

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
«ПОЛТАВСЬКА ПОЛІТЕХНІКА ІМЕНІ ЮРІЯ КОНДРАТЮКА»
Факультет філології, психології та педагогіки
Кафедра українознавства, культури та документознавства
Кваліфікаційна робота

МІЖНАРОДНА СПІВПРАЦЯ ЩОДО ОЦИФРУВАННЯ
ДОКУМЕНТНОЇ СПАДЩИНИ ЯК ЕФЕКТИВНИЙ ІНСТРУМЕНТ
УДОСТУПНЕННЯ АРХІВНИХ РЕСУРСІВ (НА БАЗІ ДЕРЖАВНОГО
АРХІВУ ПОЛТАВСЬКОЇ ОБЛАСТІ)

Магістрантка 6 курсу групи 601-ГД
спеціальності 029 «Інформаційна, бібліотечна та архівна та справа»

_____ Радченко Дар'я Іванівна

Науковий керівник
к. філол. н., доцент

Чередник Людмила Анатоліївна

Завідувач кафедри
д.і.н., професор

Передерій Ірина Григоріївна

Полтава 2025

Деканові факультету філології,
психології та педагогіки
Національного університету
«Полтавська політехніка
імені Юрія Кондратюка
Анні АГЕЙЧЕВІЙ

студентки групи 601-ФД
спеціальності
029 «Інформаційна, бібліотечна
та архівна справа»
Радченко Дар'ї Іванівни

ЗАЯВА

Прошу затвердити тему кваліфікаційної роботи «Міжнародна співпраця щодо оцифрування документної спадщини як ефективний інструмент удоступнення архівних ресурсів (на базі Державного архіву Полтавської області)».

Науковим керівником прошу призначити Чередник Людмилу Анатоліївну, кандидата філологічних наук, доцента, доцента кафедри українознавства, культури та документознавства.

28.08.2024

1. Завідувач кафедри

2. Керівник

Ірина ПЕРЕДЕРІЙ
Людмила ЧЕРЕДНИК

Національний університет «Полтавська політехніка імені Юрія Кондратюка»

Факультет філології, психології та педагогіки
Кафедра українознавства, культури та документознавства
Спеціальність 029 «Інформаційна, бібліотечна та архівна справа»

ЗАТВЕРДЖУЮ

Завідувач кафедри українознавства,
культури та документознавства
_____ Ірина ПЕРЕДЕРІЙ
« ____ » _____ 2024 року

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ МАГІСТРАНТУ**

Радченко Дар'ї Іванівні

1.Тема роботи: Міжнародна співпраця щодо оцифрування документної спадщини як ефективний інструмент удоступнення архівних ресурсів (на базі Державного архіву Полтавської області).

Керівник роботи: кандидат філологічних наук, доцент, доцент кафедри українознавства, культури та документознавства Чередник Людмила Анатоліївна.

2. Термін подання роботи: 13.01.2025.

3.Мета та завдання кваліфікаційної роботи: проаналізувати роль міжнародної співпраці в галузі оцифрування документної спадщини; розглянути нормативно-правове регулювання такого партнерства; оцінити стан українського внеску в оцифрування архівних документів; окреслити пропозиції щодо оптимізації наявних цифрових технологій; визначити особливості процесу оцифрування документів, а також розглянути актуальні проблеми забезпечення довготривалого збереження цифрових архівних фондів; розглянути іноземний досвід оцифрування та їхній внесок у розвиток таких проєктів на території України; визначити ефективність співпраці Державного архіву Полтавської області з міжнародною організацією FamilySearch в контексті діджиталізації архівних фондів; проаналізувати технічні параметри, що впливають на якість оцифрованих матеріалів, а також особливості програмного забезпечення, що використовується для оцифрування документної спадщини; розробити практичні рекомендації щодо вдосконалення функціонування ПЗ DСam.

Дата видачі завдання 28.08.2024.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Термін виконання	Примітки
1	Теоретична частина	01.11.2024 – 20.11.2024	46%
2	Аналітична частина	21.11.2024 – 20.12.2024	54%
3	Виготовлення ілюстративного матеріалу та підготовка до захисту	22.12.2024 – 13.01.2025	100%
4	Захист роботи	20.01.2025	

Магістрантка _____

Дар'я РАДЧЕНКО

Керівник роботи _____

Людмила ЧЕРЕДНИК

АНОТАЦІЯ

Радченко Д.І. Міжнародна співпраця щодо оцифрування документної спадщини як ефективний інструмент удосконалення архівних ресурсів (на базі Державного архіву Полтавської області). Спеціальність 029 «Інформаційна, бібліотечна та архівна справа», спеціалізація «Документознавство та інформаційна діяльність». Національний університет «Полтавська політехніка імені Юрія Кондратюка», Полтава, 2024.

У кваліфікаційній роботі досліджено сутність процесу оцифрування та його практичної реалізації, розглянуто міжнародний досвід кооперації задля реалізації відповідних ініціатив, систематизовано міжнародну нормативно-правову та технічну базу щодо діджиталізації документних матеріалів, що регламентує як процеси, так і технічні вимоги досліджуваного питання, розкрито різноманітність спектру всесвітніх проєктів в архівній галузі, проаналізовано специфіку співпраці Державного архіву Полтавської області з некомерційною генеалогічною організацією FamilySearch.

Визначено основні технічні вимоги до спеціалізованого обладнання та отриманих цифрових копій, а також роздільної здатності та форматів збереження зображень. Відповідно до зазначених критеріїв автором запропоновано оптимальні параметри та формати задля найбільш деталізованого, якісного та довговічного результату оцифрування архівних оригіналів. На основі практичного досвіду користування спеціалізованим програмним забезпеченням DСam від корпорації FamilySearch виконано загальний огляд роботи продукту, виокремлено переваги та недоліки його апарату, а також запропоновано внесення якісних змін до інтерфейсу ПЗ, його ключових систем функціонування та усунення багів.

Ключові слова: архів, міжнародна співпраця, оцифрування, діджиталізація.

103 с., 27 рис., 3 табл., 82 джерела.

ABSTRACT

Radchenko D.I. International cooperation on the digitization of documentary heritage as an effective tool for making archival resources available (on the basis of the State Archive of the Poltava Oblast). Speciality 029 “Information, Library and Archives”, specialisation “Documentation and Information Activity”. National University “Yuri Kondratyuk Poltava Polytechnic”, Poltava, 2024.

The qualification work examined the essence of the digitization process and its practical implementation, considered the international experience of cooperation for the implementation of relevant initiatives, systematized the international legal and technical framework for the digitization of documentary that regulates both the processes and technical requirements of the researched issue, the diversity of the spectrum of worldwide projects in the archival industry is disclosed, the specifics of the cooperation between the State Archives of the Poltava Region and the non-commercial genealogical organization FamilySearch are analyzed.

The main technical requirements for specialized equipment and received digital copies, as well as resolution and image storage formats are defined. In accordance with the specified criteria, the author proposed optimal parameters and formats for the most detailed, high-quality and long-lasting result of digitization of archival originals.

Based on the practical experience of using the specialized DCam software from the FamilySearch Corporation, a general overview of the product's operation was performed, the advantages and disadvantages of its operation were highlighted, and qualitative changes to the software interface, its key operating systems and bug fixes were proposed.

Keywords: archive, international cooperation, digitization, digitalization.

103 p., 27 pic., 3 tab., 82 sources.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ.....	
ВСТУП	9
РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ	14
1.1 Сутність поняття та процесу оцифрування документів.....	14
1.2 Новітні способи забезпечення збереженості архівних документів.....	23
1.3 Міжнародне нормативно-правове регулювання процесу оцифрування архівної спадщини.....	30
РОЗДІЛ 2. ДІДЖИТАЛІЗАЦІЯ ДОКУМЕНТНОЇ СПАДЩИНИ: ВІД ГЛОБАЛЬНИХ ПРОЄКТІВ ДО РЕГІОНАЛЬНОЇ ПРАКТИКИ.....	41
2.1 Досвід іноземних проєктів з оцифрування архівів.....	41
2.2 Співпраця Державного архіву Полтавської області з некомерційною організацією FamilySearch щодо діджиталізації архівних фондів	48
2.3 Стан, проблеми та перспективи українського внеску в оцифрування документної спадщини	55
РОЗДІЛ 3. РЕКОМЕНДАЦІЇ ЩОДО ОПТИМІЗАЦІЇ ПРОЦЕСУ ОЦИФРУВАННЯ НА ОСНОВІ ПРАКТИЧНОГО ДОСВІДУ	63
3.1 Аналіз параметрів оцифрування: роздільна здатність, формати файлів, використання штучного інтелекту.....	63
3.2 Переваги та недоліки DСam як програмного забезпечення для оцифрування документів	71
3.3 Рекомендації щодо створення нової версії DСam: впровадження змін на основі практичного досвіду	79
ВИСНОВКИ.....	89
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ ТА ДЖЕРЕЛ	94

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

БД – база даних

ДАКО – Державний архів Київської області

ДАПО – Державний архів Полтавської області

ДАХО – Державний архів Хмельницької області

ІКТ – інформаційно-комунікаційні технології

МН – машинне навчання

НАФ – Національний архівний фонд

ПЗ – програмне забезпечення

Укрдержархів – Державна архівна служба України

ШІ – штучний інтелект

CNN – convolutional neural networks (згорткові нейронні мережі)

DPI – dots per inch (точок на дюйм)

OCR – optical character recognition (оптичне розпізнавання тексту)

RL – reinforcement learning (навчання з підкріпленням)

RNN – recurrent neural networks (рекурентні нейронні мережі)

ВСТУП

Інформаційна епоха змінює традиційні підходи до вивчення архівних матеріалів. Сьогодні дослідники мають змогу створювати альтернативні архіви у віртуальному просторі, що розширює можливості для аналізу та інтерпретації даних.

Оцифрування архівів – це важливий крок до збереження всесвітньої культурної спадщини та забезпечення доступу до неї майбутнім поколінням. Завдяки міжнародній співпраці в галузі оцифрування ми отримуємо безпечний, економічний і зручний доступ до архівних матеріалів, що значно спрощує пошук і використання необхідних даних. Створення електронного фонду дозволяє об'єднати всі архівні матеріали в одну доступну базу даних. Такий підхід відкриває нові можливості для досліджень, що сприяє розвитку науки, освіти та культури, а також зумовлює **актуальність кваліфікаційної роботи.**

Посилення загроз для історико-культурної спадщини нашої країни вимагає негайних дій щодо її збереження. Тож оцифрування архівних фондів – це один із найефективніших способів захисту надбань наших предків від забуття, знищення чи викривлення.

Саме завдяки таким перевагам діджиталізація в Україні набирає обертів. Незважаючи на численні виклики, темпи цифровізації зростають, а результати цієї роботи виводять нашу країну на перше сходинки відповідних світових рейтингів. Це стало можливим лише завдяки тісній співпраці з міжнародними партнерами, адже завдяки їм наші архівісти отримали цінні знання, досвід оцифрування та навіть необхідне обладнання.

Дослідження порушеної проблеми передбачає глибокий аналіз досвіду партнерства державних архівів з міжнародними організаціями, передусім в контексті спільних проєктів з оцифрування історичної спадщини. Це дозволить нам виявити нові можливості для задоволення

зростаючих інформаційних потреб суспільства, виявлення нових інструментів та підходів до вирішення актуальних проблем сьогодні, пов'язаних із збереженням та доступністю документів в цифрову епоху.

Питаннями оцифрування світової історико-культурної спадщини присвячено праці багатьох дослідників. Вивчення цієї теми обрали для себе такі вітчизняні науковці, як: М. Долик, С. Литвинська, О. Менська, Н. Радченко, Н. Ревко, Р. Романовський та ін.

Проблеми збереженості архівних документів під час війни порушували у своїх роботах Л. Грінберг, Г. Мусіна, Н. Курило та багато інших.

Практичними аспектами оптимізації процесу оцифрування цікавились О. Кириленко, Д. Сіденко, Ю. Цирульнев та ін.

Можливості застосування штучного інтелекту та машинного навчання для оцифрування документів вивчали С. Анде, Б. Гена, Н.Кавадкар, М. Комова, В. Курило, З. Лу, Ф. Мен, Д. Рані та ін.

Варто зауважити, що проблеми оптимізації програмного забезпечення з оцифрування документів та підвищення зручності його інтерфейсу, як і базовий огляд варіацій такого програмного продукту від різних розробників, натеper досліджено недостатньо, оскільки переважну більшість наукових розвідок присвячено питанням споживання або поширення вже оцифрованих матеріалів. Тоді як сам процес і його складники, що безпосередньо впливають на результативність, не є надто популярною темою серед науковців і дослідників. Також замало розробленим є і питання престижу міжнародної співпраці Держкомархіву України в галузі оцифрування і впливу такої взаємодії на кількісні та якісні показники успішності оцифрування архівних матеріалів.

Мета роботи – проаналізувати роль міжнародної співпраці в галузі оцифрування документної спадщини; розглянути нормативно-правове регулювання такого партнерства; оцінити стан українського внеску в

оцифрування архівних документів; окреслити пропозиції щодо оптимізації наявних цифрових технологій.

Завдання дослідження:

1. Визначити особливості процесу оцифрування документів, а також розглянути актуальні проблеми забезпечення довготривалого збереження цифрових архівних фондів.

2. Розглянути іноземний досвід оцифрування та їхній внесок у розвиток таких проєктів на території України.

3. Визначити ефективність співпраці Державного архіву Полтавської області з міжнародною організацією FamilySearch в контексті діджиталізації архівних фондів.

4. Проаналізувати технічні параметри, що впливають на якість оцифрованих матеріалів, а також особливості програмного забезпечення, що використовується для оцифрування документної спадщини.

5. Розробити практичні рекомендації щодо вдосконалення функціонування ПЗ DСam.

Об'єкт дослідження – доступність (відкритість) архівних документних ресурсів.

Предмет дослідження – міжнародна співпраця з оцифрування документної спадщини як елемент удосконалення архівних ресурсів.

Отримання наукових теоретичних та практичних результатів дослідження зумовлено використанням загальнонаукових та спеціальних **методів**, основними з-поміж яких є: методи системного аналізу, синтезу, статистичний, емпіричний, порівняння.

Зокрема, загальнонаукові методи застосовано для аналізу наукової та методичної літератури щодо проблеми дослідження, вивчення стану теоретичного висвітлення проблеми. Метод аналізу – для дослідження нормативно-правового регулювання процесу оцифрування архівної спадщини. Метод синтезу обрано для узагальнення проведеного дослідження щодо специфіки оцифрування архівних фондів. Статистичний

метод – для дослідження позитивної динаміки результативності роботи Державного архіву Полтавської області протягом обраного періоду (3 роки), для порівняння здобутків оцифрування кількох архівів України (ДАПО, ДАХО, ДАКО), а також рейтингу країн світу за темпами оцифрування документного ресурсу. Емпіричні методи застосовано для виявлення наявних проблем в роботі програмного забезпечення з оцифрування, а також зіставлення з бажаним функціонуванням. Метод порівняння застосовано під час зіставлення наявного та бажаного стану програмного продукту, а також впливу змін на зручність користування.

Матеріалом дослідження є технічна та нормативно-правова документація щодо питань оцифрування в Державному архіві Полтавської області, а також некомерційній організації FamilySearch.

Наукова новизна полягає у систематизації наукових даних щодо досліджуваного питання, аналізові актуального стану процесу оцифрування в Державному архіві Полтавської області та Україні загалом, описі та аналізі роботи програмного забезпечення оцифрування архівних документів DСam від FamilySearch.

Практичне значення одержаних результатів. Під час дослідження нами було запропоновано внесення низки змін до ПЗ DСam на основі практичного досвіду його користувачів, що визначило шляхи оптимізації роботи задля підвищення рівня якості кінцевого контенту та швидкості його створення.

Апробація результатів дослідження. Результати виконаних наукових розробок з теми «Міжнародне нормативно-правове регулювання процесу оцифрування архівної спадщини» було апробовано під час роботи ІХ Міжнародної науково-практичної конференції «Документно-інформаційні комунікації в умовах глобалізації: стан, проблеми та перспективи» (м. Полтава, 21 листопада 2024 р.).

1. Радченко Д.І. Міжнародне нормативно-правове регулювання процесу оцифрування архівної спадщини. *Документно-інформаційні*

комунікації в умовах глобалізації: стан, проблеми та перспективи : матеріали ІХ Міжнар. наук.-практ. конф., 21 листоп. 2024 р. / редкол. : І.Г. Передерій [та ін.]. Полтава, 2024. С. 225 – 231.

Логіка дослідження зумовила таку **структуру кваліфікаційної роботи**: вступ, три розділи, висновки, список використаних джерел із 82 найменувань. Загальний обсяг роботи – 103 сторінки.

РОЗДІЛ 1. ТЕОРЕТИЧНІ ОСНОВИ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ

1.1 Сутність поняття та процесу оцифрування документів

На всіх етапах розвитку суспільства завдання захисту інформації є пріоритетним, і його вирішення має забезпечити збереження знань, накопичених попередніми поколіннями, та надання нових знань прийдешнім. Обсяг інформації, що надається в електронному вигляді, збільшується не тільки завдяки наявності нової інформації, яка вже існує, а й за рахунок оцифрування інформаційних ресурсів, створених у минулому. Надання такої історичної інформації в цифровому форматі вирішило багато проблем, пов'язаних зі збереженням інформації, і створило нові можливості для доступу до неї та її оброблення.

Впровадження системи електронного документообігу в діяльність архівних установ вимагає створення оптимальних умов для зберігання електронних документів та інформаційних ресурсів на національному рівні. У зв'язку з цим виникла необхідність організації Центрального державного електронного архіву України, який було засновано 12 травня 2007 року як центральне сховище електронних документів Національного архівного фонду України. Центральний державний електронний архів України здійснює зберігання електронних документів та інформаційних ресурсів Національного архівного фонду України, гарантує їх збереженість і цілісність, створює умови для їх використання [28].

Насамперед, на нашу думку, слід визначитися з поняттям «електронний архів». Найбільш поширеним є таке: «Електронний архів – це структурована система зберігання електронних документів, що забезпечує надійне зберігання, конфіденційність, розмежування прав доступу, відстеження історії використання документів та швидкий і зручний доступ до них» [28].

Науковці та практики суголосні, що метою створення електронних архівів є:

- зменшення витрат на зберігання: значна частина довідкових запитів може бути виконана за допомогою електронних зображень документів, що значно зменшує ресурсні витрати;
- скорочення ризиків втрати або пошкодження архівних документів за рахунок зменшення потреби в оригіналах документів;
- спрощення процесу передачі документів з оперативного зберігання на архівне [28].

З огляду на це, можна окреслити основні функції, які виконує електронний архів, найголовнішими з-поміж них є:

- Приймання архівних справ на зберігання та формування. Це відбувається шляхом перенесення документів з інших інформаційних систем або шляхом оцифрування та введення архівних документів у систему.
- Управління документами, що зберігаються як в електронному, так і в паперовому вигляді.
- Зручний і швидкий пошук документів і справ за назвами та атрибутами в тексті.
- Облік публікації оригіналів документів і справ.
- Надання доступу до архівних матеріалів в електронному вигляді [28].

Також, на нашу думку, важливо з'ясувати сенс поняття «оцифрування», оскільки вже вживали його в ході викладу матеріалу. Отже, майже всі дослідники одноголосно стверджують, що «оцифрування, оцифровування, або діджиталізація (англ. digitization) – це перетворення інформації в цифрову форму. Більш технічним визначенням буде «цифрова трансмісія даних об'єкта, закодованих у дискретні цифрові сигнальні імпульси» [42]. Результатом операції є послідовність чисел, що описує серію дискретних точок, які представляють об'єкти, зображення, звуки,

документи і сигнали (зазвичай аналогові сигнали). Такий результат перетворень прийнято називати «цифрове представлення» або, точніше, «цифрове зображення», у випадку об'єктів, і «цифровий формат» – у випадку сигналів. У сучасній практиці оцифровані дані мають вигляд, наприклад, двійкових чисел, що полегшує усі види комп'ютерних операцій щодо такої інформації. Загалом, оцифрування означає перетворення аналогового джерела в цифровий формат. Хоча аналогові дані, зазвичай, надійніші, цифрові дані можна схарактеризувати «більш легким обміном та отриманням до них доступу, а також перетворенням в новий стабільний формат» [42].

Як уже зазначалося вище, оцифрування – це «перетворення аналогових даних (наприклад, зображень, відео чи тексту) у цифровий формат»[42]. Іншими словами, оцифрування – це перетворення не цифрових об'єктів на цифрові дані або продукти. На практиці це означає, що, до прикладу, паперовий документ сканується і зберігається вже як цифровий (наприклад, у форматі PDF). Потім інформація з таких комп'ютеризованих аналогів традиційних документних носіїв може бути використаною в різних ситуаціях «залежно від потреб споживачів цього продукту» [42].

Англiцизм «дiджиталiзацiя» набув широкого вжитку в українській мові, і сьогодні його можна часто почути в офіційних матеріалах преси та у виступах різних посадовців. Словник сучасної української мови та сленгу «Мислово» обрав цей термін як найбільш уживане слово 2019 року. Він визначається як «зміни, що відбуваються в усіх сферах суспільного життя завдяки використанню цифрових технологій» [53].

