

<https://doi.org/10.62405/osiconf.2026.10>

Жук Олександр,

науковий співробітник

Видавничого дому «Академперіодика»

Національної академії наук України, Київ, Україна,

ORCID: 0000-0001-9631-3906, e-mail: o.m.zhuk@nas.gov.ua

Гавриш Андрій,

кандидат технічних наук, доцент,

старший науковий співробітник Видавничого дому «Академперіодика»

Національної академії наук України, Київ, Україна,

ORCID: 0000-0001-6474-6803, e-mail: Gavrysh@nas.gov.ua

РОЗРОБКА АРХІТЕКТУРИ ТА МЕТОДОЛОГІЧНИХ ЗАСАД СИСТЕМИ СЕМАНТИЧНОЇ РЕТРОКОНВЕРСІЇ НАУКОВИХ ВИДАНЬ НАН УКРАЇНИ ДЛЯ ПРОСТОРУ OPEN SCIENCE

АНОТАЦІЯ. Дослідження спрямовані на збереження масиву наукової спадщини НАН України та її інтеграцію в сучасне цифрове наукове середовище. Пропонується комплексне методологічно-технологічне рішення (стандартизовану методику оцифрування, опису та семантичного збагачення наукових видань, включно з формуванням верифікованих інтероперабельних метаданих українською та англійською мовами), створення інтелектуальної програмної системи, яка автоматизує ключові етапи обробки: розпізнавання текстів, їх структуризацію, нормалізацію, перетворення у машиночитані формати та інтеграцію з міжнародними платформами відкритої науки. Втілення такого технологічного рішення у закритому контурі на базі ВД «Академперіодика» гарантує безпеку та захищеність цифрових даних та прав інтелектуальної власності, а підготовка та поширення розроблених рішень гарантуватиме гнучкість та адаптацію продукту до потреб українського наукового простору. За результатом цифрового відтворення видань передбачене формування цілісної інфраструктури для їх включення у глобальний науковий простір. Оцифровані матеріали будуть доступні у відкритому доступі, придатні для індексації, пошуку, цитування та автоматизованого аналізу, відкривши результати українських досліджень, сприятиме їх включенню до міжнародних баз даних і розширить можливості для наукової комунікації.

КЛЮЧОВІ СЛОВА: ретроспективні архіви; оцифрування; Open Science; наукові видання; наукова комунікація; інтероперабельність; метадані.

ВСТУП. Значна частина наукового доробку НАН України, створеного впродовж останніх десятиліть, залишається фактично «невидимою» для сучасної науки через відсутність цифрового представлення та інтеграції у глобальні дослідницькі інфраструктури. Науковцями ВД «Академперіодика» НАН України пропонується системне рішення цієї проблеми шляхом створення інструментів і підходів, що перетворюють такі видання на повноцінний елемент цифрового наукового середовища на засадах відкритої науки, зокрема через автоматизоване збирання метаданих із подальшою експертною верифікацією.

Набутий досвід дозволяє спрямувати дослідження на вирішення проблеми обмеженої доступності та низької інтегрованості ретроспективної колекції наукових видань НАН України у міжнародний науково-інформаційний простір. Запропонований підхід передбачає створення інтелектуальної системи, що забезпечить перетворення архівних матеріалів у структуровані, семантично збагачені та машиночитані дані, придатні для використання в інфраструктурах відкритої науки.

Реалізація поставлених завдань забезпечить повернення до активного наукового обігу значного масиву знань, а також створить підґрунтя для впровадження нових підходів до роботи з науковими

текстами, зокрема їх автоматизованого аналізу та використання у системах забезпечення якості досліджень. У довгостроковій перспективі це сприятиме зміцненню позицій української науки у міжнародному просторі та формуванню сталих практик відкритої науки.

Також, запропоновані рішення сприятимуть підвищенню доступності, видимості та цитованості наукових публікацій, розвитку ефективної наукової комунікації та інтеграції української науки у світовий дослідницький простір.

Наукова новизна роботи полягає у формуванні концептуальної архітектури та методологічного фреймворку замкненого технологічного контуру ретроконверсії. Вперше запропоновано системну інтеграцію інструментів структуризації (GROBID), квантованих локальних великих мовних моделей (LLM) та контуру експертної валідації (Human-in-the-loop) для забезпечення інтероперабельності українськомовного академічного корпусу у стандарті JATS XML.

