the collected data is updated. Also, modern aggregators based on collected data corpora, using machine learning methods, bibliometrics, webometrics, altmetrics, semantometrics, provide a range of useful services to researchers. As developers of a number of open access scientific digital libraries, we are already registered as data providers in these systems. Therefore, we are familiar with their requirements in practice. In this article, we have tried to summarise these requirements to some extent, namely: support for standard data transfer protocols, support for standard metadata sets, support for standard character encodings for content, and others.

Key words: open access, service provider, data provider, OAI-PMH protocol.

Вступ

Під час виконання частини проекту НАНУ «Відкрита наука» були проведені дослідження із створення системи інтеграції або харвестера ресурсів із різних від-

критих академічних джерел з метою отримання зручного і потужного інструменту пошуку та доступу до інформації для користувачів. А також інтеграції цих зібра-

них наукових ресурсів України у світовий/європейський науковий інформаційний простір.

Як правило, перед системою інтеграції стоять три основні задачі:

1. Збір і інтеграція метаданих із різних джерел електронних ресурсів.

2. Організація пошуку і видачі відповідних ресурсів.

3. Передача метаданих із власного харвестеру іншим харвестерам.

Проаналізувавши низку програмних систем, на платформі яких можна розгорнути агрегатор наукових робіт, ми зупинились на системі VuFind, про яку розповіли в роботі [1]. Наразі ми вже розгорнули таку систему¹ і почали її експлуатацію. Наразі до неї підключено близько десяти діючих провайдерів даних, з яких було зібрано близько 400 тис метаданих документів, в основному це – статті. Ми індексуємо метадані документів усіх видів академічно релевантних ресурсів – таких, як журнали, інституційні репозитарії, цифрові колекції тощо, які надають інтерфейс OAI та використовують протокол Open Archives Initiative Protocol for Metadata Harvesting² (OAI-PMH) для надання свого контенту [2] як одного з найбільш використовуваних на даний час варіантів інтеграції. Проіндексовані дані зберігаються на серверах Інституту програмних систем НАНУ. Апробовано підключення та передачу метаданих з VuFind до іншого агрегатора. Вивчені можливості VuFind із перегляду, пошуку та завантаження статей та підготовлена інструкція для кінцевого користувача. Розроблені схеми метаданих основних видів наукових інформаційних ресурсів НАНУ.

Зараз продовжується налагодження цієї системи, відпрацьовуються алгоритми завантаження та оновлення даних, регламент роботи агрегатора, вимоги та інструкції для провайдерів даних по підключенню до агрегатора. Тому вивчення світового досвіду з експлуатації таких агрегаторів та їхніх вимог до своїх постачальників даних

є необхідним кроком доведення нашої системи до сучасних зразків.

Ми як розробники низки наукових електронних бібліотек з відкритим доступом вже зареєстровані провайдерами даних у таких системах як BASE³, OpenAIRE⁴, CORE⁵. Тому знайомі з вимогами провайдерів сервісів до провайдерів даних у практичній площині. У цій роботі ми спробували дещо узагальнити ці вимоги та вказівки щодо найкращого налаштування своїх репозитаріїв для індексування у таких агрегаторах.

Приклади сучасних глобальних агрегаторів документів відкритого доступу

BASE (Bielfield Academic Search Engine) – глобальна служба збору метаданих. Вона збирає репозитарії та журнали через OAI-PMH і надає зібраний вміст через API та набір даних. Наразі агрегатор налічує близько 400 млн описів документів від понад 11 тис постачальників даних. Наукова електронна бібліотека періодичних видань НАН України⁶, яку ми створили і підтримуємо, підключена до цього агрегатора як провайдер даних з січня 2012 року. Тут^{7,8} наведені приклади, яким чином у цьому агрегаторі представлені наші статті та сам провайдер даних.

OpenAIRE – це мережа постачальників даних відкритого доступу, які підтримують політики відкритого доступу. Незважаючи на те, що в минулому проєкт був зосереджений на європейських репозитаріях, нещодавно він розширився за рахунок інституційних і предметних репозитаріїв з-за меж Європи. Основним напрямком OpenAIRE є допомога Європейській Раді в моніторингу дотримання її політики відкритого доступу. Дані OpenAIRE доступні

¹<https://harvester.nas.gov.ua/>

²<https://www.openarchives.org/OAI/openarchivesprotocol.html>

³<https://base-search.net/>

⁴<https://www.openaire.eu/>

⁵<https://core.ac.uk/>

⁶<http://dspace.nbuv.gov.ua>

⁷<https://www.base-search.net/Search/Results?q=coll:ftvamadskynl&refid=dctableen>