Оксфордський словник англійської мови (OED) стверджує, що перше використання термінів «digitization», тобто «оцифрування», а також «digitalization», «цифровізація», датується ще 1954 роком та «пов'язане безпосередньо з процесом комп'ютеризації» [68].

Діджиталізація означає перехід на цифрові технології або впровадження або вдосконалення процесів з використанням цифрових

технологій і оцифрованих даних. Тобто, діджиталізація вимагає попереднього оцифрування.

Відомо, що цифровізація підвищує продуктивність та ефективність процесів і знижує витрати, але не змінює самі процеси. Іншими словами, діджиталізація перетворює процес, у якому людина контролює дію або послідовність дій, на програмований процес й оптимізує внутрішні послідовні етапи (наприклад, автоматизація завдань, мінімізація паперу тощо). Цифрову трансформацію можна визначити як «організаційні зміни, які здійснюються переважно за допомогою цифрових технологій з метою адаптації до вимог соціального середовища та зацікавлених сторін» [62, с. 65].

Категорії організаційної трансформації, що визначаються цифровою конверсією, містять організаційну форму, організаційну структуру, людські ресурси, процеси, технології, що використовуються для збору та управління інформацією, спектр послуг та моделі співпраці з зовнішнім середовищем. Прикладом цифрової трансформації, до прикладу, є «перехід від локального управління фізичними процесами до віддаленого моніторингу та управління між процесами» [62, с. 67].

Важливим питанням сучасності є оцифрування документної спадщини України, тобто «переведення документів, фільмів, фотографій, відео- та аудіоматеріалів в електронний формат» [59, с. 35]. Цей процес вважається нам напрочуд актуальним на тлі викликів сьогодення. Наразі зберігається серйозна загроза втрати важливої документної інформації, що пов'язано зі збройною агресією, терористичними актами та війною. Саме тому важливість оцифрування документної спадщини України вже ні в кого не викликає сумнівів і є актуальною та невідкладною задачею для її власників, а результати такої роботи будуть корисними для громадян, держави та міжнародної спільноти.

Варто зауважити, що ідея створення цифрового архіву впливає з серйозних проблем, які виникають в роботі архівів, а саме:

– Збереження матеріалів. Існує можливість зробити копії матеріалів так, щоб оригінали використовувалися лише один раз. Надалі ці копії можна відтворювати, перевозити, дублювати, використовувати в друкарнях і типографіях, досліджувати і виставляти без знищення оригіналів.

– Зменшення кількості персоналу. Накопичення матеріалів вимагає збільшення кількості персоналу, залученого до архівування. Цифрові архіви можуть працювати з оцифрованими копіями замість оригіналів без необхідності залучення додаткового персоналу.

– Сталі системи архівування. У процесі зміни системи опису створюється конвертер, який переносить описи наявних в архіві матеріалів до нової системи.

– Доступність. Зазвичай, деякі архівні матеріали не завжди доступні. Для вирішення цієї проблеми можна використовувати копії.

– Геолокалізація. Доступність архіву через інтернет уможлиблює працювати з документами з віддалених місць.

– Час дослідження. Під час створення складних баз даних для оцифрування архівів необхідно впроваджувати методи обробки цієї інформації. Наукове дослідження, яке проводилося протягом кількох місяців у різних архівних установах, тепер можна здійснити за кілька хвилин за допомогою бази даних [1, с. 18; 34, с. 277].

З огляду на наукові дослідження, слід наголосити, що в архівному секторі оцифрування здійснюється за такими чотирма взаємопов'язаними напрямками:

- створення та розвиток моделей оцифрованих фондів;
- створення колекцій цифрових інформаційних ресурсів;
- розвиток цифрових послуг;
- підготовка фахівців з оцифрування підвищення їхньої кваліфікації [40, с. 4].

За словами академіка О.С. Онищенко, «проблема наукової та культурної національної спадщини полягає в її збереженні. Більшість

літератури останніх трьох століть була надрукована на кислотному папері, термін експлуатації якого добігає кінця» [40, с. 4].

Процеси діджиталізації детально вивчаються дослідниками і, на їхню думку, можна окреслити такий спектр можливостей щодо цього:

– Оцифрування науково-інженерної інформації (графіки, карти та інше). Для цього існує низка програмних застосунків, таких, наприклад, як «Plot digitizer» та «Engauge Digitizer», що належать до відкритого програмного забезпечення Open Source.

– Оцифрування звуку. Перетворення аналогового сигналу в цифровий відбувається з допомогою спеціального пристрою «аналогово-цифрового перетворювача» (наприклад, звукова карта комп'ютера). Крім того, цифровий сигнал часто генерують за допомогою цифрових синтезаторів. Для програвання цифрового сигналу він знову перетворюється у аналоговий, і передається на гучномовець.

– Оцифрування відеоматеріалу, телесигналу або кінофільму з таких фізичних носіїв, як відеокасета, відеодиск тощо.

– Книжкове оцифрування – переведення паперових видань в електронний вигляд [42].

Безумовно, спектр можливостей можна продовжувати й далі.

Відомо, що ключовим компонентом процесу оцифрування є діджитайзер – пристрій, що перетворює відеосигнали з аналогової форми в цифрову. Передусім він використовується для введення графічної інформації в комп'ютери та перетворення даних зі старих носіїв (наприклад, магнітних стрічок, касет VHS) у цифровий формат. Такі спеціальні пристрої надають найбільш якісний результат наприкінці роботи з оцифрування, наприклад, для «дрібних елементів карти точність розпізнавання може досягати декількох тисяч міліметрів» [42].

Наразі існує два найбільш вагомні етапи процесу оцифрування:

1. Обов'язковий, що означає отримання копії сторінки у вигляді графічного (зазвичай растрового) зображення.

2. OCR (optical character recognition) – оптичне розпізнавання символів. Насамперед, технологія полягає в тому, що OCR з наступним збереженням розпізнаного тексту в одному з форматів електронних книг надає подальшу можливість повнотекстового пошуку книг та індексування великих серій електронних копій. «Однак при цьому складно відтворити оригінальну верстку, малюнки, діаграми та формули. Крім того, виникають помилки при розпізнаванні інформації» [32, с. 134; 34, с. 277].

Для подальшого дослідження нас необхідно з'ясувати всі тонкощі цього процесу. Як уже зазначалося, сканування документів – це «процес створення електронної копії паперового документа» [42]. Тут ми говоримо саме про електронні копії, а не створення фізичних копій. Щоб відсканувати документ, вам потрібно скористатись відповідним технічним обладнанням, наприклад, сканером – пристроєм, призначеним для обробки світлових променів, що відбиваються або проходять через поверхню об'єкта, для отримання його зображення. У вузькому сенсі так називається пристрій, що «використовується для отримання цифрового зображення комп'ютера, такі як PDF або JPG» [64].

Відсканований документ можна зберігати на комп'ютері або в хмарному сховищі. Якщо документ складається з кількох сторінок, їх можна об'єднати в один файл.

Слід нагадати, що сканування документа зберігає дані в електронному форматі, але текст у документі не можна редагувати. І за потреби внести зміни в документ це можна зробити тільки вручну. Безумовно, це доволі трудомісткий процес, який потребує великих ресурсів. Крім того, сканування паперового документа не дозволяє читати, аналізувати чи редагувати його в електронному вигляді як текст або шукати уривки в повному тексті документа. «Результатом сканування є електронний файл, що містить графічну копію паперового документа»[61].

Важливо також пам'ятати, що для здійснення вже повнотекстового пошуку документа його текст повинен бути перетворений зі сканованого растрового зображення в символи ANSI/ASCII, які можуть бути оброблені комп'ютером як текст (те саме OCR), а електронна копія документа повинна бути пов'язана з метаданими Результатом OCR і стане електронний файл, що містить текст документа в символах ANSI/ASCII. Електронна копія такого документа повинна бути проіндексована для забезпечення швидкого пошуку її за допомогою ключового (індексного) поля. Результатом індексування є електронний (індексний) файл, що містить інформацію про відповідність між електронною копією та індексним полем. Для забезпечення можливості пошуку інформації про документ електронна копія має бути пов'язана з метаданими (даними про документ). Така послідовність процесів оцифрування документа зображена на рисунку 1.1. Результатом індексування є електронний файл, що містить «інформацію щодо відповідності між електронною копією та метаданими документа» [60, с. 241].

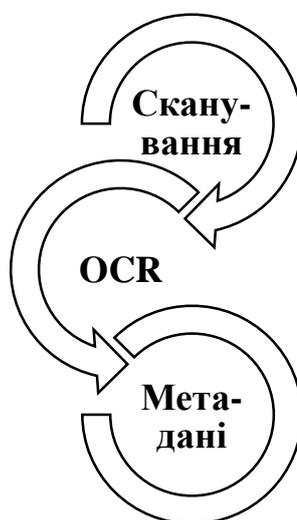


Рис. 1.1 – Взаємозв'язок етапів оцифрування документа
Розроблено автором

Останніми роками (особливо з появою формату DjVu) все частіше використовуються гібридні підходи, у яких поєднуються переваги обох

підходів: автоматичне розпізнання тексту книги та накладання його на растрове зображення оригінальної сторінки.

Значного впливу на оцифрування також надає нерівність у співвідношенні наявних ресурсів архівних установ та загального обсягу документів, які бажано піддати необхідним перетворенням для сучасної ери технологій. Про цей дисбаланс також зазначає дослідник Ю.Б. Цирульнев та називає його «залповим оцифруванням» [58, с. 121]. Він також зауважив, що «оцифрування великих обсягів документів, які утворювалися в попередні періоди («Залпове оцифрування») має сенс здійснювати ресурсами компаній, які мають спеціалізовані призначені для оцифрування великих обсягів документів технології, обладнання, програмне забезпечення та персонал» [58, с. 124].

З досвіду нашої роботи, варто підкреслити, що оцифрування документів, створених у процесі повсякденної роботи, доцільно здійснювати власними силами організації, використовуючи автоматизовані робочі місця та ділянки оцифрування документів, продуковані компаніями з відповідними компетенціями. Безумовно, пакетне оцифрування вимагає спеціального обладнання та програмного забезпечення (ПЗ), яке надає низку ключових функцій для забезпечення якості та швидкості сканування. До прикладу, для безперервного оцифрування «можна використовувати стандартне обладнання та програмне забезпечення, що постачається разом зі сканером» [58, с. 122-123].

Отже, натеper актуальною є робота щодо створення електронних копій архівних справ. Найголовнішою перевагою та першопричиною оцифрування є передача документної інформації та забезпечення кращої збереженості оригінала. Саме тому відповідно до кожного виду документа фахівці знайшли свої способи оцифрування. Так, найбільш розповсюджені текстові документи проходять два етапи обробки: створення растрових зображень та подальше оптичне розпізнавання тексту.

1.2 Новітні способи забезпечення збереженості архівних документів

Натепер оцифрування архівних документів є одним з найважливіших напрямків діяльності архівних установ у всьому світі, оскільки це стосується збереження історико-культурної спадщини будь-якої країни і України, зокрема.

Як уже зазначалося в першому розділі, джерелом інформації для українських архівів є Національний архівний фонд (НАФ). На наше глибоке переконання, архів має бути сакральним місцем збереження національної пам'яті та її увіковічення.

Крім того, інформатизація архівної справи має на меті, передусім, вдосконалення управління та розвитку архівних установ, забезпечення міжархівного співробітництва та налагодження зв'язку з користувачами архівної інформації за допомогою технічних засобів комунікації. Така діяльність може також бути продуктивною і в новій царині двостороннього діалогу з наявними вебсайтами архівознавчого та історичного характеру.

Відомо, що процес інформатизації України розпочали понад 30 років. Важливим етапом стала Державна програма «Національна архівна інформаційна система (НАІС) «Архівна та рукописна Україніка». Початок її розробки датується 1992 роком, а головним завданням стало створення єдиної інформаційної системи, яка б об'єднала всі архіви, що зберігають документи українського походження. «Архівна україніка» – це зібрання архівних документів, пов'язаних з Україною, що зберігаються за кордоном. Потенціал її використання є досить значним, особливо зважаючи на те, що її включення в сучасне соціальне та комунікаційне середовище надасть нових контекстів загальновідомим історичним подіям. Саме тому така робота є «загальнонаціональним завданням» [26, с. 24-25].

Збройні конфлікти у світі завжди впливають на всі сфери життя суспільства. Українські інформаційні сховища, як соціальні та комунікаційні інституції, усвідомили виклики війни. З метою подолання

проблем і загроз були визначені пріоритетні напрями діяльності цих установ та шляхи адаптації до реалій воєнного часу. У ці складні для України часи в діяльності архівів передбачено комплекс заходів, спрямованих на покращення інформаційного обміну на національному та регіональному рівнях, забезпечення інформаційно-комунікаційної підтримки суспільно значущих процесів і подій. Нині важливим завданням національних та регіональних бібліотек і архівів є збір інформації щодо «пошкодження та знищення культурної спадщини загалом та документної спадщини різних регіонів та громад України зокрема» [19, с. 79].

Загалом, збереження культурної та національної спадщини набуває сьогодні особливого значення в умовах російсько-української війни. Перед органами державної влади та керівниками архівів постала нагальна проблема розроблення правил і рекомендацій щодо збереження архівних фондів.

Трагічний досвід окупованих територій та неможливість відновлення архівної інформації вимагає чіткого визначення послідовних завдань збереження архівних документів. Міжнародні документи та норми міжнародного права передбачають, що документна спадщина є недоторканною, але наразі ми є свідками постійних порушень цих норм із боку окупантів. На національному рівні було започатковано низку програм з оцифрування та збереження архівних документів. Унаслідок, всі установи: чи то бібліотеки, музеї чи архіви – стали відповідальними за оцифрування рідкісних видань та стародруків. Проект створено за ініціативи Міністерства культури та інформаційної політики України. Також у майбутньому планується створення єдиного електронного ресурсу оцифрованих документів.

У період війни з'явилася ще одна проблема: вивезення ворогами документів із завойованих територій. Сумним прикладом цього може слугувати Херсонський державний архів, який втратив 40% документів

НАФ. «Ці документи були вивезені окупаційною владою до Криму» [38, с. 147].

Національний архівний фонд України постійно працює над створенням єдиного інформаційного ресурсу для доступу до оцифрованих архівних матеріалів. Наразі вебсайт надає доступ до дев'яти архівів через єдине вікно. Кінцевою метою, як зазначалося в документах, є «підключення всіх національних архівів до єдиного вікна» [38, с. 148].

Наразі Інтернет-сайти зробили архіви більш доступними та відкритими для громадськості. З огляду на дослідження науковців, сучасна американська та європейська історіографія вже давно переосмислює завдання, які стоять перед сучасними архівістами. Учені дійшли висновку, що «зміна архівної парадигми спонукає архівістів не лише бачити себе хранителями документної спадщини, але й усвідомлювати особливе значення своєї участі в історичних процесах, утверджувати та поглиблювати свою активну роль у формуванні суспільної пам'яті людства» [49, с. 151]. З цією метою було «створено різні онлайн-платформи для подальшого поглиблення співпраці між архівними установами та місцевими громадами» [49, с. 152].

Безсумнівно, що документи Національного архіву посідають важливе місце як джерело історичних фактів і пам'яті українського народу. Порівняння історичних фактів минулого та сьогодення дозволяє краще проаналізувати причини поразки та перемоги нашої країни у боротьбі за незалежність. Згідно з чинною «Програмою оцифрування архівних інформаційних ресурсів на 2022–2025 роки», першою оцифровує та публікує документи система державних архівів. Фактично – це програма-мінімум на ці чотири роки, «із щоквартальним моніторингом виконання» [5].

«Попри відкриту збройну агресію росії на території нашої країни, українські архівісти продовжують працювати та оцифрувати документи Національного архівного фонду задля їх збереження та відкритості архівної

інформації для користувачів», – зазначає Державна архівна служба України на своєму офіційному вебпорталі [22].

Війна – це великий ризик не лише для життя людей, але й для збереження культурної спадщини, зокрема й архівів. Це усвідомлює й Надзвичайний фонд спадщини ЮНЕСКО, який схвалив проєкт вартістю 1,5 мільйона доларів, щоб допомогти Міністерству культури та інформаційної політики України «розробити інструменти оцифрування у сфері культурної спадщини в Україні, створити необхідну серверну інфраструктуру, апаратне та програмне забезпечення, а також залучити міжнародного консультанта, який супроводжуватиме всі процеси і допоможе розробити стратегію цифрової інтеграції» [63].

Безумовно, що найважливішим першим кроком у створенні цифрової колекції є визначення потреб і пріоритетів. Варто підкреслити, що під час оцифрування цілого архіву вирішальну роль відіграє низка чинників, а саме:

- нагальна потреба обмежити доступ до оригіналів документів через підвищену загрозу їхній фізичній цілісності;

- великий обсяг отриманих матеріалів і потреба збільшення штату задля їхнього опрацювання;

- збережені документи досить важко вписати в існуючу систему ідентифікаторів, що ускладнює також постійне внесення змін до наявної процедури присвоєння архівних шифрів;

- каталоги створюються для різних архівів різних країн, і системи ідентифікаторів у них також мають свої відмінності;

- велика кількість дослідників не може працювати з оригіналами одночасно, що також підтверджує правильність процесу оцифрування як такого;

- географічним розташуванням фізичних оригіналів архівних документів також створюють значні обмеження [34, с. 277].

Ще одним поштовхом до оцифрування архівів є розуміння можливого знищення документів через вплив часу. Деякі фізичні носії інформації хоч і

зберігаються людством не одну сотню років, але і назвати їх вічними ми не можемо (див. рис. 1.2). Найбільшу частку архівних справ становлять документи, виготовлені з паперу. Їхній стан залежить не тільки від епохи створення, але й від місця зберігання та катаклізмів, яких вони могли застати (потопи, пожежі, напади короїдів).

Інший фізичний носій, такий як магнітна стрічка, в основному використовується для архівного зберігання інформації в цифровому форматі. Існує значний досвід зберігання аудіо- та відеоінформації на магнітних носіях, але зразки магнітних стрічок, на які було записано аналоговий звук, зберігають інформацію вже понад 50 років [34, с. 276]. В останні десятиліття системи запису на магнітну стрічку були значно вдосконалені.

Відомо, що терміни зберігання галогенідно-срібної мікроплівки становить кілька десятиліть. Водночас ця технологія не дозволяє записувати мультимедійну інформацію чи відтворювати дані на високій швидкості. Також ведуться роботи з оцифрування мікрофільмів.

Із розроблених на сьогоднішній день технологій дослідники виділяють оптичні методи запису і збереження інформації як найкращі відповідники існуючим вимогам створення носіїв інформації тривалого зберігання. Оптичні диски на скляних підкладках з одношаровим WORM-покриттям мають найдовший термін зберігання з-поміж усіх оптичних дисків [34, с. 278-279]. За словами самих розробників, термін зберігання записаної інформації становить 20-30 років для типового компакт-диска, але може досягати навіть 100 років. Оцифрований матеріал використовується для наповнення баз даних (БД), створених за допомогою різного програмного забезпечення. Оцифровані матеріали також можуть бути розміщені на відповідних порталах в мережі Інтернет.



Рис. 1.2 – Залежність часу зберігання інформації від фізичного носія документа
Розроблено автором

Важко також переоцінити важливість створення історичних медіаархівів. Фотографії, фільми та звукозаписи, які з часом зіпсувалися, принаймні витримують випробування часом, якщо не назавжди, після оцифрування. А головне – вони будуть доступні широкому загалу. Це зручно як для дослідників, так і для тих, кого зацікавили старі фотографії та фільми.

Є ще один важливий момент: оцифровані зображення і звук стають інструментами мислення для художників, журналістів і кінематографістів, які візуалізують історію і тим самим беруть участь у ширшій суспільній дискусії. Український режисер Сергій Буковський, наприклад, зробив це у своєму фільмі «Назви своє ім'я», заснованому на спогадах тих, хто пережив Голокост в Україні. В основу фільму ліг відеоархів, створений за ініціативи Стівена Спілберга, в якому були записані інтерв'ю зі свідками трагедії. Можна було б згадати багато подібних проєктів, особливо мистецьких, які пов'язані з травматичною пам'яттю. Таке мистецьке осмислення історії є

надзвичайно важливим, оскільки «не тільки змушує людей задуматись, але й висвітлює цілі історичні періоди та змінює світогляд людини»[50].

Такими ж принципами керувались і під час ухвалення Закону України «Про доступ до архівів репресивних органів комуністичного тоталітарного режиму 1917-1991 років» у 2015 році [3]. Насправді, Україна є чи не єдиною країною, яка прийняла рішення щодо повного оприлюднення цього періоду, включно з архівами спецслужб. Таке ставлення до цього питання також пов'язано зі страхом світової спільноти щодо оприлюднення списків агентів, які співпрацювали з радянськими спецслужбами.

Але практика свідчить, що такі документи втратили свою цінність ще в останні роки радянської влади за негласного наказу до знищення, який був проведений як «зміна терміну зберігання для певного переліку документів» [17]. В Україні ж Архіви СБУ, яким перейшли у спадок архіви КДБ, «зберігають найбільшу кількість документів радянських спецслужб за 1930-1940-і роки, які іноді зберігалися краще, ніж за період 1970-1980-х років» [17].