У результаті опрацювання широкому загалу стануть доступні малотиражні монографічні видання НАН України, які побачили світ після 1991 р. і не мають цифрових відповідників і метаданих українською та англійською мовами. Також завдяки оцифруванню й створенню файлів у машиночитаних форматах буде забезпечено збереження таких видань. Одним із побічних результатів реалізації стане суттєве розширення корпусу текстів, доступних для перевірки на наявність некоректних текстових запозичень (плагіату й самоплагіату), що сприятиме підвищенню рівня академічної доброчесності. Напрацьовані методичні підходи та набутий фахівцями досвід, а також апаратні рішення, стануть у пригоді для подальшого масштабування набутих практик на інші види видань (збірники наукових праць, наукові журнали), інші наукові установи і ЗВО.

Мета дослідження полягає в розробці архітектури, концептуальної моделі та методологічних засад інтелектуальної системи семантичної ретроконверсії. Запропоноване рішення обґрунтовує конвеєр автоматизованої обробки, структуризації та перетворення наукових видань НАН України (з 1991 р.) у машиночитані формати для їх подальшої інтеграції в екосистему Open Science.

МЕТОДИ. Методологія досліджень базується на комплексному поєднанні методів комп'ютерного зору (Computer Vision), обробки природної мови (NLP) та системного адміністрування видавничих платформ. В основі лежить створення автономного технологічного контуру ретроконверсії, що складається з таких етапів:

1. Методика неруйнівного цифрового захоплення (Data Acquisition). Для перетворення паперових оригіналів у цифровий формат використовується метод планетарного сканування. На відміну від планшетного сканування, дана методика передбачає використання V-подібної колиски та алгоритмів апаратної корекції вигину сторінок (Flattening Curve). Це дозволяє зберегти цілісність крихких палітурок видань 1990-х років та забезпечити високу роздільну здатність (400 DPI) для мінімізації помилок на етапі OCR. Наукові книги 1990-х часто не мають електронного відповідника взагалі. До того ж у зв'язку з малими тиражами видань значна частина книг того періоду зберігається в архіві Науково-видавничої ради НАН України лише в одному примірнику. Це зумовлює необхідність застосування щадних методів оцифрування. Безперечно, видавництва НАН України здійснювали розсилку обов'язкових примірників на підставі чинного законодавства, але відповідна постанова КМУ з'явилася тільки у липні 1992 р. (замінена 2002). Контрольні примірники надходили переважно до десятка основних бібліотек країни, деякі з них перебувають на тимчасово окупованих територіях (АР Крим, Луганськ, Донецьк), а бібліотеки в інших містах нині є zagrożеними об'єктами, питання збереження і доступності для користувачів їх фондів залишається невирішеним. Тому наявні в Науково-видавничій раді НАН України і ВД «Академперіодика» НАН України примірники фактично залишаються єдиними збереженими та доступними для оцифрування.

2. Метод структурної деконструкції документа (Document Parsing). Центральним елементом методології є перехід від простого розпізнавання тексту до семантичного парсингу. З використанням

інструментарію GROBID (на базі каскадних моделей умовних випадкових полів і нейромереж) буде реалізовано метод автоматичного виявлення логічної структури документа. Це допоможе ідентифікувати структурні частини сканованих документів: заголовки, списки авторів, афіліації, анотації та бібліографічні посилання як окремі об'єкти даних.

3. Інтелектуальна обробка та семантичне збагачення. Для обробки неструктурованих даних буде застосовано методику локального розгортання великих мовних моделей (LLM), а саме метод контекстуальної корекції, коли ШІ-модель аналізує результати OCR для виправлення типових помилок розпізнавання української мови та метод адаптивного перекладу та анотування. Техніку системного промпт-інжинірингу буде використано для генерації професійних англійських анотацій, адаптованих під вузьку наукову термінологію конкретної галузі знань.

Алгоритм передбачає розгортання локальної відкритої LLM (Llama 3 або Mistral), оптимізованої через квантування для роботи у внутрішній інфраструктурі без використання хмарних сервісів, її донавчання методом LoRA на структурованих наукових текстах у форматі JATS XML та налаштування через системний промпт для семантичної валідації й структуризації даних. Завершальний етап включає конвеєр Human-in-the-loop, де результати OCR і первинної обробки (GROBID) формуються у JATS XML і за потреби коригуються експертом для забезпечення придатності до міжнародної індексації.