⁸https://www.base-search.net/about/en/about_source.php?id=2328

через API. Наша Наукова електронна бібліотека періодичних видань зареєстрована у цій системі з 2018 року з рівнем сумісності v2.0 з інструкціями OpenAIRE (Рис. 1). На сьогодні стоїть питання щодо оновлення нашої системи до рівня сумісності v3.0 або v4.0. Оскільки нова інтегрована платформа каталогу EOSC Portal Catalog та Marketplace тепер обслуговуватиме лише зареєстровані джерела даних OpenAIRE, які сумісні з версіями 3.0 та 4.0 інструкцій OpenAIRE. Це дасть нам змогу отримувати додаткові послуги. Наприклад з'явиться можливість отримання звітів про використання досліджень. OpenAIRE збирає дані про використання, а потім об'єднує їх і надає стандартизовані звіти.

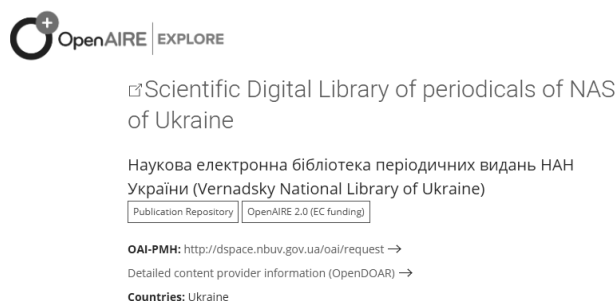


Рис. 1. Наукова електронна бібліотека періодичних видань НАНУ у агрегаторі OpenAIRE

CORE (COncnecting REpositories) об'єднує наукові статті у відкритому доступі від тисяч постачальників даних з усього світу, включаючи інституційні та предметні репозитарії, журнали відкритого доступу та гібридного доступу. CORE є найбільшою колекцією літератури з відкритого доступу – на момент написання цієї статті вона забезпечує єдину точку доступу до наукової літератури, зібраної від понад одинадцяти тисяч постачальників даних з усього світу, і ця колекція постійно зростає. Вона надає кілька способів доступу до своїх даних як для користувачів, так і для машин, включаючи безкоштовний API і повний дамп своїх даних [3]. На наш погляд, CORE – найцікавіший агрегатор із точки зору застосування у ньому низки семантичних та наукометричних технологій, наприклад, [4,5].

Наразі ми маємо дві наші бібліотеки, зареєстровані у CORE. Це – вищезга-

дана Наукова бібліотека періодичних видань НАНУ⁹, що працює на програмному забезпеченні DSpace¹⁰, і електронна бібліотека Інституту програмних систем НАНУ¹¹, що працює на програмному забезпеченні EPrints¹².

Майже всі ці служби агрегації документів відкритого доступу мають більш стандартні процедури реєстрації та вимоги до провайдерів даних.

Підтримка протоколу OAI-PMH

Обов'язково переконайтеся, що ваш репозиторій видимий через OAI-PMH. Майже всі сучасні системи агрегації публікацій використовують протокол OAI-PMH для регулярного збирання та оновлення інформації, яку вони зберігають про своїх постачальників даних. Щоб пройти індексацію, постачальники даних повинні надати інформацію про свої ресурси через OAI-PMH.

Відносно поширеною проблемою є те, що кінцева точка OAI-PMH постачальника даних неправильно налаштована або не функціонує. Це може статися навіть тоді, коли інші функції репозитарію працюють без проблем. Тому важливо, щоб репозитарії стежили за тим, аби їхня кінцева точка OAI-PMH залишалася функціональною. Це можна зробити кількома способами.

Постачальники даних, які розглядають можливість реєстрації в агрегаторі. Можна використовувати інструмент перевірки Open Archives OAI-PMH¹³. Інструмент очікує базову URL-адресу OAI-PMH і проводить низку перевірок, щоб оцінити коректність конфігурації кінцевої точки (Рис. 2). Якщо валідатор повертає помилки, їх потрібно буде виправити до реєстрації в агрегаторі. Рекомендується виконати цей крок до спроби реєстрації в агрегаторі.