На нашу думку, саме тоді почалась незворотна тенденція відкритості, модернізації, доступності та оцифрування українського НАФ.

Крім того, слід зазначити, що однією з вагомих цілей оцифрування є усунення рутинної корупції в архівах. Користувачі вже давно намагаються спокусити архівістів внести зміни в архівні документи. До прикладу, так ще 10-15 років тому найпоширенішими випадками були «кримінальні справи, зокрема, фальсифікація національності у свідоцтві про народження або реєстрація як жертви нацизму для отримання компенсації від держави» [17]. Діджиталізація практично ліквідувала такі випадки. Це тому, що документи доступні онлайн і можуть бути перевірені за 2 – 3 секунди.

Отже, архіви – це не лише основа для спадщини та історичних досліджень. Інформація, що зберігається в архівах, використовується для того, щоб робити висновки про той чи інший період історії. Наразі ми можемо говорити про те, що весь цей масив даних може перейти в

збереженому вигляді до майбутніх поколінь завдяки нинішньому процесу оцифрування. Саме таких висновків ми дійшли, спираючись на усі можливі загрози існування архівних колекцій.

І ще один важливий, на нашу думку, аспект: у всьому світі архіви ніколи не були належно оцінені, але з кожним роком цікавість громадян до НАФ усе зростає. Тож необхідно крокувати в напрямку задоволення існуючих запитів користувачів та подовження циклу життя архівних документів.

1.3 Міжнародне нормативно-правове регулювання процесу оцифрування архівної спадщини

Оцифрування архівних документів, як і будь-яка діяльність архівних установ, потребує свого унормування та правового підґрунтя, яке відіграє ключову роль і забезпечує організований та системний підхід до процесу. У багатьох країнах зарубіжжя вже розпочато розроблення нормативно-правових актів, що регулюють саме процес оцифрування. Вивчення такого досвіду свідчить про важливість цифрових технологій для збереження світової спадщини.

Насамперед, нормативно-правові акти визначають основні принципи та стандарти оцифрування, що сприяє уніфікації процесів та забезпеченню високої якості результатів. Завдяки встановленню чітких вимог до форматів файлів, параметрів сканування та метаданих гарантується довговічність та сумісність цифрових копій у майбутньому.

Крім того, законодавчі чи нормативно-правові акти у сфері оцифрування є важливими й у таких аспектах, як:

– захист авторських прав: оцифрування творів вимагає дотримання певних правил, щоб гарантувати, що права авторів не порушуються;

– захист культурної спадщини: оцифрування архівів, бібліотек і музеїв може зберегти культурну спадщину для майбутніх поколінь, але воно також порушує питання доступу та використання такої інформації;

– цифрова безпека: оцифровані дані повинні бути надійно захищені від несанкціонованого доступу, втрати або пошкодження;

– електронне урядування: оцифрування адміністративних процедур сприяє підвищенню ефективності роботи державних установ, але при цьому необхідно забезпечити конфіденційність персональних даних.

Також варто підкреслити, що нормативно-правове регулювання сприяє залученню державного фінансування та грантової підтримки для реалізації проєктів з оцифрування. Документи, що регулюють цей процес, можуть містити положення щодо співпраці з міжнародними організаціями та партнерськими установами, що сприяє обміну досвідом і використанню передових технологій.

Неабияке значення має і той факт, що оцифрування архівів також сприяє розвитку електронного урядування та уможливорює оперативно отримувати необхідну інформацію для прийняття управлінських рішень. Завдяки нормативному врегулюванню створюються передумови для формування єдиних електронних архівів, що сприяє підвищенню ефективності діяльності державних органів.

Перед аналізом міжнародних вимог до оцифрування ми маємо звернути увагу на базові законодавчі акти України, що регулюють процес діджиталізації, з-поміж яких вагомими є такі, як:

– Цивільний кодекс України: містить статті, пов'язані з авторським правом, суміжними правами та захистом персональних даних [16].

– Закон України «Про авторське право і суміжні права»: регулює «права та обов'язки авторів, виконавців, виробників об'єктів авторського права, а також встановлює порядок використання об'єктів авторського права» [2].

– Закон України «Про захист персональних даних»: стосується обробки персональних даних, «встановлює права суб'єктів персональних даних та обов'язки операторів» [6].

– Інше законодавство: залежно від сфери оцифрування може застосовуватися інше законодавство, наприклад, Закон України «Про Національний архівний фонд та архівні установи» [8], Закон України «Про електронну ідентифікацію та електронні довірчі послуги» [4].

З-поміж нормативно-методичних документів, що регулюють використання технологій оцифрування, можна виділити низку міжнародних нормативних документів, зокрема, такі, як: стандарти та технічні звіти Міжнародної організації зі стандартизації; національні нормативні документи, передусім, різноманітні стандарти, які має кожна країна, методичні рекомендації та інструкції. Зазначимо, що існує чимало різних нормативних та методичних документів, що регулюють процес оцифрування. Назвемо деякі з них (Рис. 1.3).

«Стандарт оцифрування S6»	Визначено повноваження Головних архівістів Сполученого королівства та союзних держав (Нова Зеландія, Австралія і Канада) щодо використання стандартів для управління архівними документами
Технічний звіт ISO/TR 13028-2010	Містить рекомендації щодо створення та підтримки документів у виключно електронному вигляді
AS/NZS ISO 13028: 2012	Міжнародний стандарт для створення цифрових зображень паперових джерел, що мають якість, яка забезпечує їхню функціональну еквівалентність паперовому джерелу
ISO/TS 19264-1:2017	Передовий досвід фіксації електронних графічних образів матеріалів культурної спадщини

Рис. 1.3 – Міжнародні нормативно-методичні документи, що регламентують технічний аспект оцифрування
Розроблено автором

Безумовно, що дотримання правильних технічних стандартів (табл. 1.1) щодо оцифрування архівних матеріалів є фундаментальним кроком для довгострокового збереження світової культурної спадщини.

Таблиця 1.1 – Огляд на деякі стандарти оцифрування

Розроблено автором

Назва	Опис	Технічні вимоги	Переваги
ISO 19264	Міжнародний стандарт, розроблений для забезпечення високої якості зображень, отриманих в результаті оцифрування. Він встановлює критерії оцінювання якості цифрових копій, а це, насамперед, дозволяє гарантувати, що вони точно відображають оригінали.	Аналіз якості зображень: Стандарт визначає методики оцінювання різних характеристик зображень: тони та шум; кольори; деталі; геометрія. Тестові діаграми: вимоги до тестових діаграм, які використовуються для оцінювання якості сканування. Загальні рекомендації щодо організації процесу оцифрування, від підготовки матеріалів до зберігання цифрових копій.	Об'єктивне оцінювання якості. Покращення якості цифрових копій. Співпраця між архівами: стандарти полегшують обмін даними між різними архівами та установами.
METS	Міжнародний стандарт, який використовується для створення метаданих, що містять інформацію щодо формату файлу, його роздільної здатності, колірного профілю, авторського права, дати створення та інших важливих характеристик.	Структурування метаданих: METS забезпечує чітку структуру для організації метаданих, що полегшує їхнє розуміння та використання. Інтерпретація: визначає правила інтерпретації метаданих, що забезпечує їхню однозначність. Обмін метаданими.	Співпраця між організаціями. Довготривале збереження. Покращення пошуку та доступу. Підтримка різних форматів.
OAI-PMH	Протокол, що дозволяє автоматично збирати метадані з різних цифрових сховищ. Він забезпечує стандартизований спосіб отримання метаданих, що полегшує пошук і доступ до оцифрованих матеріалів.	Сервер OAI-PMH: Система, яка надає доступ до метаданих відповідно до протоколу OAI-PMH. Клієнт OAI-PMH: Система, яка збирає метадані з сервера OAI-PMH. Реєстр OAI-PMH: Каталог серверів OAI-PMH, який дозволяє знайти потрібні сервери.	Покращення пошуку та доступу. Збільшення видимості. Співпраця між організаціями. Зменшення дублювання зусиль.

Українські та зарубіжні науковці окреслили головні чинники, які необхідно враховувати в процесі оцифрування:

- сумісність цифрових копій з майбутніми технологіями та мінімізація ризиків втрати даних через застарілість формату;
- дотримання стандартів гарантує, що оцифровані матеріали можуть бути інтегровані в різні інформаційні системи та доступні широкому колу користувачів;
- параметри сканування, стиснення та метаданих забезпечують максимально точне відображення оригіналу;
- спільні стандарти полегшують обмін цифровими архівами між різними країнами та організаціями;
- перед початком оцифрування бажано проконсультуватися з експертами в галузі архівної справи та інформаційних технологій;
- перед оцифруванням слід провести оцінку стану матеріалів [62, с. 73].

Отже, з рисунку 1.3 та таблиці 1.1 ми можемо зрозуміти, що світова спільнота приділяє достатню увагу питаннями розробленості нормативно-методичної та технічної документації щодо процесу оцифрування. Така залученість свідчить про важливість проблеми та її вагомість як для сьогоденного суспільства, так і майбутніх нащадків.

Для проведення нашого дослідження нам необхідно розглянути деякі документи, що мають фундаментальне значення як для міжнародної спільноти загалом, так і кожної окремої країни зокрема.

Передусім слід назвати UNESCO. Відомо, що це міжнародна організація, яка докладає значних зусиль для захисту та збереження світової культурної спадщини. Щодо масштабів діяльності UNESCO свідчить значна кількість документів, розроблених і прийнятих під її егідою, а саме:

- «Конвенція про захист культурних цінностей у випадку збройного конфлікту» (1954 р.) [9];

– «Конвенція про заходи, спрямовані на заборону та запобігання незаконному ввезенню, вивезенню та передачі права власності на культурні цінності» (1970 р.) [10];

– «Конвенція про охорону всесвітньої культурної і природної спадщини» (1972 р.) [11];

– «Конвенція про охорону нематеріальної культурної спадщини» (2003 р.) [12];

– «Конвенція про охорону та заохочення розмаїття форм культурного самовираження» (2005 р.) [13] тощо.

Так, «Конвенція про охорону всесвітньої культурної і природної спадщини» визначає заходи, яких має вживати кожна держава-учасниця для забезпечення ефективного захисту, збереження та популяризації культурної і природної спадщини на своїй території, зокрема:

– здійснювати політику, «спрямовану на надання культурній і природній спадщині певних функцій у житті суспільства та створювати служби охорони, збереження і популяризації спадщини»;

– сприяти науково-технічним дослідженням і розробкам з метою усунення можливих загроз;

– здійснювати «комплекс наукових, законодавчих, технічних і фінансових заходів для виявлення, збереження, охорони, популяризації та реставрації спадщини»;

– створювати і розвивати національні та регіональні навчальні центри з питань охорони, збереження, популяризації та реставрації об'єктів спадщини;

– започатковувати і розгортати «національні й регіональні навчальні центри з питань охорони, збереження, популяризації та реставрації об'єктів спадщини» [60, с. 233].

Починаючи з 1990-х років, ЮНЕСКО приділяє особливу увагу концептуальним засадам збереження культурної спадщини в епоху глобалізації, охороні документальних свідчень, концепції цифрової культури та записаної цифрової спадщини. Саме визнання значущості «захисту записаної спадщини та необхідності координації відповідних дій між країнами спричинило започаткування у 1992 р. Програми ЮНЕСКО «Пам'ять світу» [25, с. 33; 45]. Загальновідомо, що основною метою діяльності Програми є захист і збереження об'єктів спадщини, які можуть бути втрачені назавжди. Програма затвердила основні напрямки діяльності, з-поміж яких найголовнішими є такі:

- 1) сприяння збереженню зафіксованої спадщини з використанням найбільш прийнятних методів;
- 2) підтримка доступу громадськості до спадщини шляхом створення цифрових копій;
- 3) сприяння розумінню важливості документної спадщини в міжнародному масштабі.

Відповідно до цілей Програми було створено Реєстр «Пам'ять світу», який включає перелік об'єктів спадщини, що мають глобальне значення. Програма також заохочує країни захищати зафіксовану спадщину, створювати відповідні реєстри, підвищувати обізнаність, а також розробляти та координувати політику щодо забезпечення захисту спадщини, що потрапляє чи перебуває під загрозою.

Програма також заохочує країни захищати зафіксовану спадщину, створювати відповідні реєстри, підвищувати обізнаність, а також розробляти та координувати політику щодо забезпечення захисту спадщини, що потрапляє чи перебуває під загрозою зникнення. Саме за такого сприяння багато країн світу розробили власні реєстри національної культурної спадщини (див. рис. 1.4).

Грузія	<ul style="list-style-type: none"> • нерухомі, рухомі пам'ятки культури та нематеріальна культурна спадщина
Естонія	<ul style="list-style-type: none"> • Реєстр культурних пам'яток
Польща	<ul style="list-style-type: none"> • Реєстр пам'яток
Філіппіни	<ul style="list-style-type: none"> • Філіппінський реєстр культурних цінностей
Хорватія	<ul style="list-style-type: none"> • Реєстр культурних цінностей Республіки Хорватія
Чехія	<ul style="list-style-type: none"> • Централізований список культурних пам'яток Чехії
Україна	<ul style="list-style-type: none"> • Державний реєстр нерухомих пам'яток України; • Національний перелік елементів нематеріальної культурної спадщини в Україні; • Державний реєстр національного культурного надбання. Перелік унікальних музейних предметів Музейного фонду України; • «Інформація про викрадені культурні цінності з музеїв державної та комунальної форм власності за період 2010–2015 рр.

Рис. 1.4 – Приклади реєстрів культурної спадщини деяких країн
Розроблено автором

Аналогічну мету має Закон України «Про внесення змін до Закону України «Про культуру» щодо впровадження та ведення електронних реєстрів обліку та управління культурною спадщиною та культурними цінностями України» (2023 р.) [1]. Однак, у законі є певні недоліки. Наразі відсутнє правове регулювання збору, зберігання, захисту, обліку, репрезентації, відображення та ведення реєстраційних даних про культурну спадщину (нерухоме, нематеріальне та рухоме майно). Відсутніми є також і

нормативно-правові акти «про електронну систему як інформаційно-комунікаційну систему, що гарантує надання інформації, оброблення та відповідну реєстраційну інформацію, а також про державні інституції, що гарантують створення, функціонування та ведення таких публічних електронних реєстрів, створення сучасних центрів культурних послуг як багатофункціональних закладів культури та надання сучасних культурних послуг є проблематичним» [36, с. 60-61].

Безсумнівно, що доступ до оцифрованих архівних документів в онлайн-режимі відкриває безмежні можливості для досліджень, освіти та збереження культурної спадщини. Однак у цьому процесі також необхідно враховувати різні юридичні нюанси.

Найперше, це стосується питання щодо авторського права. Слід враховувати таку деталь: «якщо термін дії авторського права на документ вже закінчився (зазвичай через 70 років після смерті автора), документ можна вільно використовувати і поширювати»[2], оскільки він стає суспільним надбанням.

Наступний пункт: якщо авторське право все ще чинне, необхідно отримати дозвіл від власника авторських прав або використовувати документ за ліцензією.

Не варто забувати й про колективне управління авторськими правами: для використання творів (музики, фотографій) більш ніж одного автора доведеться звернутися до відповідної організації.

Існують також такі поняття, як «суміжні права», «права виконавця» (якщо документ містить виконання музичного твору, необхідно враховувати права виконавця), «права на запис»(використання записів також вимагає дозволу власника авторських прав).

Закон унормовує також права інтелектуальної власності щодо бази даних, оскільки оцифрований архів вважається базою даних і на нього поширюються спеціальні положення про захист; нових творів (якщо в

процесі оцифрування створюються нові твори (наприклад, переклади та адаптації), вони також є об'єктом авторського права).

Усі ці нюанси в Україні регулюються Законами «Про авторське право і суміжні права» [2], «Про інформацію» [7], «Цивільним кодексом України» [16].

Варто наголосити, що під час оприлюднення в мережі тих масивів інформації, що оцифровують за міжнародної співпраці, усі учасники процесу мають дотримуватись не тільки вимог законодавства своєї країни, але й правих актів країн своїх колег та міжнародних стандартів.

Зважаючи на все викладене, хочемо наголосити на важливості розвитку архівної галузі в Україні, який закріплений у «Стратегії захисту документної спадщини як запоруки збереження національної ідентичності та державності на період до 2027 року та затвердження операційного плану заходів з її реалізації у 2025-2027 роках» [15]. Цим документом наша країна затверджує створення сучасної, ефективної та інноваційної архівної системи, яка відповідатиме актуальним потребам українського суспільства.

Ця Стратегія розроблена з урахуванням необхідності відмовитися від застарілих радянських підходів до управління архівною справою та адаптувати її до нових реалій, що уможливило реалізація трьох головних стратегічних цілей: «цифровізація сфери архівної справи; забезпечення збереженості наявих та майбутніх документів; вдосконалення комунікаційної політики» [15].

Отже, діджиталізація та забезпечення вільного доступу до документної спадщини нерозривно пов'язані з рівнем інформатизації країни та розвитком інформаційного суспільства.

Ставлення міжнародної спільноти до оцифрування як до пріоритетного напрямку в загальному розвитку держави зумовило прийняття низки відповідних законів. Нормативно-правове регулювання процесу оцифрування є необхідною умовою для ефективного збереження, доступу

та використання архівних матеріалів, що, насамперед, сприяє розвитку суспільства та збереженню історичної пам'яті нації.

Аналіз лише невеликої кількості документів дає підстави стверджувати, що світове співтовариство доклало значних зусиль для розвитку та інтеграції єдиного культурного простору. Культурна політика в цій царині характеризується наявністю ключових напрямів захисту та поширення культурної спадщини для майбутніх поколінь.

Пріоритетними завданнями для забезпечення ефективності культури в інформаційному суспільстві є охорона культурної спадщини та цифрова безпека, використання ІКТ для забезпечення доступу до культурної спадщини, а також використання оцифрованих даних та права на ці об'єкти.

РОЗДІЛ 2. ДІДЖИТАЛІЗАЦІЯ ДОКУМЕНТНОЇ СПАДЩИНИ: ВІД ГЛОБАЛЬНИХ ПРОЄКТІВ ДО РЕГІОНАЛЬНОЇ ПРАКТИКИ

2.1 Досвід іноземних проєктів з оцифрування архівів

Натепер міжнародні ініціативи в сфері оцифрування культурної спадщини набирають обертів, демонструючи, як важливо об'єднувати зусилля для збереження світового культурного надбання. Завдяки цим проєктам, унікальні архівні артефакти стають доступними для широкого кола користувачів, а країни отримують цінний досвід для збереження своєї культурної ідентичності в цифровому вигляді. Тобто, сьогодні оцифрування є невід'ємною частиною стратегій збереження таких цінностей. Міжнародні проєкти демонструють, як технології можуть допомогти нам зберегти наші надбання для майбутніх поколінь.

Україна, як країна з багатою та різноманітною культурою, має значний потенціал для оцифрування своєї спадщини. Однак, війна та політичні потрясіння останніх десятиліть призвели до значних втрат культурних цінностей та ускладнили процес їхнього збереження. Саме тому співпраця з міжнародними проєктами з оцифрування набуває особливої актуальності.

На нашу думку, важливим аспектом нашого дослідження є аналіз досягнень міжнародної спільноти щодо збереження культурної спадщини. Представимо деякі напрацювання у цьому напрямку.

Так, колекція стародруків з антверпенського музею Плантена–Моретуса, де колись працював один із найвідоміших європейських друкарів, незабаром стане доступною для мільйонів людей завдяки ініціативі Google. Компанія береться за масштабний проєкт з оцифрування понад 100 000 стародруків з бельгійського міста Антверпен. Ця цінна колекція, що охоплює понад 32 000 книг з Музею Плантена–Моретуса та 60 000 видань з Бібліотеки Гендрика Консьянса, незабаром буде опублікована для широкого загалу в онлайн-форматі [61].

Стало відомо, що протягом наступних кількох років Google оцифрує величезну колекцію стародруків, опублікованих між XVI та XIX століттями. Оцифровані книги будуть доступні для вільного доступу через Google Books та каталоги самих бібліотек. Особливо цінним є те, що книги, які вже не захищені авторським правом, можна буде знайти за допомогою повнотекстового пошуку. Це значно полегшить роботу дослідників і звичайних читачів. Оцифрування триває з 2021 року і добігає кінця. Отож зовсім скоро цей архів буде повністю доступним для користувачів.

Варто також зауважити, що Google Books – це амбітний проєкт американської транснаціональної корпорації Google, який було створено у 2004 році і який спрямований на демократизацію доступу до знань. Шляхом оцифрування книг і журналів, Google надає можливість кожному користувачу інтернету безкоштовно отримати доступ до нескінченних джерел інформації.

Співпраця держави з культурними інституціями є ключовим фактором розвитку культурної сфери. Яскравим прикладом такого співробітництва є платформа Europeana, яка об'єднує зусилля тисяч європейських установ, які прагнуть зробити культурну спадщину доступною для всіх.

Europeana – це найбільша в Європі платформа цифрової культурної спадщини, що надає доступ до колекцій понад 4000 установ. Тут користувачі мають змогу переглянути мільйони цифрових творів мистецтва, таких як картини, книги, музичні записи, відео, 3D-моделі тощо. «База даних Europeana містить понад 31 мільйон зображень, 24 мільйони текстів, 635 тисяч звукозаписів, 362 тисячі відео та понад 5 тисяч 3D-моделей» [69].