4. Методика стандартизації та валідації. Отримані під час сканування дані буде піддано автоматичній трансформації у формат JATS XML. Методологія передбачає використання підходу Human-in-the-loop: автоматично згенерований XML-пакет проходитиме етап верифікації фахівцем через вебінтерфейс валідації. Це гарантуватиме 100%-ву точність перед фінальною інтеграцією у видавничу платформу.

Стратегія управління та збереження даних передбачає багаторівневе резервування цифрового масиву з використанням RAID 6, контроль цілісності через контрольні суми та диференційоване зберігання (SSD/NAS/резервні копії), що у поєднанні зі стандартизованим форматом JATS XML забезпечує довгострокову збережуваність, відтворюваність і технологічну сумісність даних.

5. Протоколи інтеграції та відкритого доступу. Роботи завершуватиме процес автоматизованого передання даних на платформу Open Monograph Press (OMP) за допомогою протоколів API. Для забезпечення «видимості» архівів буде налаштовано протоколи OAI-PMH, що забезпечить глобальним агрегаторам (OpenAIRE, Google Scholar) можливість автоматично «збирати» метадані з побудованого локального вузла. Це також допоможе інтегрувати підготований масив інформації з іншими європейськими і світовими системами збереження і надання у відкритий доступ наукової інформації.

РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ. Сьогодні світова наукова спільнота перейшла до парадигми Відкритої науки (Open Science), де доступ до знань визначається не лише наявністю PDF-файлу в мережі, а його «машиночитаністю» та інтеграцією у глобальні семантичні мережі. Проте для української науки, зокрема для масиву фундаментальних монографій НАН України 1991—2020 рр., існує критична проблема «цифрової невидимості». Сучасні наукові інфраструктури базуються на використанні технологій штучного інтелекту, семантичних графів знань і стандартизованих моделей метаданих. Наприклад, найбільша у світі відкрита бібліотека Internet Archive, цифровий агрегатор понад 50 млн різних об'єктів, зокрема публікацій, або Europeana з відкритим API для інтеграції з іншими системами. Питаннями оцифрування архівів ще від 1990-х років активно займаються також Бібліотека Конгресу США, Gallica, найстаріший волонтерський проєкт Project Gutenberg та ін. Вони використовують різне обладнання і програмне забезпечення, різні формати збережуваних файлів і стандарти метаданих. Ретельне вивчення їхніх умов і досвіду, а також побудови відкритих ресурсів дає нам можливість обрати такі параметри, що уможливлять ефективну інтеграцію та краще збереження оцифрованих наукових монографій, у яких представлено результати роботи українських науковців. Водночас ретроспективні архіви національних наукових установ залишаються

недостатньо інтегрованими в ці глобальні екосистеми, що створює суттєвий технологічний та інфраструктурний розрив. Український науковий доробок за період 1991–2020 років характеризується значною цифровою фрагментацією та обмеженою видимістю. Переважна частина наукових публікацій зберігається у вигляді паперових архівів або простих PDF-сканів у внутрішніх репозитаріях, без структурованих метаданих і англomовного представлення, що суттєво ускладнює їх індексацію у глобальних наукометричних та пошукових системах. Додатковим обмеженням є недостатній рівень стандартизації та структуризації метаданих, а також відсутність єдиних підходів до бібліографічного опису наукових видань, створених у різні періоди та за різними правилами оформлення. Це ускладнює автоматизований пошук, ідентифікацію авторства та відтворення наукових зв'язків між публікаціями, а також є причиною недоступності цього корпусу текстів для систем перевірки текстових запозичень. Тобто технологічна проблема полягає у необхідності розробки та впровадження інтелектуального технологічного конвеєра, який би поєднав сучасне сканувальне обладнання, локальні обчислювальні потужності для ШІ-аналізу та міжнародні видавничі стандарти для подолання «цифрової невидимості» української наукової спадщини, виведення її з ізоляції, забезпечення від знищення й загроженості цифрових резервних копій, надання користувачам стабільного доступу до неї на засадах відкритої науки.