⁹ <https://core.ac.uk/data-providers/944>

¹⁰ <https://dspace.lyrasis.org/>

¹¹ <https://core.ac.uk/data-providers/953>

¹² <https://www.eprints.org/uk/>

¹³ <https://www.openarchives.org/Register/ValidateSite>

Реєстрація постачальників даних у агрегаторі. Відвідайте сторінку реєстрації постачальника даних та введіть URL-адресу OAI-PMH свого репозиторію. Якщо URL-адреса OAI-PMH є дійсною, репозиторій буде зареєстровано за умови проходження додаткових перевірок, описаних

нижче. Важливо зазначити, що перевірка правильності конфігурації кінцевої точки OAI-PMH не гарантує правильне відображення всіх записів метаданих із системи репозиторію. Це можна перевірити лише після того, як агрегатор спробує отримати дані з репозиторію.

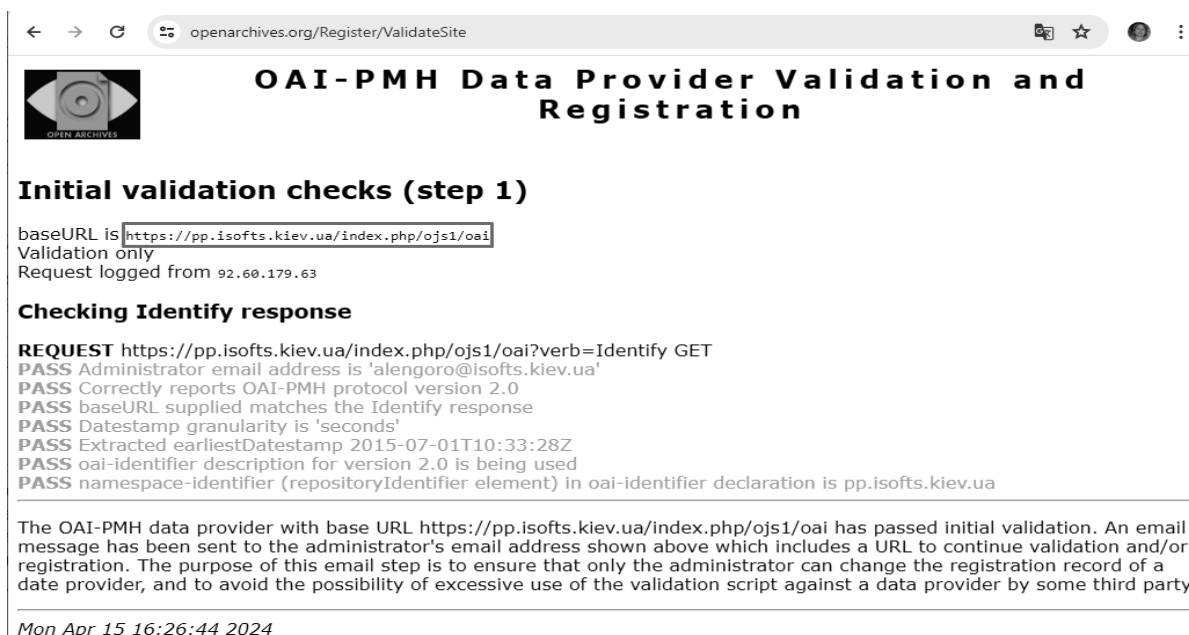


Рис. 2. Зразок перевірки OAI-PMH для журналу Проблеми програмування

Постачальники даних, які вже зареєстровані в агрегаторі. Зареєстровані постачальники даних можуть отримати доступ до інформаційної панелі агрегатора. Панель управління надає доступ до звіту про збір даних, який містить таку інформацію, як останній випадок успішного збору даних, кількість знайдених та проіндексованих метаданих і повнотекстових записів (якщо повні тексти збираються агрегатором), будь-які помилки або попередження, що виникли в процесі збору даних.

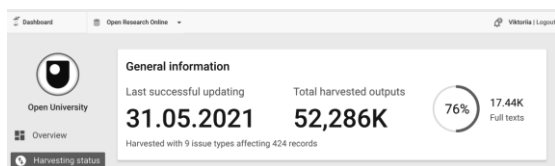


Рис. 3. Приклад показу збору даних агрегатором CORE на інформаційній панелі для репозитарію Відкритого університету¹⁴

Нерідко кількість проіндексованих записів метаданих у агрегаторах може де-що відрізнятись від кількості, яку постачальник даних може бачити у своїй системі репозитарію, оскільки інформаційна панель надає незалежний зовнішній погляд на репозитарій. Відхилення можуть виникати через низку причин. Наприклад, через те, що ваш репозиторій не відкриває всі записи через інтерфейс OAI-PMH, деякі записи відключені або їхні метадані не відповідають мінімальним вимогам щодо якості. Тому постачальникам даних рекомендовано регулярно перевіряти і виправляти будь-які помилки та попередження, що з'являються на інформаційній панелі.