За свідченням Дослідницької служби Європейського парламенту, платформа Europeana висловила солідарність з Україною, розмістивши на своїх ресурсах оцифровані колекції української культурної спадщини. Паралельно з цим, проєкт «Збережемо українську культурну спадщину онлайн» (SUCHO) мобілізував понад 1300 фахівців із різних галузей, щоб

«врятувати від знищення унікальні об'єкти культурної спадщини України та перенести їх у цифровий простір» [36, с. 64].

Головна ідея застосунку Time Machine та електронного масиву Big Data of the Past полягає в об'єднанні даних від різних проєктів з метою створення єдиної цифрової платформи, яка відобразить всю різноманітність європейської історії. Time Machine – це монументальний проєкт, який «використовує передові технології збору, обробки та візуалізації архівних даних» [65] для створення багатовимірної моделі європейської історії. За допомогою оцифрування, машинного навчання та інтерактивних інтерфейсів, проєкт «дозволить дослідникам і широкій аудиторії зануритися в минуле і отримати нові знання про нашу спільну спадщину» [65].

Завдяки співпраці з американським Музеєм і бібліотекою рукописів Хілл, Острозький музей книги та друкарства розпочав масштабний процес оцифрування своїх цінних рукописів. Ідея цього проєкту виникла після вдалого експерименту в 2019 році, коли острозькі відвідувачі могли віртуально «гортати» оцифровані книги на планшетах. Саме тоді «представники американського музею звернули увагу на потенціал наших фахівців та запропонували свою допомогу» [55].

У межах міжнародного культурного проєкту Острозький музей передає цифрові копії своїх рукописів до Університету Абатства Сент-Джон у США. Ця співпраця дозволила зберегти та убезпечити цінні історичні документи, а також зробити їх доступними для широкого кола дослідників. Оцифровані рукописи з Острога поповнили багату онлайн-колекцію американського університету, яка вже містить понад 300 тисяч копій рукописів з різних куточків світу.

Російська агресія завдала непоправної шкоди українській культурній спадщині. Однак, завдяки підтримці міжнародних партнерів, зокрема Чеської Республіки, вдається поступово відновлювати втрачене. Програма «Ковчег» стала важливим кроком у цьому напрямку. Завдяки сучасному обладнанню, яке було передано в постраждали регіони, ми змогли

оцифрувати унікальні документи та зберегти їх для нащадків. Тоді як запланована програма «Ковчег-2» відкрила «нові можливості для збереження та популяризації української культурної спадщини на міжнародному рівні» [56].

Важливим є й інший спільний проєкт США та Міжнародного альянсу із захисту спадщини в зонах конфлікту (ALIPH), спрямований на підготовку співробітників Державної архівної служби до роботи з архівними документами в надзвичайних ситуаціях. У межах цього проєкту архівісти матимуть змогу навчатися безпечно поводитися з пошкодженими документами та переводити їх у цифровий формат, щоб захистити від втрати. Збережені архівні матеріали українською мовою є надзвичайно цінними, оскільки вони відображають унікальну історію України та нашу ідентичність. Ця ініціатива, що фінансується на суму 645 000 доларів США, «є частиною більш широкої програми підтримки культурної спадщини України» [54].

Потреба адаптуватися до нових викликів змусила польські архіви змінити свої пріоритети. Ці зміни знайшли своє відображення у низці стратегічних документів, першим з яких стала «Програма оцифрування культурної спадщини та збирання, зберігання і надання доступу до цифрових об'єктів у Польщі 2009–2020». Цей документ «заклав фундамент для подальшого розвитку оцифрування в країні» [80].

Нагадаємо, що процес створення польських цифрових ресурсів розпочався з ініціативи Міністерства культури та національної спадщини Польщі. Виконання робіт з оцифрування архівів було доручено Генеральній дирекції державних архівів. Було спільно розроблено Програму на 2009-2020 роки, що окреслила стратегічні завдання цього масштабного проєкту, а саме:

«По-перше, збереження для майбутніх поколінь доробку польської культури шляхом масового оцифрування архівних документів, музейних

предметів, творів мистецтва, писемних та аудіовізуальних матеріалів і безпечне зберігання цифрових ресурсів.

По-друге, розроблення та впровадження технічних стандартів та стандартів метаданих, що забезпечують функціонування цифрових ресурсів у польських та світових мережах.

По-третє, забезпечення загального доступу польських громадян до цифрових ресурсів культурної спадщини Польщі.

По-четверте, збільшення ефективності і конкурентоспроможності польських підприємств та установ, які діють у сфері культури» [51, с. 115].

Усі готові оцифровані матеріали потім викладаються на вебпорталі «Szukajwarchiwach.pl», де надається доступ до понад 38 мільйонів оцифрованих архівних документів з різних установ Польщі. Зазначимо також, що ресурс містить детальні описи документів, що дозволяють легко знайти необхідну інформацію. Привабливим є те, що «користування порталом є безкоштовним і не вимагає реєстрації» [80].

Польські дослідники з «Центру діалогу імені Юліуша Мерошевського» та «Інституту національної пам'яті» продемонстрували ефективність міжнародної співпраці в сфері історичних досліджень. Створена ними вебплатформа «Жертви антипольського терору в Радянському Союзі 1934–1938» стала «важливим ресурсом для вивчення трагічних сторінок польсько-радянських відносин» [79].

Мета створення цієї БД – заповнити прогалини в історичних дослідженнях, скориставшись інформацією з родинних архівів. Цей ресурс не лише допомагає вченим, а й є важливим джерелом для нащадків репресованих. Проект став яскравим прикладом того, «як наука може слугувати інструментом збереження історичної правди та протистояння фальсифікаціям» [49, с. 152].

Щоб отримати повну об'єктивну картину щодо тогочасних подій, дослідникам необхідно було проаналізувати широкий спектр документальних джерел НКВС, зокрема особові справи, протоколи слідчих

дій та інші матеріали. Наразі вільний доступ до цих архівів забезпечено лише на території України. Тоді як аналогічні документи, що перебувають на території росії, піддаються системному знищенню задля приховування реальних масштабів терору та репресій тих часів.

Безсумнівно, архіви сьогодні є не лише сховищами інформації, але й важливими елементами збереження культурної спадщини. Завдяки таким міжнародним ініціативам, як до прикладу, діяльність найбільш потужної в світі генеалогічної організації FamilySearch, ми маємо можливість зберегти пам'ять про наших предків та передати її наступним поколінням. Це важливо для збереження культурної ідентичності та розуміння нашого місця в світі загалом.

Відомо, що FamilySearch (США) – це міжнародна організація, яка допомагає людям досліджувати свою родову історію. Її заснувала Церква Ісуса Христа Святих останніх днів. На сайті FamilySearch зібрані мільйони записів з архівів і бібліотек різних країн, включно з України. Це можуть бути церковні записи, переписи населення, документи про шлюби та народження, військово-облікові документи, імміграційні записи, а також фотографії з їхнім описом. Крім того, FamilySearch має «мережу центрів по всьому світу, де волонтери надають безкоштовну допомогу у проведенні генеалогічних досліджень» [71].

Платформа пропонує користувачам пошук інформації про предків та можливість створення персоналізованих генеалогічних дерев. Реєстрація на сайті надає доступ до глобальної бази даних, що містить понад 1,2 мільярда імен.

Крім того, ресурс використовується для релігійних цілей, зокрема, для пошуку предків, що відійшли у засвіти. Ці персоналії використовують під час обрядів за померлих у мормонських храмах.

Загалом, історія розвитку платформи тісно пов'язана з історією Церкви Ісуса Христа Святих останніх днів і розпочалася ще у XIX столітті. Упродовж цього часу було зібрано та оцифровано величезну кількість

історичних документів, які зберігаються у спеціально обладнаних сховищах. Натепер мережа центрів сімейної історії охоплює весь світ, надаючи користувачам зручний доступ до генеалогічних ресурсів.

Ще у лютому 2014 року FamilySearch розширив свої можливості завдяки партнерству з провідними генеалогічними платформами, з-поміж яких Ancestry.com, MyHeritage, Findmypast, Geneanet, Foursquare, що, передусім, уможливило створити масштабну мережу обміну даними і значно збільшити кількість доступної інформації для дослідників родоvodu.

Прикметно, що нині на сайті FamilySearch зібрано величезну колекцію оцифрованих генеалогічних документів з-понад 110 країн.

Крім доступу до архівів, користувачі FamilySearch можуть скористатися безкоштовними освітніми матеріалами, форумами та іншими інструментами, які допоможуть їм самостійно вивчати свою сімейну історію. Зауважимо, що всі ці документи було створено на основі мільйонів мікрофільмів зі знаменитого сховища, що має назву «Гранітна гора» [71].

Загальновідомо, що сховище документів у Гранітній горі – це підземний архів Церкви Ісуса Христа Святих останніх днів, створений для довготривалого зберігання мільйонів мікрофільмів, що містять генеалогічні дані. Цей унікальний об'єкт, розташований на глибині 180 метрів. Знаходиться він неподалік від міста Солт-Лейк-Сіті (штат Юта, (США)). Сховище обладнано системою безпеки найвищого рівня, що містить озброєну охорону та 14-тонні двері, здатні витримати навіть ядерний вибух. Також «забезпечуються оптимальні умови для збереження документів завдяки клімат-контролю та посиленому захисту» [19]. Варто підкреслити, що сховище у Гранітній горі входить до десятки найбільш безпечних і захищених місць на земній кулі.

Отже, об'єднання зусиль з іншими країнами у рамках міжнародних ініціатив є ключовим для ефективного цифрового збереження української архівної спадщини. Цей підхід дозволить не лише зберегти наше культурне надбання, але й інтегрувати його у світовий культурний контекст.

Сучасні реалії підштовхують нас до пошуку ефективних методів цифровізації, які дозволять створити єдиний доступ до культурної спадщини та зберегти її для майбутніх поколінь.

2.2 Співпраця Державного архіву Полтавської області з некомерційною організацією FamilySearch щодо діджиталізації архівних фондів

Міжнародна співпраця Державного архіву Полтавської області (ДАПО) у сфері оцифрування фондів є важливим аспектом сучасного розвитку архіву та сприяє збереженню культурної спадщини регіону. Співпраця з міжнародними партнерами є досить корисною і дозволяє ДАПО мати низку привілеїв, зокрема:

– Отримати доступ до новітніх технологій: завдяки співпраці з провідними світовими архівами та науковими установами, ДАПО може ознайомитися з найсучаснішими розробками в галузі оцифрування, зберігання та доступу до цифрових архівів.

– Підвищити якість оцифрування: обмін досвідом з міжнародними партнерами дозволяє вдосконалювати технологічні процеси оцифрування, забезпечуючи високу якість цифрових копій архівних документів.

– Розширити доступ до архівних матеріалів: спільні проєкти з міжнародними партнерами сприяють популяризації архівних фондів ДАПО на світовому рівні, забезпечуючи ширший доступ до них для науковців, істориків та широкої громадськості.

– Уможлиблює залучення додаткового фінансування: міжнародні проєкти часто передбачають фінансування, яке дозволяє ДАПО придбати необхідне обладнання та програмне забезпечення для оцифрування, а також оплатити роботу фахівців.

Слід нагадати, що задля збереження фондів існує загальнодержавна «Програма оцифрування архівних інформаційних ресурсів на 2022–2025 роки», що є логічним продовженням заходів із пріоритету Державної

архівної служби «Відкриті архіви». У ній записано, що «прозора, зрозуміла та ефективна комунікація з партнерами та стейкхолдерами є невід'ємним складником забезпечення безперешкодного доступу до архівних інформаційних ресурсів, підвищення якості архівних послуг відповідно до потреб користувачів» [14].

Крім того, важливо, що ця програма, зорієнтована на очікування сьогоdnішнього користувача, й охоплює два самостійні та, водночас, взаємопов'язані напрями: оцифрування документів Національного архівного фонду й переведення в електронну форму довідкового апарату, які віднині носитимуть системний і цілеспрямований характер.

Наразі ця діяльність відбувається за допомогою техніки, яку надали державні структури, уряд Німеччини за програмами допомоги архівним структурам України та співпраці з корпорацією «FamilySearch».

Як уже зазначалося вище, FamilySearch як міжнародна некомерційна організація «має на меті дослідження генеалогічних зв'язків в усьому світу, збереження важливих сімейних записів і забезпечення вільного доступу до них в мережі Інтернет» [67]. Саме тому співробітництво з архівами України було пріоритетним завданням для FamilySearch.

Варто підкреслити, що впродовж майже двох десятиліть, з 1993 по 2011 рік, Церква Ісуса Христа Святих останніх днів здійснювала масштабний проєкт з мікрофільмування українських архівів. Після завершення цього етапу через позицію Державної архівної служби України, «усі отримані мікрофільми були оцифровано за допомогою сучасного обладнання та розміщені в онлайн-доступі на платформі FamilySearch» [71].

Згодом співпрацю Укрдержархіву та FamilySearch International було відновлено в червні 2020 року, і основною її спрямованістю стало оцифрування генеалогічних документів. Завдяки цій ініціативі, архіви кількох областей України долучилися до проєкту, з-поміж них був і Державний архів Полтавської області.

Безумовно, що така співпраця стала можливою завдяки компаніям-підрядникам, які підпорядковуються головному офісу правління через кураторів проєктів та безпосередньо перебувають на території державних архівів. Особливості комунікації та обов'язки сторін таких трьохсторонніх домовленостей можна окреслити так:

- архів виділяє приміщення та інвентар; також зобов'язується надавати документні матеріали відповідно до умов Договору та його додатків; здійснює консультування операторів щодо нюансів зазначення метаданих;

- підрядник навчає операторів та несе відповідальність за їхню роботу щодо всього процесу оцифрування; здійснює посередницьку комунікацію;

- корпорація надає установки та техніку, ліцензійне ПЗ; встановлює технічні норми до готового інформаційного продукту та здійснює аудит якості; надсилає оцифровані масиви архіву; проводить оплату праці всіх задіяних працівників [72].

Після початку проєкту в ДАПО стали отримувати діджиталізовані копії своїх оригіналів та могли зберігати їх власними силами, а також надавати користувачам за їхнім запитом замість фізичних одиниць зберігання. Архів також розміщує посилання на ці оцифровані ресурси на своєму офіційному сайті у вигляді гіперпосилань на цифрові архіви FamilySearch.

Так, за умовами договору з корпорацією FamilySearch оцифровано та виставлено 296 391 справ із фондів Державного архіву Полтавської області (станом на 14.10.2024) [24].

«У 2023 році оцифрували 141 632 одиниці зберігання, 4 193 364 аркуші архівних документів, тоді як у 2022 році було досягнуто показника в 136 360 одиниці зберігання та 3 339 978 аркушів. А всього у співпраці з генеалогічною корпорацією оцифровано біля 400 тисяч одиниць зберігання, більше 9 млн. аркушів, що складає 80% від кількості документів генеалогічного характеру і майже 30% від загальної кількості справ, що

перебувають на обліку у державному архіві», – зі звіту директора Державного архіву Полтавської області за 2023 рік, що відображено на рис. 2.1 [47].

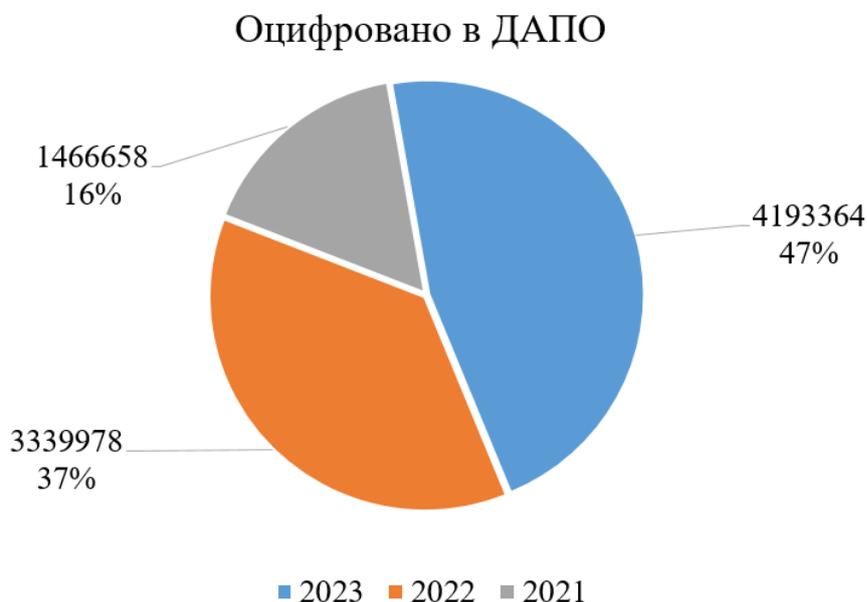


Рис. 2.1 – Кількість оцифрованих аркушів зберігання за весь час співпраці з некомерційною організацією FamilySearch
Розроблено автором

Раніше Анатолій Хромов у своїх заявах стверджував, що «на повне оцифрування НАФ України необхідно не менше ніж 1000 років» [17]. Але після декількох років сумлінної праці операторів з оцифрування голова архівної служби отримав реальні показники успішності роботи, виконав обчислення та зробив нову заяву. «Лідером із виконання умов договорів серед архівів 2023 року лишається Державний архів Полтавської області. У співпраці з генеалогічною корпорацією FamilySearch робота з оцифрування документів виконується з такою швидкістю, що станом на сьогодні документи архіву вже можна оцифрувати за 8 років», – зазначив Анатолій Хромов у своєму щорічному звіті «Про роботу Укрдержархіву, архівних установ і спеціальних установ страхового фонду документації у 2023 році та пріоритети на 2024 рік» [44].

Варто обов’язково підкреслити, що співпраця Державного архіву Полтавської області з міжнародною генеалогічною корпорацією

FamilySearch може слугувати яскравим прикладом успішного міжнародного співробітництва в галузі оцифрування архівних фондів. Це партнерство відкриває нові можливості для дослідників, генеалогів та широкого загалу користувачів отримати доступ до унікальних історичних документів, що зберігаються в архіві.

Загалом, слід наголосити, що співпраця з FamilySearch є важливою для вирішення таких проблем:

– глобальний доступ до інформації: оцифровані документи, розміщені на платформі FamilySearch, стають доступними для користувачів з усього світу, що значно розширює аудиторію дослідників;

– збереження культурної спадщини: оцифрування допомагає зберегти оригінали документів від пошкоджень і забезпечує їхнє довготривале збереження;

– просування генеалогічних досліджень: співпраця сприяє розвитку генеалогічних досліджень, дозволяючи людям досліджувати свою родову історію;

– підвищення престижу українських архівів: така співпраця демонструє відкритість українських архівів до міжнародного співробітництва та сприяє підвищенню їхнього престижу на світовій арені.

Під час нашого дослідження ми мали змогу пересвідчитися, що переважно оцифровують документи особового походження, такі, як: метричні книги, переписи населення, листування, що дозволяють досліджувати родинні зв'язки, історію окремих сімей та громад.

Для користувачів така співпраця надає низку переваг, з-поміж яких можна окреслити такі:

1. Загальний пошук: інтерфейс платформи FamilySearch дозволяє здійснювати пошук за різними критеріями, що надає більш чіткі результати.

2. Онлайн-доступ: документи доступні онлайн у будь-який час і з будь-якого пристрою з доступом до Інтернету.

3. Можливість завантаження копій: користувачі можуть завантажувати копії документів для подальшого використання.

4. Мультимовний інтерфейс: платформа підтримує кілька мов, що робить її доступною для широкого кола користувачів.

Однак оцифрування має і свої недоліки, найголовнішими з яких, на нашу думку, є такі:

- оцифрування великих обсягів архівних матеріалів дорого вартує, оскільки потребує значних фінансових інвестицій;

- це тривалий процес, який може тривати роками, особливо для великих архівів;

- збереження цифрових даних вимагає постійного оновлення обладнання та програмного забезпечення, а також в сучасних українських реаліях існує ризик втрати даних через збої або кібератаки;

- якість оцифрованих документів може бути різною і залежить від багатьох факторів, до прикладу, таких, як: стан оригіналу, обладнання для сканування та налаштування процесу.

З огляду на практичне застосування платформи з пошуку генеалогічної інформації від FamilySearch, вона має низку власних специфічних недоліків, які перешкоджають вільному доступу до інформації. Цій проблемі присвятив чимало досліджень, зокрема, Гаррет Маст. Під призмою вивчення Шотландії, він, насамперед, порівнював FamilySearch Research Wiki, що надає цінний ресурс генеалогічних довідок для країн з усього світу, з такими сайтами, як Wikipedia. Вебсторінки для кожної шотландської парафії відійшли на другий план, і більшість з них не отримували жодних суттєвих оновлень протягом десяти років. В аналізі проєкту, зробленому дослідником, висвітлюються певні процеси, вжиті для того, щоб оживити сторінки шотландських місцевостей Research Wiki після перевірки цивільної, церковної та спадкової юрисдикції для кожної парафії в Шотландії. Це, зокрема, містило створення дизайну інформаційного вікна та переформатування майже тисячі окремих вебсторінок. Науковець також

надав рекомендації, необхідні для забезпечення того, щоб «шотландські статті надавали точну та цінну енциклопедичну інформацію щодо генеалогічних ресурсів для генеалогів та істориків у найближчі роки» [75, с. 6].