Наукова проблема полягає у відсутності цілісного підходу до перетворення ретроспективних наукових видань у структуровані цифрові ресурси, придатні для інтеграції в сучасні інфраструктури відкритої науки. Незважаючи на значні обсяги наукової спадщини, зокрема наукових видань НАН України, ці ресурси залишаються поза межами сучасних інфраструктур відкритої науки через відсутність структурованого цифрового представлення, семантичної обробки та стандартизованих метаданих. У результаті вони не інтегровані в міжнародні наукові екосистеми, є обмежено доступними для автоматизованого пошуку, аналізу та цитування і фактично не функціонують як повноцінний елемент глобального наукового обігу. Ситуація ускладнюється відсутністю єдиної технологічної та методичної моделі трансформації ретроспективних наукових видань у машиночитані ресурси, сумісні з сучасними вимогами інтероперабельності та відкритої науки. Це призводить до фрагментації цифрових колекцій, втрати зв'язності наукових даних і зниження їх наукометричної та дослідницької цінності. Таким чином, ключовим викликом є не лише цифровізація наукової спадщини, а відсутність повноцінного інфраструктурного рішення, яке забезпечує її структуризацію, семантичну інтеграцію та включення до глобальних наукових платформ. Цей розрив обмежує видимість української науки, знижує ефективність використання накопичених знань і стримує їх включення у сучасні процеси наукової комунікації та аналізу. Інші ініціативи, вже розпочаті в Україні, зокрема у Національній бібліотеці України імені Ярослава Мудрого, Національній бібліотеці України імені В.І. Вернадського та ДУ «Книжкова палата України імені Івана Федорова», спрямовані передусім на оцифрування фондів задля збереження старовинних і рідкісних видань, документальної спадщини. Вони мають певні інструментальні обмеження, а також інші алгоритми і мету. Наприклад, не передбачають збору й укладання інтероперабельних метаданих за допомогою семантичних систем. Запропоновані дослідження орієнтовані на створення цілісного цифрового середовища для наукових видань, що поєднує процеси оцифрування, структуризації та семантичної обробки з формуванням машиночитаного корпусу даних, придатного для інтеграції в глобальні інфраструктури відкритої науки та міжнародні системи індексування. Вченими ДНТБ України проведені дослідження, результати якого свідчать, що впровадження відкритого доступу до наукових монографій є ключовим стратегічним вектором інтеграції України до Європейського дослідницького простору та обов'язковою умовою виконання вимог програми «Горизонт Європа». Попри наявні виклики, переваги відкритого доступу, зокрема зростання цитованості результатів досліджень на 50–200% та розширення географічного охоплення аудиторії, значно переважають фінансові та юридичні труднощі його імплементації.

У результаті проведеного дослідження запропоновано концептуальну архітектуру та методологічну модель автономної системи семантичної ретроконверсії. Спроектowana архітектура включає п'ять

взаємопов'язаних модулів — цифрового захоплення, структурної деконструкції, семантичної корекції та анотування, верифікації та конверсії, дистрибуції та інтероперабельності. Запропонована методологічна модель обґрунтовує використання закритого локального контуру, що гарантує захист прав інтелектуальної власності та відмовостійкість збереження даних (RAID 6 / NAS).

Запропонована архітектурна модель розрахована на пропускну спроможність до 1500 наукових праць на рік з мінімальним залученням ручної праці. Буде створено централізований цифровий репозитарій, тобто модернізовану платформу на базі Open Monograph Press (OMP), що міститиме повний масив ретроспективних видань (наукових монографій) НАН України з 1991 року та цифрове сховище страхового фонду; сформований на базі NAS-систем відмовостійкий архів майстер-копій (TIFF/PDF) об'ємом понад 10 ТБ, захищений від втрати даних і деградації цифрових носіїв.