Реєстрація у відкритих міжнародних реєстрах

Рекомендована реєстрація репозитарію у відкритих міжнародних реєстрах. Агрегатори використовують визнані відкриті реєстри репозитаріїв, зокрема, та-

¹⁴<https://www.open.ac.uk/>

кі як OpenDOAR¹⁵, для виявлення нових постачальників даних із відкритим доступом. Тому бажано, щоб ваш репозитарій був включений до такого міжнародного списку репозитаріїв, і щоб базова URL-адреса OAI-PMH підтримувалася в актуальному стані в цьому реєстрі. Наприклад, для реєстрації будь-якого провайдера даних у агрегаторі OpenAIRE, попередня реєстрація у реєстрі OpenDOAR є обов'язковою вимогою. Нерідко зустрічаємо і таку ситуацію, коли одні агрегатори автоматично індексують інші агрегатори, як, наприклад, серед провайдерів даних агрегатора CORE є агрегатор DOAJ (Каталог журналів відкритого доступу), кілька серверів препринтів та наукових праць, які мають ідентифікатори публікацій, зареєстрованих у Crossref [3].

Конфігурація метаданих

Для того, щоб агрегатор міг успішно збирати дані зі сховища, як було сказано вище, сховище повинно мати кінцеву точку, сумісну зі стандартом OAI-PMH. Більшість поширених програм, таких як EPrints, DSpace або Open Journal Systems¹⁶ (OJS) підтримують OAI-PMH. Крім того, для успішного індексування важливо, щоб метадані записів у репозитарії були правильними та у підтримуваному форматі.

Обов'язкова підтримка стандартів метаданих. Агрегатори можуть підтримувати кілька форматів метаданих для опису наукових документів [1]. Як правило, метадані мають бути представлені в одному з наступних форматів.

1. Dublin Core / Extended Dublin Core (мінімум). Дублінське ядро є однією з найпростіших і найпоширеніших схем метаданих. Розширена версія була формалізована у вигляді термінів метаданих DCMI¹⁷ 2019 року, з приблизно 70 полями даних. Хоча Dublin Core та Extended Dublin Core є достатніми для індексації, зокрема, в агрегаторі BASE, схема надає лише обмежені можливості для опису нау-

кових ресурсів порівняно з OpenAIRE Guidelines¹⁸ та Rioxx¹⁹.

2. OpenAIRE Guidelines була створена для підтримки стратегії відкритого доступу Європейської Комісії та для задоволення вимог інфраструктури OpenAIRE. Ця нова версія настанов, відповідно до розширення цілей ініціативи OpenAIRE та її інфраструктури, має ширшу сферу застосування. Впроваджуючи ці настанови, менеджери репозитаріїв дозволять авторам, які розміщують публікації в їхніх репозитаріях, відповідати вимогам Європейської Комісії щодо відкритого доступу.

3. RIOXX. Агрегатор CORE рекомендує репозитаріям використовувати формат метаданих Rioxx (the research outputs metadata schema) як найбільш придатний профіль метаданих для опису результатів наукових досліджень. Ця схема метаданих була розроблена для інституційних репозитаріїв з метою обміну метаданими про наукові ресурси, які вони містять. Спочатку розроблена для задоволення вимог до звітності Дослідницьких рад Великої Британії (Research Councils UK, RCUK)²⁰, Rioxx також виявилася корисною як стандарт для обміну метаданими між репозитаріями та мережевими сервісами, такими як великомасштабні агрегатори метаданих, наприклад, CORE. Rioxx фокусується на застосуванні узгодженості до полів метаданих, що використовуються для ідентифікаторів наукових результатів, людей і організацій, спонсорів досліджень і проєктів/грантів, і призначений для підтримки послідовного відстеження наукових публікацій з відкритим доступом у різних наукових системах. Rioxx має кілька переваг над іншими схемами. Зокрема, ця схема була розроблена з особливим акцентом на забезпеченні ефективного обміну даними між репозиторіями і машинними агентами. Це робить збір даних швидшим і точнішим завдяки уникненню неоднозначності у зв'язках метаданих, що описують ресурси такі, як повні тексти, набори даних і програмне забезпечення, а також самі ресурси.