З огляду на це глибоке дослідження, можна зробити висновок про те, що усунення цих та інших незручностей істотно б покращило користування інформаційними ресурсами корпорації.

Під час написання кваліфікаційної роботи ми скористалися порадами Гаррета Маста і проаналізували деякі проблеми, які змогли визначити в роботі FamilySearch. Схарактеризуємо їх.

1. Складність інтерфейсу самої платформи достатньо значна. Багато дослідників та користувачів мають проблеми з навігацією та пошуком затребуваної інформації.

2. Обмежений доступ до оригіналів, що можуть бути необхідні для більш детального дослідження. Такий варіант можливий за умов технічних збоїв чи обмеженням публікації документів у зв'язку з термінами секретності.

3. Помилки в індексації імен та інших даних, що значно ускладнює пошук потрібної інформації. Можливий також варіант часткової відсутності метаданих, що пов'язано з їхнім незазначенням операторами, відсутністю на оригіналах, людським фактором або технічними проблемами.

4. Не всі архіви повністю оцифровані, тому деякі дані можуть бути недоступні на платформі.

5. Оцифрування рукописних документів може бути проблематичним через труднощі з розпізнаванням почерку. Це стосується як визначення метаданих, так і самої вміщеної в архівній справі інформації.

Отже, співпраця ДАПО та FamilySearch продовжується з метою удосконалення архівних документів для широкого кола користувачів. Завдяки оцифруванню все більшої кількості фондів та вдосконаленню

платформи, дослідники зможуть глибше зануритися в історію, а звичайні користувачі – знайти інформацію щодо своїх предків.

Наявні недоліки сканування, опублікування інформації та її пошуку мають стати для фахівців пріоритетним завданням для подальшого розвитку та вдосконалення всієї системи оцифрування. Це дозволить зробити архівні документи ще більш доступними для дослідників та зберегти їх для майбутніх поколінь.

Незважаючи на значні досягнення, перед проєктом стоять нагальні завдання, зокрема такі, як: покращення якості сканування для забезпечення максимально точної передачі інформації; розширення метаданих, що допоможе користувачам точніше знаходити необхідну інформацію; вдосконалення інструментів пошуку; співпраця з іншими архівами та об'єднання зусиль, що дозволить створити єдиний цифровий архів України та світу.

2.3 Стан, проблеми та перспективи українського внеску в оцифрування документної спадщини

Незважаючи на війну та постійні перебої з електроенергією, українські архіви продемонстрували вражаючі результати в оцифруванні документів. На круглому столі в Українському інформаційному агентстві «Укрінформ» 25 вересня 2023 року було оголошено про значні досягнення України в галузі оцифрування архівних документів. За словами голови Державної архівної служби України Анатолія Хромова, у 2023 році українські архіви оцифрували понад 21 мільйон документів, що дозволило країні стати лідером за темпами оцифрування у світі. Україна продовжує активно працювати над збереженням історичної пам'яті та планує збільшити обсяги оцифрування в поточному році.

За результатами 2023 року, Україна увійшла до топ-9 країн за темпами оцифрування архівних документів, перевівши в цифрову форму понад 21 мільйон документів (див. рис. 2.2). Перше місце посіла Фінляндія, яка

включила до своєї статистики не лише оцифровані паперові документи, а й цифрові документи, створені електронно. Однак, якщо порівнювати виключно кількість оцифрованих традиційних архівних документів, то Україна є безумовним лідером серед цих країн. Такий висновок зробив директор Українського науково-дослідного інституту архівної справи та документознавства Віталій Скальський [57]. «Ми рахували кількість цифрових копій, тобто зроблених кадрів. Не документів, не справ чи фондів. Йдеться, зазвичай, про розвороти сторінок, по дві сторінки на один кадр. Якщо ж справи розшиті, то один кадр робиться з однієї сторінки, зокрема пустої та незаповненої», – зазначив Віталій Скальський [57].

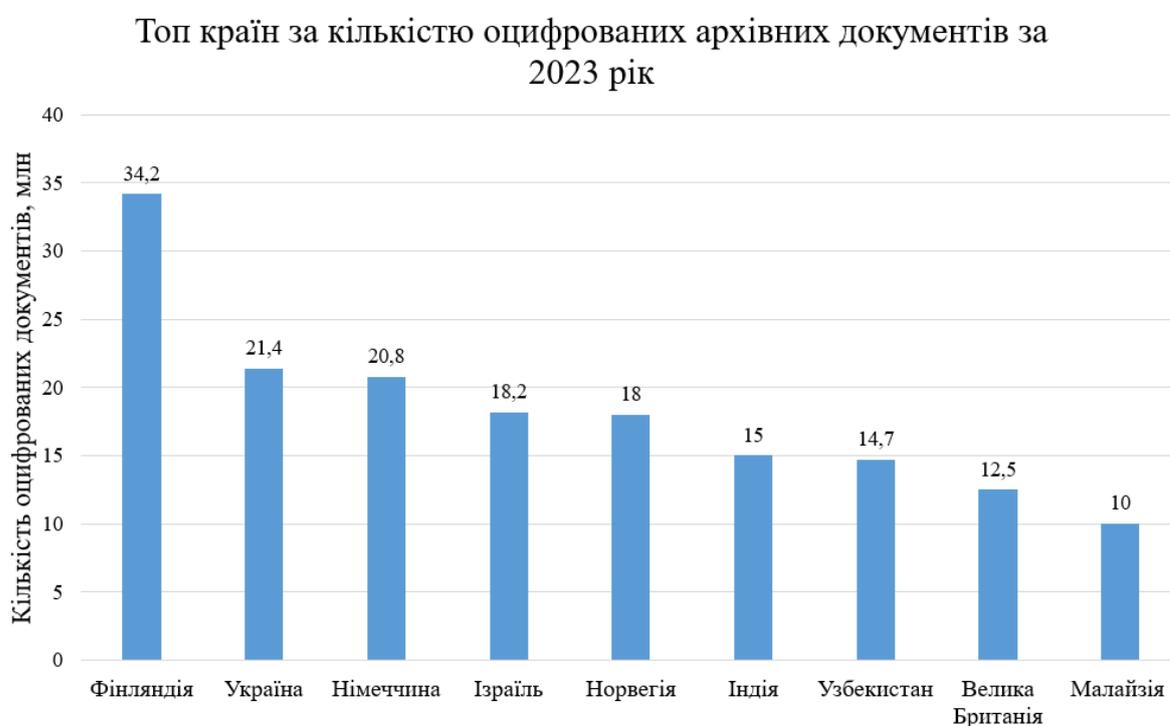


Рис. 2.2 – Топ країн за кількістю оцифрованих архівних документів за 2023 рік
Розроблено автором

Безумовно, під час війни в країні виникають проблеми з фінансуванням окремих галузей, оскільки основні кошти зосереджені на ЗСУ. Однак українські архіви досягли значних результатів в оцифруванні документів Національного архівного фонду навіть в тяжкі для країни часи. Про це свідчать дані Укрдержархіву. Важливість цієї роботи 25 вересня

2023 року підкреслили учасники круглого столу, серед яких були представники Верховної Ради, міжнародні партнери та керівники профільних міністерств.

Крім того, делегація українських архівістів на чолі з головою Укрдержархіву представляли здобутки на міжнародному симпозиумі FamilySearch International 9-10 вересня 2024 року у Потсдамі (Федеративна Республіка Німеччина) [23].

Анатолій Хромов очолив українську делегацію на заході, присвяченому обговоренню процесу оцифрування архівних документів. Разом із колегами він поділився досвідом збереження важливих історичних матеріалів в умовах кризи. Крім того, пан Хромов провів окремі зустрічі з керівництвом FamilySearch, обговоривши подальші кроки співпраці та можливості залучення до спільного проєкту всіх державних архівів України.

Кількість українських архівів, що долучилися до спільного проєкту з FamilySearch International, невпинно зростає. Станом на вересень 2024 року їх вже 27 (5 центральних та 22 обласних), що відтворено на рис. 2.3 [21].

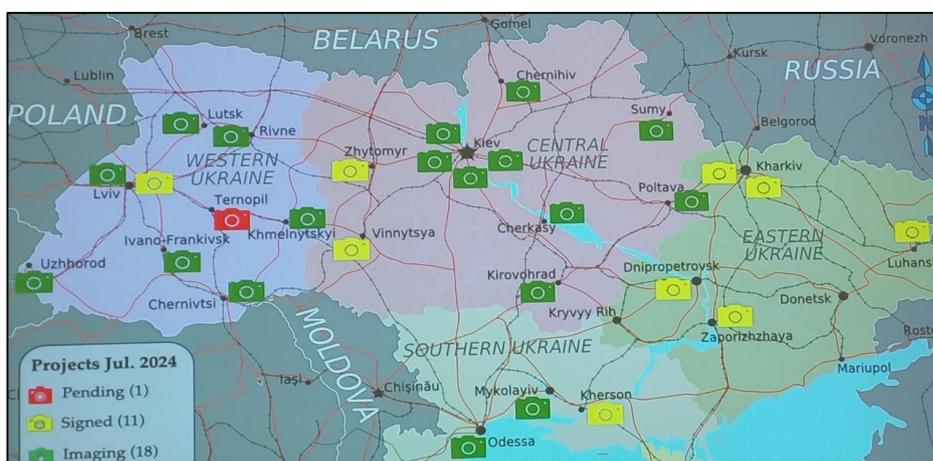


Рис. 2.3 – Мапа проєктів оцифрування FamilySearch в Україні [23]

Так, наприклад, у межах співпраці з FamilySearch, розпочатої 21 жовтня 2020 року, Державний архів Київської області (ДАКО) оцифрував 77 846 справ, що відповідає 2 337 934 файлам цифрових копій архівних документів. Відомо, що ці дані стали доступними для дослідників через інформаційно-пошукову систему «Е-Архів» [23, 41].

Продовжується співпраця з некомерційною організацією FamilySearch і в Державному архіві Хмельницької області (ДАХО). У межах співпраці вже оцифровано 102892 одиниць зберігання (57 фондів) та виготовлено 1825195 кадрів [37]. «Договір (з FamilySearch) підписали 26 листопада 2021 року, але війна скорегувала наші плани і ми почали працювати влітку 2022. Спочатку були задіяні два оператори, які оцифровували фільтраційні справи репресованих з фондів СБУ. Наразі він вже повністю оцифрований, а операторів у нас тепер четверо», – зазначила Лілія Тараненко, завідувачка сектору цифрових та інформаційних технологій Держархіву області [35].

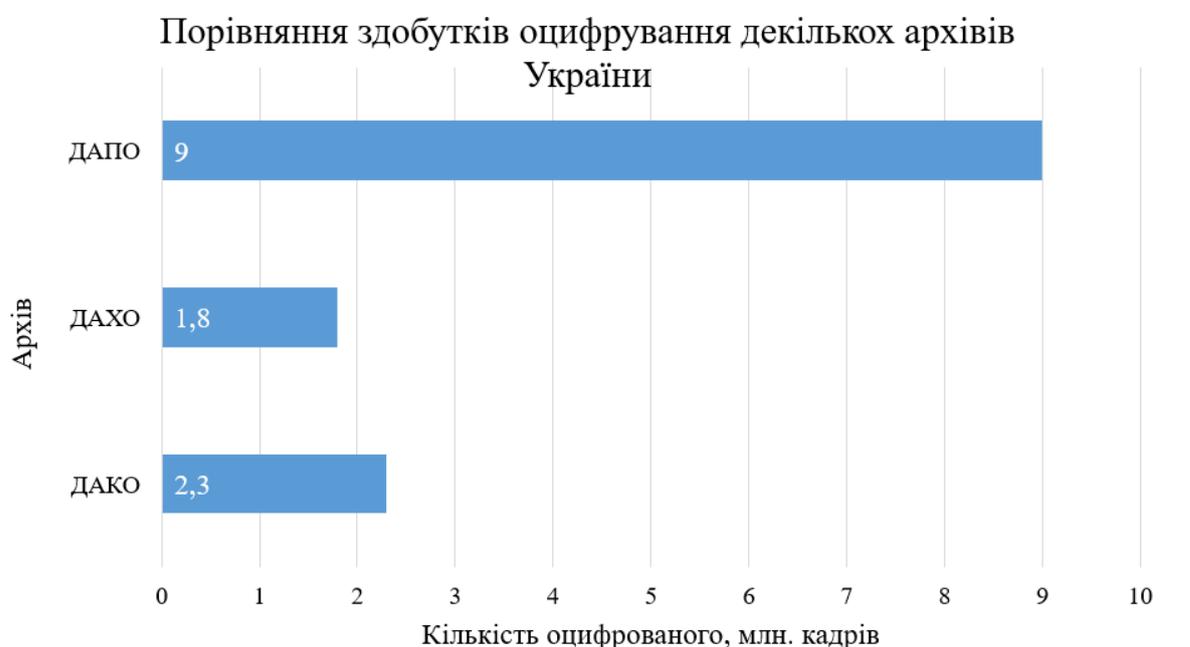


Рис. 2.4 – Порівняння здобутків оцифрування декількох архівів України
Розроблено автором

Такі результати цих обласних архівів ми можемо порівняти з лідерами галузі, аби точніше розуміти масштаби проведеної роботи. Звісно, таке порівняння не є напрочуд точним, оскільки загальний час оцифрування не є однаковим в усіх проектах. Також на загальну кількість впливає штат операторів: чим більше фахівців сканує, тим вищі показники отримуємо. Варто звернути увагу і на досвід таких працівників, бо швидкість та якість напрацьовуються та вдосконалюються з часом. І, звісно ж, стан документів,

який відіграє неабияку роль у швидкості та складності роботи. На це також впливає й формат документів: першочергово оцифровують метричні книги, які є найважливішими для генеалогічної компанії та найлегшими для оцифрування; тоді як судові справи, партійні протоколи та дані військкоматів потребують особливих навичок та значних зусиль під час оцифрування.

Варто наголосити, що сама процедура сканування документів передбачає попередню перевірку їх упорядкування та опису в електронній базі даних. Винятками є фото, фільми та аудіозаписи, які описуються вже після створення цифрової копії, що дозволяє виявити більше деталей та забезпечити точнішу інформацію про їхній зміст. Для визначення фізичного стану документів перед скануванням застосовується комплексна оцінка та розподіл на 4 групи, а саме:

Група 0. Ці документи готові до сканування відразу. Вони в такому стані, що навіть незначні пошкодження не завадять отримати чітке цифрове зображення. Весь текст буде добре видно.

Група 1. Ці документи потребують невеликої попередньої роботи. Консерватор повинен підготувати їх до сканування: вирівняти сторінки, склеїти невеликі розриви (не більше 10 аркушів). Це необхідно, щоб відкрити ті частини тексту, які зараз закриті.

Група 2. Стан документів оцінюється як середній. Багато сторінок пошкоджено і зім'ято, що унеможлиблює їх безпосереднє сканування. Для якісного оцифрування необхідно провести масштабні роботи з відновлення документів.

Група 3. Документи знаходяться у критичному стані. Великі ділянки тексту не читаються через пошкодження та попередні, невдалі спроби реставрації. Оцифрування в такому стані спричинить ще більші втрати інформації. Перед скануванням потрібна тривала і копітка робота реставратора [51, с. 123 – 124].

Якщо архівні матеріали перебувають у незадовільному фізичному стані, фахівці з реставрації повинні провести оцінювання та визначити, чи потрібна їм спеціальна обробка перед оцифруванням, а також обрати оптимальну технологію для цифрової копії. Документи, які скріплені або зшиті, бажано оцифрувати без роз'єднання. У випадку, коли розшивання є необхідним, наприклад, для проведення масової консервації, документи не слід зшивати знову. Якщо архівні матеріали уражені мікроорганізмами або комахами, перед оцифруванням їх обов'язково дезінфікують або дезінсекують. Крім того, «документи, виготовлені з паперу з високою кислотністю, підлягають процесу розкислення, який може бути як масовим, так і індивідуальним» [51, с. 125].

Безсумнівно, робота з відновлення документів є масштабним та тривалим процесом. Це можна зрозуміти на прикладі Кам'янець–Подільського архіву, який постраждав внаслідок пожежі 2003 року: тоді згоріло близько 70% фондів. Ця подія була визнана гуманітарною катастрофою європейського масштабу, а втрачені документи були безцінним джерелом для вивчення історії України. Зникли унікальні документи, що розкривали адміністративне, економічне та культурне життя Поділля протягом кількох століть. Хмельницькі реставратори, яким було передано вцілілі документи, неухильно просувалися вперед, опрацьовуючи щорічно від 4 до 6 тисяч аркушів. Протягом двох десятиліть, з 2003 по 2023 рік, було проведено колосальну роботу з ідентифікації та опису архівних матеріалів, в результаті якої було опрацьовано понад 40 000 справ [27].

Реставратор Юрій Руденко у своєму інтерв'ю 2013 року зазначив, що більше уваги приділялося документам, які не постраждали в пожежі. Однак, ті документи, які пережили вогонь, воду та тривале просушування, вже не зазнають значних пошкоджень. Натомість, інші документи, хоч і не горіли, але були уражені цвіллю, грибокком та іншими шкідниками, перебувають у критичному стані. Саме такі документи й потребують, на думку фахівця, «першочергової реставрації» [29].

Наразі ж хмельницькі оператори зі сканування діджиталізують, зокрема, й кам'янецькі залишки. Вони, а також спеціалісти з інших проєктів, мають складнощі в роботі через недостатню кількість реставраторів. Це особливо помітно, якщо порівняти 4 тисячі реставрованих одиниць на рік та 3 тисячі кадрів на день для кожного оператора, яких в тому ж ДАХО наразі четверо. Ця проблема тягне за собою необхідність видачі на оцифрування невідреставрованих справ, які того насправді потребують, а це, насамперед, впливає на якість готових зображень та витрати часу на кожен аркуш. Також не в усіх архівах дозволено розшити справу за потреби представлення прихованої в палітурці інформації.

Слід наголосити, що цифрова трансмісія даних документних матеріалів різного часу є вкрай необхідною для українських архівів, оскільки це питання їхнього виживання. Але важливо звернути увагу на деякі нюанси цього процесу. Необхідно також підкреслити, що цифрові копії мають служити людству ще впродовж тривалого часу, тому й технології варто обирати відповідно – з погляду перспективності діяльності.

Виникають певні питання і щодо самих копій, які, насамперед, повинні створюватись з максимальною якістю, навіть якщо для цього необхідно розшивати документи. До речі, європейські архіви надають переваги оцифруванню саме в такий спосіб – аркушний, оскільки там надають перевагу новим технологіям та відходять від паперових носіїв як від панівного формату. На нашу думку, сканування розворотами притаманне українській діджиталізації через бажання максимальної збереженості й так вже нашої знівченої історії та документальних джерел зокрема. Усі політичні режими, війни та стихійні лиха загартували українців та змусили надвідповідально ставитись до своєї спадщини, бажаючи не тільки отримати копії, а й зберегти оригінали хоча б в тому вигляді, який вони мають натепер.

Отже, актуальними для українських архівів залишаються питання подальшого оцифрування. Оскільки документи з генеалогічною

інформацією, що цікавлять FamilySearch, колись закінчатся в архівах, деякі проєкти поступово доходять до свого завершення. Тож, подекуди, в недалекому майбутньому корпорація припинить свою співпрацю та ліквідує проєкти з оцифрування. Зберігати наявні темпи діджиталізації спадщини без їхніх досвідчених операторів, ПЗ, техніки стане неможливим. Тож необхідно вже зараз планувати подальші дії, оскільки орієнтовний термін повного оцифрування може знов збільшитись з 8 до 100 років. Для цього необхідне достатнє фінансування галузі задля створення повноцінних робочих місць для досвідчених операторів чи наймання тих самих підрядників; закупівлі необхідного обладнання тощо.

РОЗДІЛ 3. РЕКОМЕНДАЦІЇ ЩОДО ОПТИМІЗАЦІЇ ПРОЦЕСУ ОЦИФРУВАННЯ НА ОСНОВІ ПРАКТИЧНОГО ДОСВІДУ

3.1 Аналіз параметрів оцифрування: роздільна здатність, формати файлів, використання штучного інтелекту

З огляду на все, про що йшлося в попередніх розділах кваліфікаційної роботи, цифрова трансформація суспільства не оминула й архівну справу, надавши нам потужні інструменти для збереження та використання історичних документів.

Безумовно, що переведення в цифровий формат величезних масивів архівних документів – це масштабне завдання, яке вимагає, передусім, значних ресурсів. Однак, воно є необхідним для збереження національної пам'яті та забезпечення доступності архівних матеріалів для майбутніх поколінь.

Нагадаємо, що головна мета оцифрування – забезпечити ефективний, безпечний і стандартизований перехід від паперових носіїв до цифрового формату, що дозволяє зберігати, пошукувати та обробляти інформацію більш зручно.