Очікувані інформаційні та наукометричні результати досліджень будуть представлені у вигляді структурованого масиву даних у форматі JATS XML. Передбачається створення понад 500+ машиночитаних пакетів наукових даних відповідно до міжнародного стандарту NISO Z39.96. Це забезпечить автоматичну індексацію в Google Scholar, Scopus, Web of Science, інших агрегаторах. Надання та активація унікальних цифрових ідентифікаторів DOI кожному виданню з масиву ретроспективних наукових видань (зокрема монографій) забезпечить сталість наукових посилань, підвищення відтворюваності результатів досліджень і коректний облік цитувань українських науковців у міжнародних наукометричних системах, а також зв'язки цих видань з авторами й науковим установами, де підготовлено видання. Завдяки застосуванню технологій штучного інтелекту для автоматизованого анотування та семантичної обробки текстів кожне видання буде забезпечене якісними англійськими метаданими. Це сприятиме підвищенню міжнародної видимості українських наукових результатів, їх інтероперабельності з глобальними науковими інформаційними системами та потенційному зростанню рівня міжнародного цитування щонайменше на 20–30 %.

У межах виконання досліджень, відповідно до очікуваних методичних і суспільних результатів буде розроблено та оприлюднено методичні рекомендації (протокол), з покроковою інструкцією із застосування локальних систем ІІІ для оцифрування, семантичної обробки та структуризації академічних архівів. Документ буде спрямований на уніфікацію підходів до цифрової трансформації наукової спадщини та може бути масштабований на інші види видань (збірники наукових праць, наукові журнали), для впровадження в установах НАН України, інших наукових організаціях і ЗВО.

Реалізація запропонованих рішень забезпечить розширення відкритого доступу до ретроспективних наукових ресурсів шляхом створення цифрової платформи відкритого доступу до структурованих наукових даних. Це сприятиме підвищенню доступності фундаментальних знань для широкого кола користувачів — дослідників, викладачів та здобувачів освіти — та підтримуватиме розвиток відкритої науки, підвищення якості наукових досліджень і освітнього процесу в Україні.

ВИСНОВКИ. Запропонований комплекс досліджень щодо інтеграції архівів наукових видань НАН України у простір Open Science як виклик нагальної потреби інтеграції української науки в європейський дослідницький простір (EOSC) забезпечить ряд похідних переваг:

- подолання інформаційної ізоляції: вирішується проблема «невидимості» через відсутність машиночитаних метаданих великого масиву фундаментальних знань для світових пошукових систем;
- відповідність вимогам Open Science, зокрема перехід на XML-розмітку та автоматизацію реєстрації DOI;
- часткове подолання ресурсних обмежень: в умовах обмеженого фінансування автоматизація рутинної праці бібліографів та редакторів є єдиним способом опрацювати такий обсяг видань у стислі терміни;

- збереження національного інтелектуального капіталу: в умовах війни паперові архівів і внутрішні репозитарії НАН України перебувають під постійною загрозою. Створення територіально розподілених цифрових копій на відмовостійких сховищах (NAS) є питанням національної безпеки та культурного виживання;
- підвищення міжнародного авторитету України: забезпечення «видимості» українських вчених у Scopus та Web of Science через коректну індексацію метаданих сприятиме зростанню наукометричних показників установ НАН України та держави в цілому.
- освітню та суспільну цінність: результати досліджень відкривають безперешкодний доступ до академічних знань для студентів, викладачів та науковців, що сприятиме розвитку наукоємної економіки та суспільства знань в Україні. Також реалізація запропонованих рішень істотно збільшить базу порівняння наукових текстів з метою виявлення некоректних запозичень (плагіату);
- економічний ефект: упровадження розробленої технології дасть установам НАН України можливість заощаджувати бюджетні кошти, через зменшення обсягів ручної праці, зменшення кількості помилок (вплив людського фактору), або закупівлю іноземних пропріетарних систем управління контентом.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