¹⁵<http://v2.sherpa.ac.uk/opensoar/>

¹⁶<https://pkp.sfu.ca/ojs>

¹⁷<https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

¹⁸<https://guidelines.openaire.eu/>

¹⁹<https://rioxx.net/>

²⁰<https://www.ukri.org/>

Вона уможливорює опис широкого спектру наукових властивостей, корисних для звітування та аналізу досліджень, таких як ідентифікатори грантів, ідентифікатори проєктів, ідентифікатори спонсорів, інформація про ліцензування тощо.

Схема Rіоxx сумісна з іншими протоколами машинного доступу до наукових документів, зокрема, Signposting²¹ та ResourceSync²², і логічно інтегрується з ними.

Вона забезпечує механізми для розрізнення ресурсів, що перебувають під управлінням репозитарію, від ресурсів під зовнішнім управлінням, що дозволяє CORE надавати перевагу посиланням на репозитарій, де це можливо (і доречно), замість зовнішніх посилань на веб-сайти видавців.

CORE надає валідатор²³ метаданих для постачальників даних, які підтримують Rіоxx.

Повнотекстова конфігурація

Як уже було сказано вище, низка потужних можливостей деяких агрегаторів заснована на здатності індексувати повнотекстовий контент. До них належать повнотекстовий пошук, рекомендації повнотекстового контенту, виявлення версій і дубльованого контенту та інші. Підтримувані повнотекстові формати – тільки дійсні файли PDF, doc або docx, які містять текст, що витягується. Якщо PDF-файл це – відскановане зображення, то використання документа в агрегаторах буде обмежено. Для успішного індексування повного тексту статті, на нього має бути посилання в один з наступних способів.

Пряме посилання на повний текст. Рекомендується, щоб постачальники даних надавали однозначне посилання на повний текст у метаданих. Правильна стратегія посилання залежить від використовуваної схеми метаданих. Там, де використовується Дублінське ядро, рекоменду-

ється надавати пряме посилання на повний текст у першому входженні dc:identifier.

```
<dc:identifier>http://oro.open.ac.uk/37823/1/jcdl2019_v7.pdf</dc:identifier>
```

Система EPrints за замовчуванням дотримується цієї рекомендації. Такий підхід значно зменшує навантаження, яке агрегатор накладає на сховище під час індексації. Якщо це неможливо, то агрегатор може індексувати вміст, якщо в метаданих є чітко визначене посилання, яке вказує на машинодоступний документ (або PDF, або формат Microsoft Word doc/docx).

Непряме посилання на повний текст. Агрегатор автоматично визначає, коли постачальник даних не вказує посилання на повний текст безпосередньо, як було описано вище. У таких випадках індексатор відвідає цільову сторінку і збере з неї посилання так само, як це зробив би користувач, намагаючись знайти посилання на документ на сторінці.

Щоб переконатися, що правильний повний текст був знайдений, тобто повний текст, що відповідає запису метаданих, агрегатор запускає процес зіставлення заголовка запису із заголовком, що міститься в pdf-файлі. Цей процес не є на 100% точним, спричиняє додаткове навантаження на сервер і вимагає більше часу на обробку одного документа. Це не працює для документів, які потребують розпізнавання тексту, і тому вони будуть відкинуті.

Вимоги щодо кодування символів контенту та інші

Кодування символів. Увесь вміст інтерфейсу OAI-PMH (заголовки, імена авторів, анотації) кодується в UTF-8. Інші кодування або дублювання кодувань спричиняють помилки у відображенні результатів пошуку з вашого джерела.

Розділення кількох записів у полі метаданих. Якщо у полі метаданих вказані кілька записів (наприклад, ім'я автора та його ідентифікатор ORCID), розділяйте їх пробілами, крапкою з комою та пробілами. Таке розділення дозволяє індексувати інформацію окремо і зробити її доступною для пошуку.

²¹<https://signposting.org/>

²²<https://www.openarchives.org/rs/1.1/resourcesync>

²³<https://core.ac.uk/membership-documentation#rioxx-validator>

Розмітка FAIR рекомендується.

Сторінки, на яких розміщені наукові статті, розроблені з думкою про людину, ці сторінки не оптимізовані для використання машинними агентами, які здійснюють навігацію в науковій мережі. Розмітка задовольняє потреби машинних агентів, надаючи метадані, посилання на документи відповідно до стандартів, використовуючи типізовані посилання²⁴ як засіб для пояснення закономірностей, що повторюються на наукових порталах.