Не викликає сумнівів той факт, що використання високопродуктивного обладнання для оцифрування документів дозволяє не тільки прискорити процеси оброблення інформації, але й забезпечити збереження оригіналів документів та отримання цифрових копій найвищої якості. Але, важливо підкреслити, що перед прийняттям рішення щодо вибору такого обладнання, слід проаналізувати його різновиди. Наведемо кілька прикладів таких технічних засобів та коротку їхню технічну характеристику:

– Fujitsu ScanSnap iX1500: 600 dpi; до 30 сторінок за хвилину; оптичне розпізнавання символів (OCR);

– Epson WorkForce ES-500W: 600 dpi; до 35 сторінок за хвилину; підтримка облікових систем хмарних служб;

– Zeutschel zeta: до 1200 dpi; до 100 сторінок за годину; інтеграція з програмним забезпеченням для обробки;

– ST ViewScan III: до 2 400 dpi; до 60 сторінок за хвилину; автоматичне розпізнавання кадрів і налаштування параметрів сканування;

– збірні конструкції, що приєднуються до системного блоку та монітору комп'ютера завдяки спеціалізованому програмному забезпеченню. Їхні параметри залежать від можливостей ПЗ та комплектуючої техніки, а також їхньої взаємодії [39; 30, с. 99].

Загальновідомо, щоб отримати найкращу якість цифрової копії документа, необхідно використовувати високу роздільну здатність сканування. Це дозволяє зберегти всі деталі, навіть найдрібніші, і «гарантує подальше використання таких копій для різних цілей, навіть таких, як збільшення, друк або подальше редагування» [31, с. 43].

DPI (dots per inch) – це важливий параметр, який характеризує щільність растрового зображення. Він визначає кількість точок (пікселів), що припадають на лінійний дюйм. DPI «безпосередньо впливає на якість зображення під час виведення на друк або відображення на екрані з високою роздільною здатністю» [70].

Фахівці визначили рекомендовану роздільну здатність відповідно до формату та складності документа, що зазначено в таблиці 3.1.

Таблиця 3.1 – Роздільна здатність різних форматів документів.
Розроблено автором

Формат документа	Мінімальна роздільна здатність, dpi	Середня роздільна здатність, dpi
A4 і більше	300	300 – 450
менше A4	600	600 – 800
Документи з дрібними деталями	600	600 – 1200
Фотодокументи	1200	1200 – 2400

Крім того, на якість готових зображень також впливає формат збереження документів. З огляду на практичних досвід багатьох фахівців, для отримання найкращої якості сканування слід використовувати формат TIFF. Цей формат зберігає всі деталі зображення без втрати якості, тоді як для повсякденної роботи та обміну файлами достатньо формату JPG з компресією 80%. Це дозволяє значно зменшити розмір файлу без помітної втрати якості зображення.

Для створення електронних книг ідеально підходить формат PDF, оскільки його файли зручні для перегляду на різних пристроях і дозволяють зберігати форматування тексту та зображень. Для оптимізації PDF-файлів для вебперегляду бажано використовувати JPG-копії з параметрами стиснення та роздільної здатності, вказаними в таблиці 3.2. Цей підхід забезпечить баланс між якістю зображення та розміром файлу.

Таблиця 3.2 – Порівняння форматів збереження оцифрованих архівних документів
Розроблено автором

Назва формату	Мета використання	Програми, які можуть відкривати файли	Рекомендований метод стиснення
TIFF	Для майстер-копій	Більшість растрових редакторів, векторні редактори, що підтримують растрові об'єкти	Без стиснення
JPEG	Для робочих і службових копій	Більшість графічних програм, інтернет-браузери	Можливо обрати ступінь стиснення, але не більше 50% від оригіналу
PDF	Для робочих і службових копій	Міжплатформовий формат файлу, що переноситься на комп'ютери будь-яких систем. Для перегляду файлу потрібна програма Acrobat Reader	LZW (стиснення без втрати якості), JPEG

Варто зауважити, що для забезпечення сумісності та зручності використання, всі PDF – файли повинні відкриватися без проблем на комп'ютерах з операційною системою Windows 2000 і новішими версіями.

Також усі необхідні елементи документа (шрифти, картинки) повинні бути вбудовані в файл. Виняток становлять тільки файли, що містять конфіденційну інформацію, для яких може бути встановлений пароль.

Так, на рисунках 3.1 – 3.3 вміщено приклади зберігання зображень у PDF, JPEG та TIFF відповідно для візуальної оцінки якості контенту.



Рис. 3.1 – PDF. Державний архів Волинської області [33]

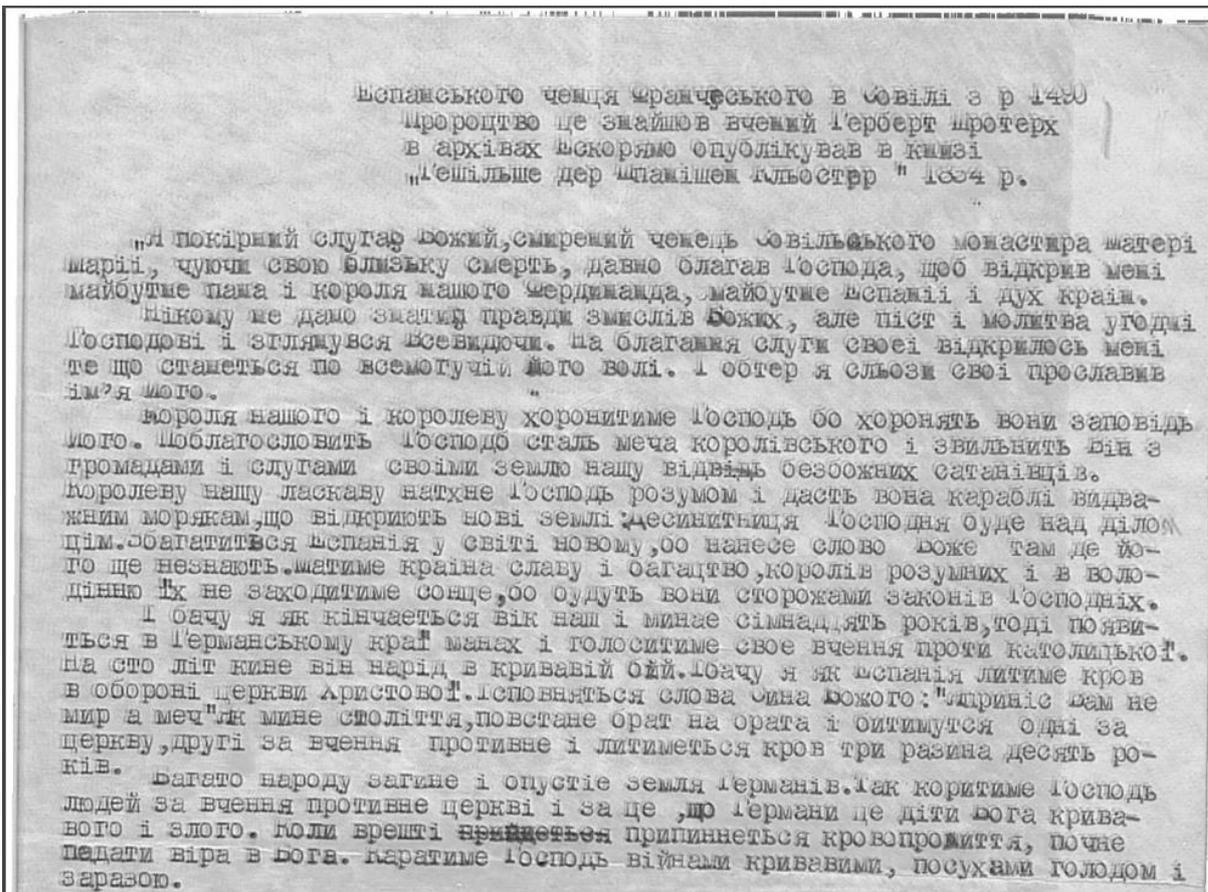


Рис. 3.2 – JPEG. Державний архів Рівненської області [46]

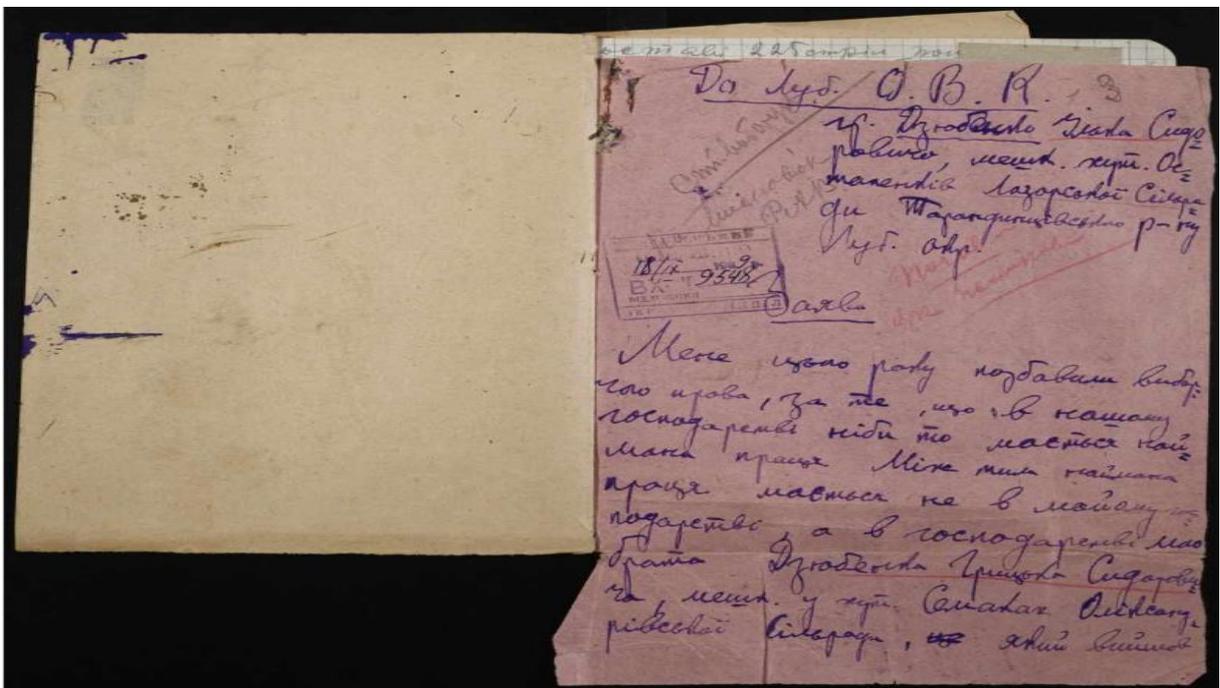


Рис. 3.3 – TIFF. Державний архів Полтавської області [81]

Тож ми можемо впевнено стверджувати, що TIFF найбільш точно передає всі, навіть дрібні, деталі оригіналу. Чи не єдиним недоліком цього формату є його розмір, оскільки для подальшого зберігання таких файлів знадобиться чимала кількість додаткової пам'яті на жорстких дисках чи хмарних сховищах.

Варто зауважити, що перевірка якості зображення базується не тільки на візуальному вигляді, що є суб'єктивним, а й на достатньо об'єктивних показниках оцінювання.

Показник оцінювання – це «інструмент, який складається з набору критеріїв, що дозволяють оцінити якість виконання завдання» [52, с. 13]. Кожен критерій деталізує вимоги до окремих компонентів завдання залежно від рівня складності. Маємо також розтлумачити поняття «критерії оцінювання» як «чітко визначений стандарт, за яким оцінюється виконання певної частини завдання» [52, с. 13]. Він дозволяє об'єктивно визначити рівень досягнення поставленої мети.

З метою об'єктивної оцінки можливостей програм для сканування документів, було визначено такі ключові критерії, як:

- універсальність сканування: «здатність обробляти як односторінкові, так і багатосторінкові документи»;
- функційність оптичного розпізнавання символів: «точність перетворення сканованого тексту в формат, який підлягає редагуванню»;
- експорт даних: «можливість копіювання та збереження розпізнаного тексту у популярних форматах»;
- пошукові можливості: «наявність інструментів для швидкого пошуку потрібної інформації в сканованих документах»;
- сегментація документів: «здатність програми розпізнавати окремі сторінки в багатосторінкових документах» [52, с. 14].

Загалом, вибір цих критеріїв обумовлений аналізом типових задач, що виникають під час роботи з оцифрованими документами, а також вивченням відповідної наукової літератури.

Враховуючи постійне збільшення обсягів інформації, що вимагає оцифрування, застосування систем штучного інтелекту (ШІ) та машинного навчання (МН) стає ключовим фактором підвищення ефективності та точності цього процесу.

Варто підкреслити, що ШІ та МН революціонізують роботу з документами. Вони дозволяють автоматично розпізнавати текст, класифікувати документи за різними критеріями, а також вирізняти рукописний текст. Це значно спрощує процес оцифрування великих обсягів інформації, забезпечуючи швидкий доступ до потрібних даних та їхній подальший аналіз.

Передусім, варто зауважити, що глибоке навчання відрізняється від інших методів МН здатністю створювати складні моделі, які імітують роботу людського мозку. Це дозволяє йому виявляти в текстових даних такі тонкі нюанси, як контекст, синтаксис і семантика, що є недоступним для традиційних методів. Ієрархічна організація цих мереж уможливорює їм «виявляти складні шаблони та зв'язки між словами в тексті, що суттєво покращує їхню здатність виконувати складні завдання розпізнавання з високим рівнем точності» [76, с. 342].

Новітні згорткові нейронні мережі (CNN) використовують спеціальний тип обчислень, згортки, щоб виявляти характерні ознаки в даних. Коли ми застосовуємо CNN до тексту, мережа шукає певні послідовності символів чи слів, які можуть бути важливими для розуміння загального змісту. Відомо, що завдяки багаторазовому застосуванню згорткових шарів, «CNN можуть виявляти як прості, так і більш складні залежності між словами, що дозволяє їм ефективно аналізувати будь-яку текстову інформацію» [74, с. 1].

Інші нейронні мережі, зокрема, рекурентні (RNN), мають внутрішню структуру, яка дозволяє їм зберігати інформацію щодо попередніх елементів послідовності. Це робить їх ефективними для задач, де контекст має вирішальне значення, до прикладу, таких, як розпізнавання мови або

генерація тексту. На відміну від звичайних нейронних мереж, «RNN можуть обробляти вхідні дані змінної довжини, що є важливою перевагою у процесі роботи з текстовими даними» [66].

Крім того, необхідно зазначити, що RNN мають внутрішній стан, який дозволяє їм зберігати інформацію про послідовність даних. Це, найперше, робить їх ідеальними для обробки текстових даних, де контекст попередніх слів має вирішальне значення для розуміння загального значення. Завдяки своїй здатності до довготривалої залежності, RNN можуть моделювати складні мовні структури, такі як синтаксис і семантика.

Ще один сучасний вид навчання, що має назву «трансферне» – це «потужний інструмент для розробки моделей розпізнавання тексту, який дозволяє зменшити час і ресурси, необхідні для навчання, та підвищити точність результатів» [32, с. 133]. Завдяки трансферному навчанню можна створювати більш точні та ефективні системи розпізнавання тексту, навіть за умови обмежених обчислювальних ресурсів.

Доволі часто використовують попередньо навчені моделі, як-от BERT чи GPT, що пройшли навчання на величезних обсягах тексту. Вони здатні виявляти складні зв'язки між словами та розуміти контекст, що робить їх універсальними інструментами для багатьох завдань обробки мови. Ці моделі працюють на основі механізмів уваги, що дозволяє їм фокусуватися на найважливіших словах у реченні та розуміти контекст загалом.

Навчання з підкріпленням (RL) – це «метод МН, який базується на принципі нагород за правильні дії та покарань за помилки» [32, с. 134]. З метою розпізнавання тексту, RL можна використовувати для покращення точності моделей, спонукаючи їх до пошуку оптимальних рішень

Якнайперше, RL – це метод навчання, який дозволяє моделі розпізнавання тексту самостійно вдосконалюватися, отримуючи зворотний зв'язок про правильність своїх відповідей.

Одним із цікавих застосувань підсилювального навчання є генерація синтетичних текстових даних, що дозволяє створювати великі обсяги

високоякісних даних для тренування моделей розпізнавання тексту, особливо коли реальної інформації недостатньо. Крім того, RL може бути використане для оптимізації стратегій розпізнавання тексту в реальному часі, дозволяючи системам адаптуватися до змінних умов та покращувати свою ефективність.

Дослідження українських і зарубіжних учених підтверджують, що системи штучного інтелекту та методи машинного навчання є потужними інструментами для оцифрування документів. Вони дозволяють точно розпізнавати, вилучати та інтерпретувати різноманітну інформацію, що міститься в документах. Постійний розвиток цих технологій обіцяє ще більш ефективні та адаптивні рішення для розвитку цифрової трансмісії даних в майбутньому.

Отже, оцифрування архівних фондів – це комплексний і ресурсомісткий процес, який потребує не лише спеціальних знань та досвіду, а й використання сучасного обладнання.

Спеціалізована техніка, правильний формат збереження зображень та достатня роздільна здатність є запорукою якісного контенту для його тривалого використання та подальшої обробки людськими силами й інструментами ШІ та МН. Такий підхід є необхідним для збереження культурної спадщини, забезпечення доступності архівних документів для широкого кола користувачів та ефективного управління інформаційними ресурсами.

3.2 Переваги та недоліки DСam як програмного забезпечення для оцифрування документів

Найперше як розглядати особливості однієї з поширених програм для оцифрування архівних документів, на нашу думку, доцільно розглянути її специфіку.

«DСam – це спеціальна комп'ютерна програма, створена FamilySearch, яка фотографує архівні записи разом із певною кількістю

пов'язаних із ними коротких історичних, географічних і технічних даних. Ці дані називаються «метаданими», – зазначено в посібнику користувача DСam [67, с. 1].

Програма надає засоби для сортування та упорядкування мільйонів архівних записів у колекції FamilySearch. Зображення й дані доступні для тих, хто шукає інформацію про своїх предків на сайті FamilySearch.

Загалом, цей інструмент забезпечує потужні можливості для сортування та організації мільйонів архівних записів на платформі FamilySearch, що дозволяє користувачам швидко знаходити необхідну інформацію про своїх предків.

Варто презентувати, як працює програма. Спочатку для безпосереднього сканування необхідно увімкнути світлодіодні світильники, камеру та системний блок з монітором. Після запуску програми DСam потрібно авторизуватись під своїми логіном та паролем для майбутньої ідентифікації оператора. Далі йде невелике оновлення даних із «Головним управлінням»: вивантаження матеріалів, отримання нових списків (або їхнє оновлення). Наступним кроком є калібрування камери для точнішої кольоропередачі. Тільки тепер оператор зможе обрати поточний проєкт та список для створення папки.

Під час занесення до шаблону метаданих зазначаються така інформація, як: архівний шифр, дати створення та завершення папки, місцезрештування (локалізація), мова, тип записів, мітки та коментарі. Два останніх поля слугують для опису стану документа чи його особливостей: це можуть бути збляклі сторінки чи помилка послідовності сторінок, що позначається міткою, або несиметрично зшитий документ чи заборона для розшиття, що детально описується в коментарі. Варто зауважити, що мітки як стандартні недоліки можна проставляти, як на весь документ, так і на окремі його сторінки, тоді як коментар слід лише прописати вручну для всього документа як нетипову особливість сканованого матеріалу.

Тільки тепер новостворену папку можна наповнювати контентом відповідно до технічних вимог сканування матеріалів. Після виконання цього етапу оператор має перевірити всю інформацію та виправити можливі неточності чи помилки своєї роботи. Це відбувається під час оцінювання у відповідній вкладці, після чого папка закривається автоматично, а оператор може перейти до наступного матеріалу, виконуючи всі зазначені вище етапи заново.

У кінці тижня всі створені папки завантажують на зовнішні жорсткі диски, які один раз на два тижні надсилають аудиторам для перевірки якості готових зображень. Готовий і тільки якісний контент аудитори вивантажують на платформу для загального доступу та користування. Контент, що містить порушення технічних вимог, повертається оператору для доопрацювання та проходить повторну перевірку аудиторамі перед його публікацією. Саме така схема руху документів забезпечує високий рівень сканованих матеріалів.

Безперечно, масштаби таких проєктів та кількість оцифрованих сторінок викликають повагу та довіру до системи та програмного забезпечення, оскільки часу на їхнє вдосконалення було достатньо. Проте і після виходу 9-ої версії DСam ще залишились багі, помилки та недоліки роботи ПЗ, які бачить чи не кожен оператор.

Так, нагадаємо, що декілька українських проєктів були під загрозою припинення сканування через відсутність відкритих списків у їхніх проєктів. Цю ситуацію не могли вирішити ні адміністратори ресурсу, ні розробники упродовж 1 – 6 місяців в різних місцях України. Проблему вирішило лиш створення нових списків після тривалого очікування операторів та роботи в неправильних списках. Унаслідок цього довелось переносити вручну кожен створену папку у відповідний список, адже робота має бути виконана бездоганно, а функція перенесення діапазону папок з одного списку в інший виявилась заблокованою в новій версії програми. Так одразу стали очевидними одночасно дві проблеми цього ПЗ:

складність створення списків адміністраторами та, можливо, баг програми, а також неможливість перенесення декількох папок.