- British Library. (n.d.). Endangered Archives Programme. <https://www.bl.uk/projects/endangered-archives>
- Cherniavskiy, S. (2026, March 25). "The shelling was the final blow." How the Book Chamber of Ukraine is trying to save 15 million printed publications. Detector Media. <https://detector.media/infospace/article/248563/2026-03-25-obstril-stav-ostannim-udarom-yak-knyzhkova-palata-ukrainy-namagaietsya-vryatuvaty-15-milyoniv-drukovanykh-vydan/> [in Ukrainian].
- Chronicling America. (n.d.). Chronicling America. Library of Congress. <https://chroniclingamerica.loc.gov/>
- Dublin Core Metadata Initiative. (n.d.). Dublin Core Metadata Initiative. <https://www.dublincore.org/>
- Europeana. (n.d.). API documentation. <https://pro.europeana.eu/page/apis>
- Gallica. (n.d.). Gallica. Bibliothèque nationale de France. <https://gallica.bnf.fr/>
- International Image Interoperability Framework. (n.d.). IIIF. <https://iiif.io/>
- Internet Archive. (n.d.). Internet Archive / Open Library. <https://archive.org/about>
- Kovtaniuk, Y. S. (2026, March 4). Experience of the NLUU. V. I. Vernadsky National Library of Ukraine. https://www.nbu.gov.ua/sites/default/files/all_files/202603_artilces_field_dopmat_files/2026.03.04_ua_yurii_s._kovtaniuk.pdf [in Ukrainian].
- Kramarenko, O. (2026). Open access to scientific monographs: The current landscape. Open Science and Innovation, 3(1), 27–34. <https://doi.org/10.62405/osi.2026.01.03> [in Ukrainian].
- Library of Congress. (n.d.). National Digital Library. <https://www.loc.gov/about/general-information>
- National Academy of Sciences of Ukraine. (n.d.). Books of the National Academy of Sciences of Ukraine. <https://books-nasu.org.ua/>
- National Library of Ukraine. (n.d.). Mobile units for the protection and digitization of library collections. <https://nlu.org.ua/article.php?id=656>
- Open Archives Initiative. (n.d.). OAI-PMH protocol. <https://www.openarchives.org/pmh>
- PDF Association. (n.d.). PDF/A standard (ISO 19005). <https://www.pdfa.org/>
- Publishing House "Akademperiodyka," National Academy of Sciences of Ukraine. (n.d.). Policy of the Publishing House "Akademperiodyka" of the National Academy of Sciences of Ukraine. <https://akademperiodyka.org.ua/uk/category/politics-ph-akademperiodyka/>
- Project Gutenberg. (n.d.). About Project Gutenberg. <https://www.gutenberg.org/about>

Publishing House “Akademperiodyka,” National Academy of Sciences of Ukraine. (n.d.). Regulations on ethics. <https://akademperiodyka.org.ua/uk/politics-ph-akademperiodyka/regulations-on-ethics/>

Save Ukrainian Cultural Heritage Online. (n.d.). SUCHO. <https://www.sucho.org/>

Ukrainian Library Association. (n.d.). Ukrainian Library Association. <https://ula.org.ua/>

Zhuk Oleksandr,

Researcher,

Publishing House «Akademperiodyka» of the National Academy
of Science of Ukraine, Kyiv, Ukraine,

ORCID: 0000-0001-9631-3906, e-mail: o.m.zhuk@nas.gov.ua

Havrysh Andrii,

PhD, Associate Professor, Senior Researcher,

Publishing House «Akademperiodyka» of the National Academy
of Science of Ukraine, Kyiv, Ukraine,

ORCID: 0000-0001-6474-6803, e-mail: Gavrysh@nas.gov.ua

DEVELOPMENT OF ARCHITECTURE AND METHODOLOGICAL FOUNDATIONS FOR A SEMANTIC RETRO-CONVERSION SYSTEM OF THE NAS OF UKRAINE SCIENTIFIC PUBLICATIONS FOR THE OPEN SCIENCE ECOSYSTEM

ABSTRACT. *This research aims to preserve the array of scientific heritage of the National Academy of Sciences of Ukraine and to integrate it into the modern digital scientific environment. A comprehensive methodological and technological solution is proposed (a standardised methodology for the digitisation, description and semantic enrichment of scientific publications, including the creation of verified, interoperable metadata in Ukrainian and English), along with the development of an intelligent software system that automates key processing stages: text recognition, structuring, normalisation, conversion into machine-readable formats, and integration with international open science platforms. Implementing such a technological solution within a closed loop based at the PH “Akademperiodyka” guarantees the security and protection of digital data and intellectual property rights, whilst the development and dissemination of the solutions devised will ensure the product’s flexibility and adaptability to the needs of the Ukrainian scientific community. Following the digitisation of the publications, a comprehensive infrastructure will be established to integrate them into the global scientific community. The digitised materials will be available via open access, suitable for indexing, searching, citation and automated analysis; by making the results of Ukrainian research accessible, this will facilitate their inclusion in international databases and expand opportunities for scientific communication.*

KEYWORDS: *retrospective archives; digitisation; Open Science; scientific publications; scientific communication; interoperability; metadata.*