Як адміністратор платформи репозитарію, на якій розміщуються результати досліджень, ви можете допомогти машинам орієнтуватися в інформаційних системах, якими ви керуєте, впровадивши профіль вказівників FAIR²⁵, який надає конкретні рекомендації, спрямовані на уніфікацію в різних системах.

Висновки

В роботі зроблено огляд доступних і популярних агрегаційних служб публікацій відкритого доступу BASE, OpenAIRE, CORE. А також проаналізовані і узагальнені їх вимоги та рекомендації щодо індексування вмісту постачальників даних. Наші електронні бібліотеки, які ми створили і підтримуємо протягом уже майже 20 років, працюють на основі програмного забезпечення з відкритим кодом Dspace, EPrint та OJS, вже присутні у цих агрегаторах як провайдери даних. Тому деякий практичний досвід ми маємо, і це дозволяє нам виконати проєкт НАНУ «Відкрита наука» із створення агрегатора публікацій, який у подальшому буде підключено до більш потужних європейських або світових агрегаторів публікацій відкритого доступу.

Матеріали цієї статті будуть корисні тим, хто планує підключити свої ресурси до нашої системи. А також будуть використані у програмному налаштуванні нашого агрегатора наукових публікацій та документації із його адміністрування та регламенту роботи.

²⁴<https://www.iana.org/assignments/link-relations/link-relations.xhtml>

²⁵<https://signposting.org/FAIR/>

Література

1. Проскудіна Г.Ю., Кудім К.О., Резніченко В.А. VUFIND: відкрите рішення для інтеграції бібліотечних колекцій // Проблеми програмування. – 2023. – № 4 – С. 15–26. <https://pp.isoftware.kiev.ua/ojs1/article/view/590>
2. 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. <http://dspace.nbuv.gov.ua/handle/123456789/291>
3. Кнот П., Херрманнова Д., Канселлієрі М. та ін. CORE: Глобальна служба агрегації документів відкритого доступу. Nature Scientific Data 10, 366 (2023). <https://www.nature.com/articles/s41597-023-02208-w>
4. Ammar, W. et al. Construction of the Literature Graph in Semantic Scholar. Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. – 2018 – Vol. 3 – pp. 84–91.
5. Кнот П., Херрманнова Д. К вопросу о семантометрии: новый критерий для оценки вклада научной публикации на основе семантического сходства // Международный форум по информатизации. – 2015. – Т.40, № 1. – С. 3-8.

References

1. Proskudina G.Yu., Kudim K.A., Reznichenko V.A. 2023. VuFind: an open solution for integrating library collections. Problems in programming. no. 4, pp. 15–26. (in Ukrainian). Available from: <https://pp.isoftware.kiev.ua/ojs1/article/view/590> [Accessed 17/04/2024].
2. Reznichenko V.A., Novytskyi O.V., Proskudina G.Yu., 2007. OAI-PMH protocol-based integration of scientific digital libraries. Problems in programming, no. 2, pp. 97–112. (in Ukrainian). Available from: <http://dspace.nbuv.gov.ua/handle/123456789/291> [Accessed 17/04/2024].
3. Knoth P., Herrmannova D., Cancellieri M. et al. CORE: A Global Aggregation Service for Open Access Papers. Sci Data 10, 366 (2023). <https://doi.org/10.1038/s41597-023-02208-w> [Accessed 17/04/2024].

4. Ammar, W. et al. Construction of the Literature Graph in Semantic Scholar. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 3 (Industry Papers): 84–91 (2018). [Accessed 17/04/2024].
5. Petr Knoth and Drahomira Herrmannova. Towards Semantometrics: A New Semantic Similarity Based Measure for Assessing a Research Publication's Contribution. D-Lib Magazine. Volume 20, Number 11/12 (2014) <https://www.dlib.org/dlib/november14/knoth/11knoth.html> [Accessed 18/04/2024].

Одержано: 29.04.2024

Внутрішня рецензія отримана: 08.05.2024

Зовнішня рецензія отримана: 08.05.2024

Про авторів:

Проскудіна Галина Юріївна,
науковий співробітник.
<http://orcid.org/0000-0001-9094-1565>

Кудім Кузьма Олексійович,
молодший науковий співробітник.
<http://orcid.org/0000-0001-9483-5495>

Місце роботи авторів:

Інститут програмних систем
НАН України,
тел. +38 (050) 368 49 27
E-mail: guproskudina@gmail.com,
kuzmaka@gmail.com