Складність оцифрування української спадщини компанією FamilySearch полягає також у термінах типових договорів та тривалості звичних для них проєктів. Відомо, що стандартна робота триває 3 місяці, за які оператори встигають оцифрувати метричні книги як головний предмет цікавості корпорації. Після цього співпраця з обраним архівом завершується, а співробітники починають виконувати подібну роботу, але вже в іншому архіві.

З Україною ж співпраця компанії вийшла на значно новий рівень, оскільки оцифрування триває вже 4-ий рік без значних перерв у роботі. Тому розрахована на більш швидкий результат система під час такої довгострокової роботи й видає назовні усі свої недоліки та нюанси.

У 2022 році Державна архівна служба України повідомила про значну подію: онлайнний ресурс FamilySearch став доступним українською мовою. Це відкрило нові можливості, насамперед, для українських архівістів та дослідників, оскільки тепер вони можуть ще ефективніше працювати з величезною базою даних сімейної історії. Такі впровадження стосувались не лише пошукового порталу FamilySearch, а й їхнього програмного забезпечення для оцифрування архівних документів. Цього чекав кожен український оператор зі сканування, оскільки раніше з доступних мов були лише англійська та російська. Це породжувало внутрішній конфлікт цінностей невикористання мови ворожнечі проти необхідності розуміння написаного, адже іноді технічна англійська не піддавалась перекладу навіть спеціалізованим програмам.

Попри довгоочікуваний переклад ПЗ українською мовою, результати роботи не є досконалими. Наприклад, поширеними є такі помилки (Рис. 3.4):

1. «Completed» замість «завершено».
2. «Open» замість «відкрито».
3. «На HD» замість «на жорсткому диску».

4. Деякі списки йдуть суто англійською мовою.

5. Окремі помилки та інструкції також містять лише англійські тези.

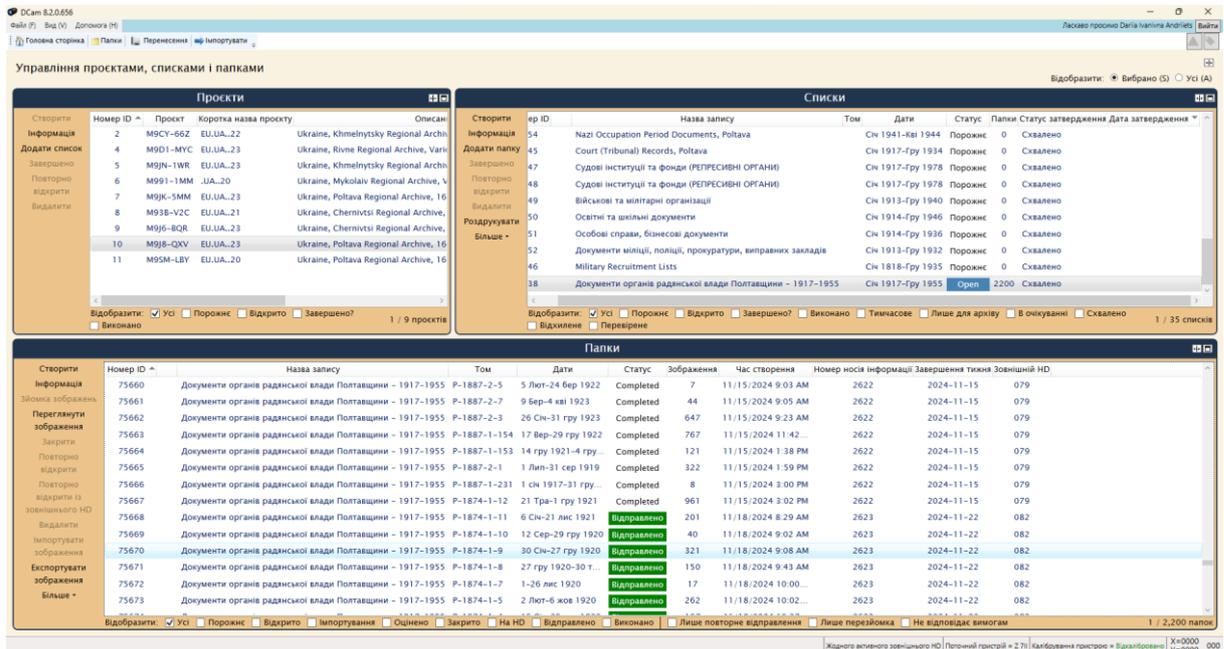


Рис. 3.4 – Основні мовні недоліки української версії ПЗ DCam
Розроблено автором

Прогнозується виправлення всіх цих недоліків перекладу інтерфейсу в наступній версії програми. Якщо цих очікуваних змін не відбудеться найближчим часом, то, на наше переконання, варто додати розроблені й озвучені нами рекомендації для майбутніх вдосконалень.

З'явилась також нова проблема пошкодження файлів під час завантаження готових матеріалів на жорсткі диски.

Під час попереднього огляду створеного за тиждень контенту та перевірки готовності папок до відправлення, тобто перевірки статусу «Закрито», все виглядає відповідним чином. Але саме під час процесу завантаження на жорсткий диск висвічується вікно помилки передачі, що зображено на рисунку 3.5.

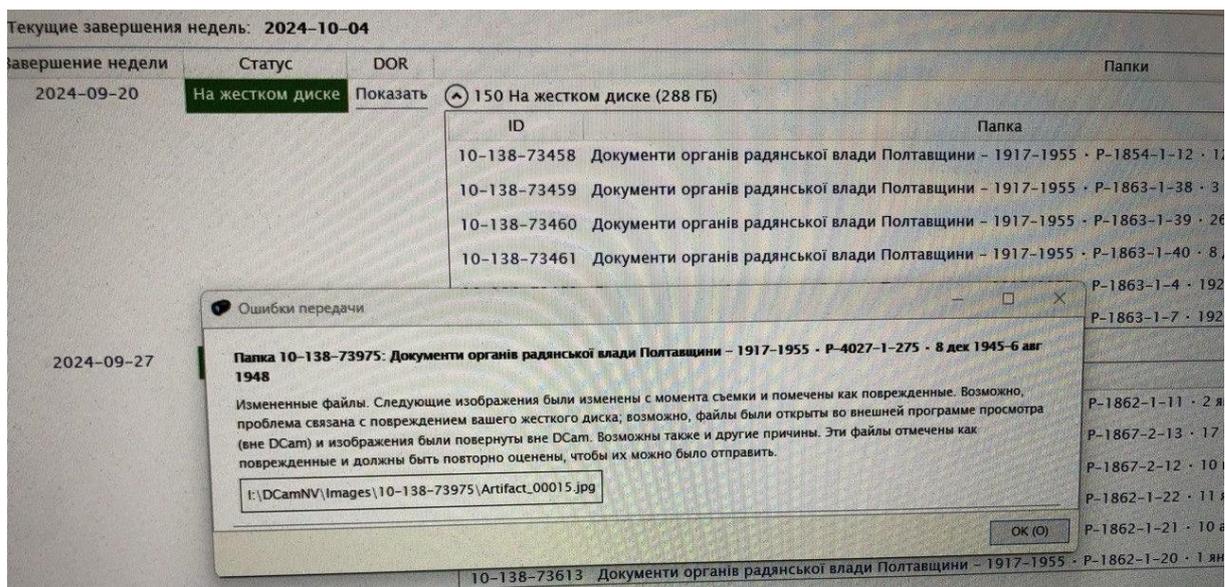


Рис. 3.5 – Помилка передачі папки з DCam на жорсткий диск
Розроблено автором

Пошкоджені файли згодом отримують статус «Повторно відправити – Оцінено» (рис. 3.6) і потребують перегляду на наявність пошкоджених програмою кадрів, а також кадрів, позначених як порожні, оскільки вони найчастіше можуть виявитись пошкодженими.

Це трапляється рідко, але, на превеликий жаль, має місце в роботі оператора.



Рис. 3.6 – Пошкоджений файл
Розроблено автором

Цікавими також є помилки роботи ПЗ, які пов'язані з контролем камери. Часто від тривалої та безперервної роботи обладнання виходить з ладу, а наразі єдиним ефективним вирішенням проблеми є перерва в його роботі. Таке перезавантаження системи є не просто улюбленим методом усунення збоїв у роботі операторів саме завдяки його ефективності, а й рекомендованим розробниками, що й зазначено на рисунку 3.7.

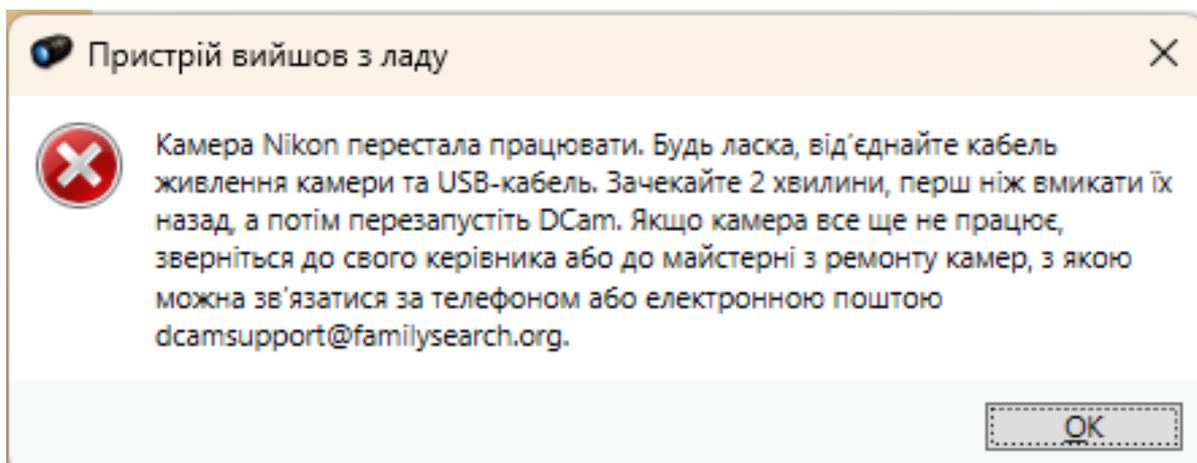


Рис. 3.7 – Помилка роботи камери
Розроблено автором

Це найефективніший та чи не єдиний метод усунення будь-яких помилок, пов'язаних з камерою. Навіть помилки калібрування (рис. 3.8) можна усунути саме завдяки перезавантаженню системи. В іншому разі може виявитись несправність камери чи об'єктиву, що потребуватиме лише заміни частини обладнання.

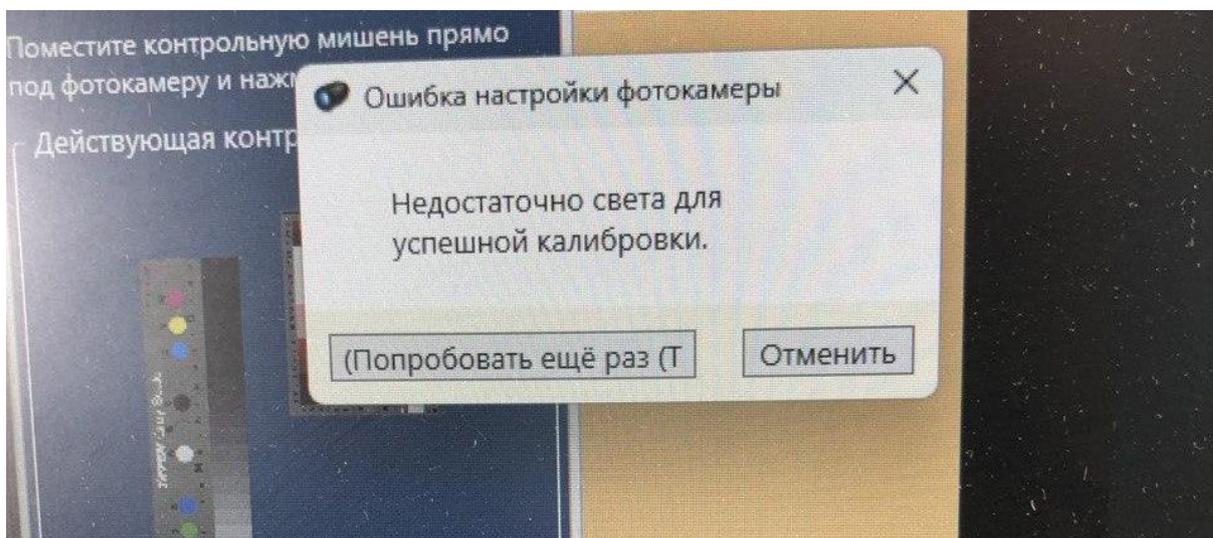


Рис. 3.8 – Помилка роботи камери
Розроблено автором

Ще один випадок несправності обладнання. Перезавантаження системи камера – системний блок усунув помилку під час калібрування, але після того сама камера не могла передавати відповідно до реального освітлення робочу зону, а кадри мали такий вигляд, як репрезентовано на рисунку 3.9.

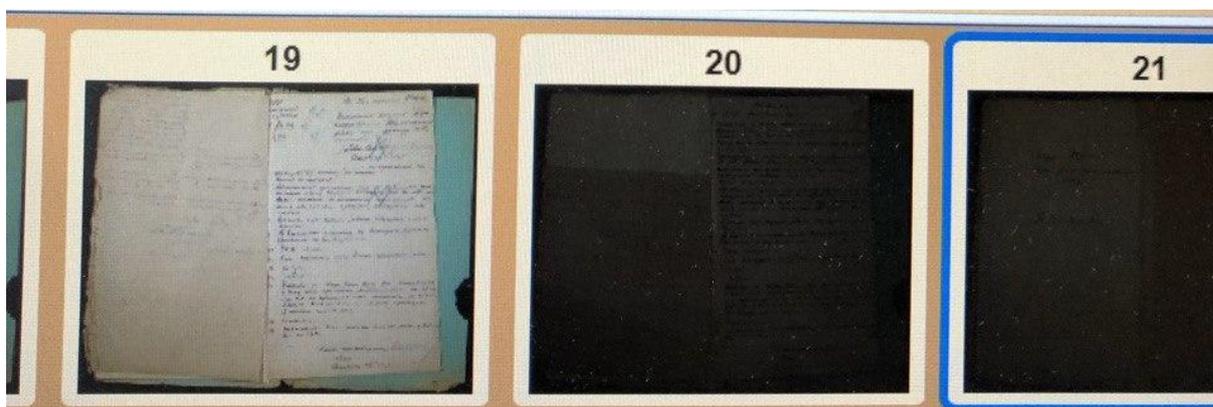


Рис. 3.9 – Наслідки несправності камери
Розроблено автором

Варто зауважити, що в посадові обов'язки оператора зі сканування не входить обслуговування технічно обладнання, тож несправну камеру відправили до спеціаліста з їхнього ремонту.

Є питання також до підбору техніки, яку надає корпорація. Наразі всіх операторів перевели на новіші камери Nikon 7Z II, але з першого ж дня роботи починаються труднощі з фокусуванням (рис. 3.10).



Рис. 3.10 – Недоліки роботи камери Nikon 7Z II
Розроблено автором

Нове обладнання важко сфокусувати на документі з першого разу, доводиться багато разів виконувати автофокусування доки не буде досягнуто прийнятного результату. Також на рисунку 3.10 видно, що іноді

під час сканування фокус збивається без підстав. Це також сповільнює оператора та вимагає підвищеної уважності до кожного кадру.

Недоліком також є відсутність звуку затвору, на який зручно орієнтуватись під час роботи, оскільки чітко зрозуміло, коли кадр вже зроблений та можна підготовлювати до сканування наступний. Система показує, що цей звук має бути увімкненим, однак в жодного з операторів він не лунає. Заміною став кольоровий індикатор зйомки, проте він висвічується на моніторі з певним запізненням, тож це також сповільнює співробітника.

Насамперед, робота з Nikon 7Z II також показала неочікувану статистику. Nikon 7Z II підтримує зйомку зображень зі швидкістю 10 кадрів на секунду [78]. Старі камери Nikon D850 знімають 7 кадрів на секунду [77]. Але, попри це, програму було створено під можливості моделі D850, тому з новішими 7Z II підвантаження знятих зображень в папку відбувається відчутно повільніше, а погодинна статистика оператора підтверджує таку тенденцію.

Отже, розробникам слід звернути увагу на низку недоліків самої програми, її помилок та сумісність з компонентним обладнанням. Саме зв'язок цих елементів надає можливість рівномірної роботи з оцифрування архівних фондів та запобігає появі перепон на цьому шляху.

Особливу увагу також слід приділити питанням збереження цифрових архівів. Програма повинна забезпечувати надійне зберігання даних, захист від втрати інформації та можливість відновлення пошкоджених файлів.

3.3 Рекомендації щодо створення нової версії DСam: впровадження змін на основі практичного досвіду

Глибокий аналіз недоліків поточної версії DСam, зібраних зі зворотного зв'язку його користувачів, технічних звітів та детального дослідження даних використання спонукають нас запропонувати деякі зміни до роботи ПЗ та самого процесу оцифрування. Визначення

пріоритетних напрямків розвитку та розробка стратегії впровадження нових функційних можливостей, спрямованих на підвищення ефективності, зручності та адаптивності програмного забезпечення до потреб операторів, – такою є мета оптимізації DСam.

На нашу думку, варто почати з достатньо глобальних проблем, які ми вже описали вище. До таких важливих моментів, які потребують пильної уваги, можемо віднести й унікальні обмеження, які має кожен список. Створені в них папки мають відповідати цим материнським даним та логічно підходити до визначеної назви (рис. 3.11). Такі списки можуть створювати або редагувати лише адміністратори ресурсу, тоді як додавати папки є безпосереднім обов'язком оператора, як і наповнювати папки **КОНТЕНТОМ**.

The screenshot shows a web application interface for managing document metadata. The interface is in Russian and includes the following sections:

- Заголовок документа** (Document title): A text input field containing "Записи державних установ - 1764-1947".
- Обозначение тома** (Volume designation): An empty text input field.
- Даты** (Dates): A date range selector showing "1764-1947" with "Год" (Year) labels and "Месяц" (Month) label. Below it are two dropdown menus for "Год" (Year) with values "1764" and "1947", and a "Кому" (To) label. A checkbox "Подтверждена правильность этих дат." (These dates are confirmed correct) is present.
- Архивный шифр** (Archival number): A text input field with a placeholder "Если для этого списка нет архивного номера, вместо него введите прочерк (-)".
- Местонахождение** (Location): A search box with "id: 5299" and "Поиск (Дистанционная база данных)". The search result is "Чернівецька, Україна" with a green checkmark and the text "Стандартизированное местоположение: Чернівецька, Україна (Область, Регион, 1991-Настоящее время)". A link "Добавьте более конкретное местоположение" (Add more specific location) is also visible.

Рис. 3.11 – Материнські метадані списку «Записи державних установ 1764 – 1947»
Розроблено автором

Такою, на жаль, є лиш теорія, оскільки на практиці багато українських проєктів мають проблеми з тим, що визначені в списку метадані не відповідають. дійсному історичному періоду створення оцифровуваних документів. Наприклад, у Чернівцях наразі сканують документи, які відповідають списку «Записи державних установ 1764 – 1947», тоді як

материнські метадані місцезрештування допускають зазначення тільки «Чернівці, Україна 1991 – теперішній час» (рис. 3.11). Як ми можемо помітити, місцезрештування зовсім не підходить під назву списку за часовими межами історичних періодів, що є проблемою для майбутнього пошуку інформації в документному масиві.

Неможливість зазначення правильних метаданих (рис. 3.12) вважається помилкою в роботі оператора, оскільки саме він вносить цю інформацію під час оцифрування. Ця помилка вважається критичною як для архіву, так і для FamilySearch. Тому під час перевірки контенту такі матеріали мають бути доопрацьовані аудитором. Такий стан справ створює додаткову роботу через неможливість зазначити правильні дані оператором. Саме через цю невідповідність ми пропонуємо декілька варіантів вирішення питання:

1. Попереднє узгодження всіх материнських метаданих зі зберігачами фондів релевантно до описів, чи з істориками відповідних архівів з огляду на правильність історичних періодів.

2. Зняття суворої заборони на внесення метаданих, що відрізняються від материнських.

3. Надання оператору можливості подання скарги на неправильні материнські метадані списку та виправлення помилок адміністратором ресурсу.

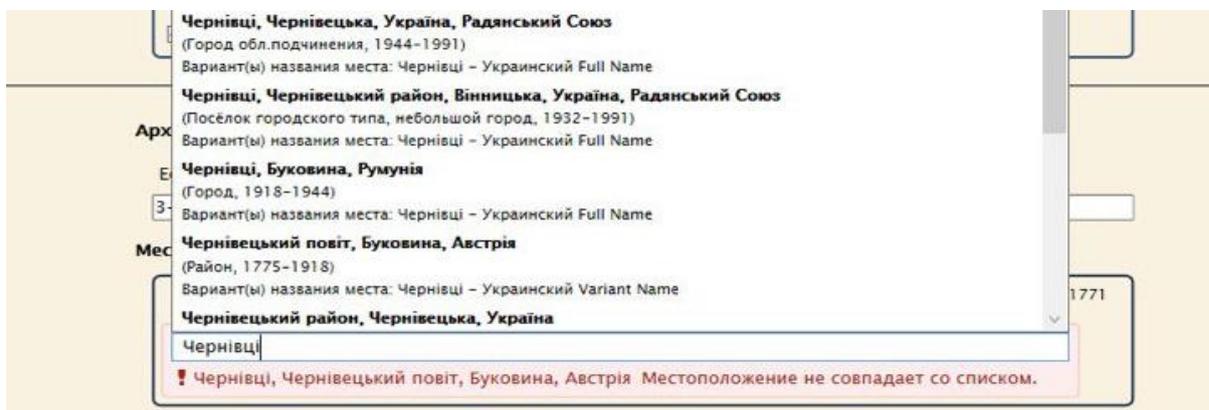


Рис. 3.12 – Помилка внесення метаданих місцезрештування
Розроблено автором

Помилки в роботі з метаданими можуть виникати й з іншими полями шаблону. Так, наприклад, допевне трапитись справа, яку необхідно відсканувати в тому чи іншому списку, але зазначені на ній роки створення чи завершення не збігаються з роками списку. У таких випадках ми практикуємо внесення материнських років списку, що не є правильним в самому задумі.

Під час роботи з типами записів також можуть виникнути труднощі. Так, на рисунку 3.13 зображено помилку роботи програми, коли головний тип видаляється програмою. У цьому разі головним типом має бути встановлений саме материнський варіант, оскільки інакше система не пропустить свого користувача на етап створення контенту.

Можливий ще один варіант збою. Оператором було зазначено додаткові типи записів для точнішої передачі документної інформації. У разі виникнення помилки, може бути забраковано системою також і материнські типи записів. У кращому разі необхідно перевірити послідовність та кількість материнських типів, зазначених в проблемній папці. Якщо відповідні зміни не допомогли, то необхідно перезавантажити систему. Траплялись випадки, коли не допомагало й оновлення, тоді оператор видаляв створену папку та створював усе заново за шаблоном раніше оцифрованих матеріалів, які пропускає система.

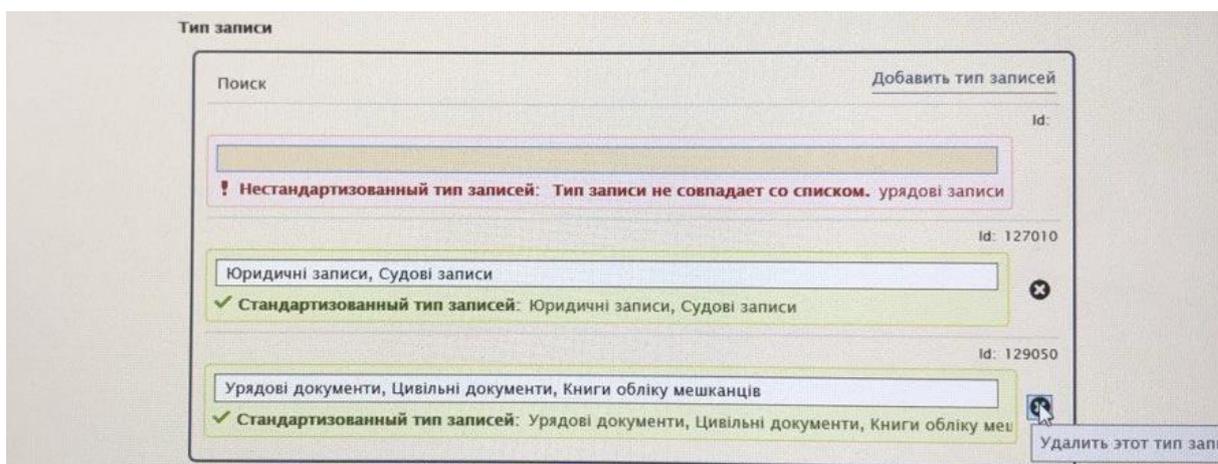


Рис. 3.13 – Помилка внесення метаданих типу записів
Розроблено автором

Зауважимо, що з кожною новою версією DСam складність внесення додаткових типів записів стає дедалі вищою. Тому без великої необхідності оператори обмежуються лиш материнськими метаданими списку, оскільки наразі кожен додатковий тип, крім заздалегідь визначених, проковує виникнення помилки в системі.

Загальні системні помилки також є цікавими для вивчення та абсолютно проігнорованими розробниками. Їхня особливість полягає в тому, що поява такого випадуючого вікна (див. рис. 3.14) не впливає на подальшу роботу. Не має значення і реакція користувача ПЗ закрити вікно чи погодитись з ним, бо там вміщена інформація для ознайомлення. Куратори проєктів дали вказівку своїм операторам погоджуватись із вміщеною інформацією на таких вікнах. Але будь-яка помилка та інформування від програми несе в собі певне смислове навантаження, тож саме розробникам варто розібратись з подібного роду ситуаціями.

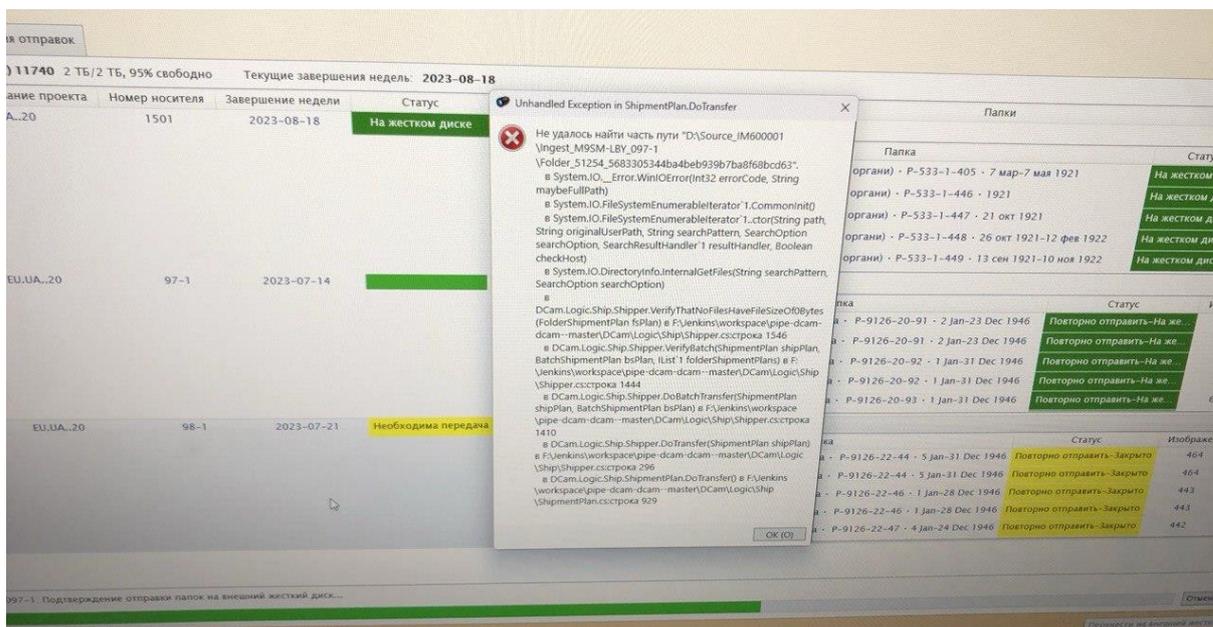


Рис. 3.14 – Помилка втрати шляху
Розроблено автором

Одним із різновидів цієї проблеми є унікальний випадок, який поки що траплявся лише на одному системному блоці. Усі взаємодії з ним є аналогічними до інших, цікаво лиш те, що ця помилка виникає на постійній основі з інтервалом в 5 хвилин. У разі перерви в роботі оператора одна

помилка нашаровується на іншу до безкінечності, тому ця проблема отримала неофіційну назву «Помилка нашарування» (див. рис. 3.15).

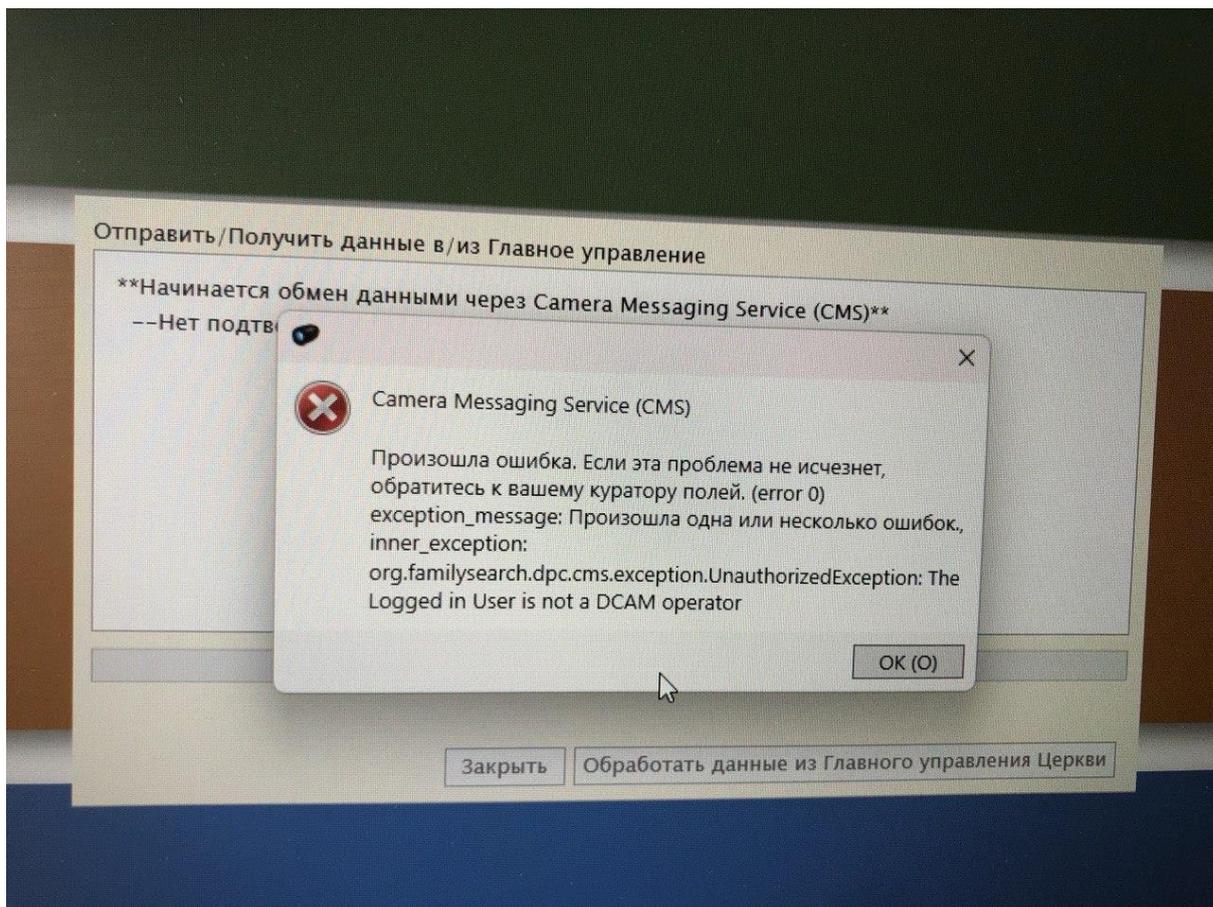


Рис. 3.15 – Помилка нашарування
Розроблено автором

Складнощі також виникають під час дотримання правила орієнтації тексту: якщо основна інформація на сторінці перевернута на 90° , 180° , 270° , а на іншій сторінці розвороту текст має інший кут нахилу (зокрема й, класичний 0°), то необхідно створити два однакових кадри та перевернути необхідні сторінки у правильному напрямку. У разі неправильної, але однакової, орієнтації обох сторінок розвороту слід створити один кадр та перевернути його.

Типовою помилкою в роботі DСam є збій перевертання діапазону сторінок, коли для частини виділених сторінок документа виконується команда, а для деяких із них – ігнорується програмою (див. рис. 3.16). На етапі оцінювання оператор має помітити цю похибку системи та виправити

її вручну. Але розробникам варто звернути увагу на такий баг та покращити роботу функції «Перевертання».

Загалом, під час перегляду оцифрованих фондів є можливість перевернути кадр самому користувачу та дослідити документ під зручним йому кутом. Тож ми вважаємо недоцільним створення додаткових кадрів та їхнє перевертання оператором, оскільки від цієї роботи менше користі, ніж витраченого робочого часу.

У разі розбіжностей у поглядах з розробниками, адміністраторами ресурсу і визначенні обов'язковості дублювання розворотів та їхнього перевертання, ми пропонуємо внесення нової функції до програми. Мова йде про кнопку повороту кадру на 180°, яка присутня в будь-якому фоторедакторі та не є надто складною для внесення нових рядків коду до ПЗ.



Рис. 3.16 – Збій роботи функції «Перевертання» на діапазоні кадрів

Розроблено автором

Якщо аналізувати помилки роботи програми, то не можна оминати й позначення пошкоджених кадрів. Беззаперечно, що вони могли стати пошкодженими через виникнення якоїсь сторонньої помилки, несправною роботою камери чи збоєм ПЗ. Такі кадри необхідно переробити, оскільки самі сторінки документа завжди містять цінну інформацію, яка має бути оцифрована та оприєвлена. Але цікавим є те, що позначення пошкоджень бувають декількох типів, хоча для жодного з них немає опису в посібнику користувача програми. Вони відрізняються зовнішніми ознаками, часом

появи та можливістю подальшого створення якісного контенту до переробки пошкоджених кадрів.

Розробниками варто звернути увагу й на самі причини пошкодження кадрів, адже інколи видимих підстав для дефектів чи збоїв у роботі складників установки не було. Такі кадри потім на етапі оцінювання система помічає хрестом (див. рис. 3.17).



Рис. 3.17 – Позначення пошкоджених кадрів. Хрест

Розроблено автором

У разі збою DСam чи камери, останні відскановані зображення пошкоджуються з поміткою у вигляді восьмикутника (див. рис. 3.18). Система не дозволяє подальшого сканування чи перезйомки таких зображень, їх необхідно просто видаляти та створювати заново.



Рис. 3.18 – Позначення пошкоджених кадрів. Восьмикутник

Розроблено автором

Проте є ще один вид пошкодження контенту. Його не можна помітити в зменшеному вигляді копій і він не помічається жодними червоними елементами привернення уваги оператора (див. рис. 3.19). Але після оцінювання нереально закрити таку папку ні автоматично, ні вручну. Для виявлення та усунення цього недоліку необхідно перевіряти вручну всі кадри, тільки так і можливо виявити причину збою.

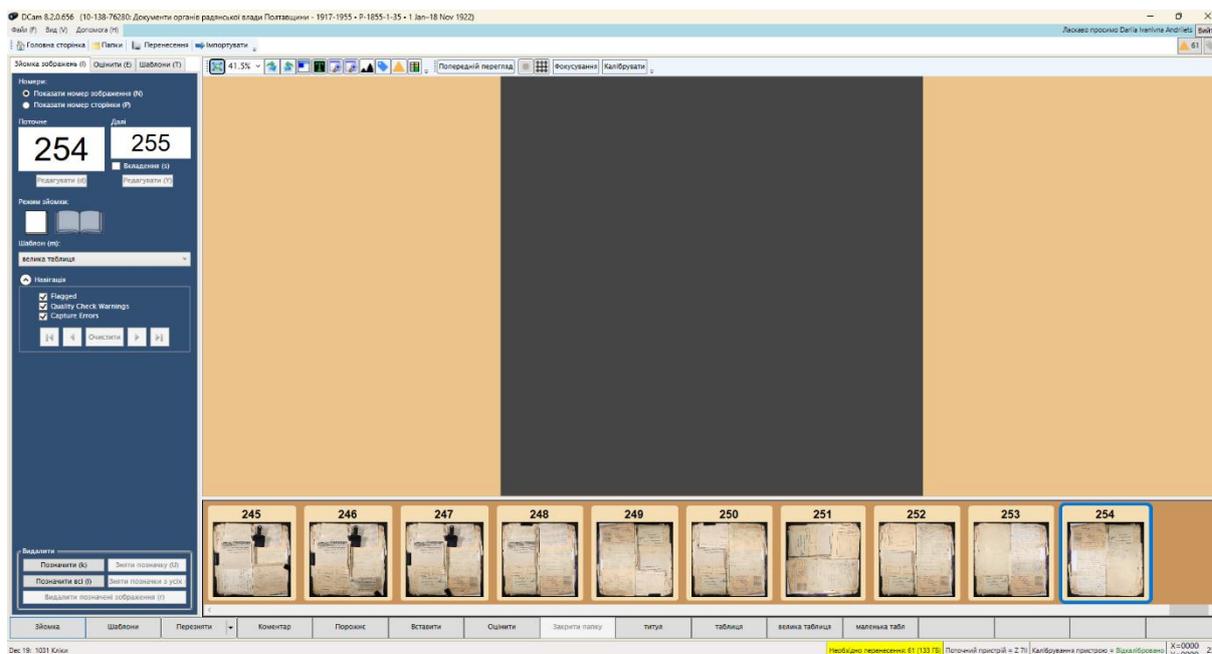


Рис. 3.19 – Позначення пошкоджених кадрів. Сірий квадрат
Розроблено автором

Піднята тема неможливості автозакриття папки стосується ще й нового недоліку останньої версії DScan. Наступного дня після завантаження контенту на зовнішній жорсткий диск правильний алгоритм дій не виконує своїх прописаних функцій. Тож необхідно кожен папку закривати вручну. Цей баг, безумовно, значно ускладнює роботу оператора, оскільки ручний режим може також не спрацювати. Але впродовж тижня все ж настає час, коли всі функції програми починають працювати відповідно до закладених можливостей.

Можемо також порадишити покращити якість роботи наявної функції автокадрування, оскільки наразі вона дублює попередню роботу камери та не видає іншого результату. Бажаним було б введення підфункції ручного кадрування та повторного автокадрування, це б уможливило редагування зображення, і зникла б необхідність переробляти їх. Так, наприклад, камера іноді захоплює відблиск на підручних матеріалах та створює ширшу рамку, тоді як описаний нами інструмент допоміг би виправити готові кадри.

Сама ж корпорація FamilySearch кожного року звітує про свої досягнення та оновлення, а також анонсує заплановані зміни та покращення

роботи ресурсів на наступний рік. На жаль, «усі новації стосуються лише відкриття нових проєктів чи оновлення пошукової системи порталу, тоді як основа їхньої роботи та зростання обсягів оцифрування не зазнавала суттєвих змін на краще ще з 7-ої версії» [72, 82].

Кожна нова версія після того приносила лиш несуттєві чи непотрібні зміни та купу нових обмежень в роботі та багів.

Ми ж пропонуємо розробникам на честь ювілейної 10-ої версії DСam усунути зазначені нами недоліки роботи програми та трохи осучаснити інтерфейс. Мова йде не про кольори чи дизайн, а про зручний та адаптивний інтерфейс, зокрема:

- надати можливість корегувати масштаб тексту для зручності читання;
- усунути дубляж функцій та панелей в інтерфейсі;
- можливість налаштування інтерфейсу під потреби оператора та закріплення додаткових вкладок і випадаючих вікон на вільному місці в бічній панелі;
- закріплення роздільної здатності зображення під номером кадру на бічній панелі;
- необхідно передбачити функціонал для колективної роботи та обміну даними між різними користувачами задля уникнення проблем із дубляжем оцифрування різними операторами.

Отже, важливо забезпечити зручність роботи користувачів власного програмного забезпечення та розвивати й покращувати DСam.

На нашу думку, програма повинна мати гнучку систему налаштувань, що дозволить адаптувати її до різних типів документів та вимог операторів.

Крім того, варто аналізувати скарги та усувати вже наявні проблеми перед випуском нової версії програми.

Загалом, DСam є потужним інструментом сканування та перспективною платформою, що потребує своєї вчасної актуалізації відповідно до ринкових пропозицій.

ВИСНОВКИ

Важливість ролі представлення документної інформації зумовлює необхідність у розробленні ґрунтовних наукових розвідок щодо міжнародної співпраці в цій царині, а також досліджень всіх етапів процесу діджиталізації архівної спадщини. Саме тому особливу роль відіграє вивчення практичного досвіду спільної реалізації проєктів з оцифрування архівів.

Під час написання кваліфікаційної роботи нами було проведено низку досліджень, пов'язаних із оцифруванням архівної спадщини. Після систематизації отриманих результатів було зроблено такі висновки.

1. У процесі роботи було визначено дефініції термінів «оцифрування» та «діджиталізація», а також «електронний архів», що є важливим для розуміння сутності досліджуваних процесів.

Ми також розглянули основні загрози втрати інформації через пошкодження фізичного носія. Ними є старіння матеріалів, з яких виготовлений носій, механічне пошкодження оригіналів, втрата інформації через катаклізми чи війну. Нами встановлено, що одним із можливих варіантів захисту цінної інформації від втрати, є оцифрування архівних документів.

Під час аналізу особливостей основних етапів оцифрування документів ми виокремили три складники загального процесу: обов'язкового, коли сканують фізичні примірники; метадатування, під час чого присвоюють відповідні відомості щодо оригіналу; оптичного розпізнавання символів для можливості подальшого повнотекстового пошуку.

2. Розглянули іноземний досвід оцифрування як ключовий інструмент для забезпечення збереження культурної спадщини для майбутніх поколінь. Міжнародні ініціативи в цій галузі доводять ефективність такої кооперації.

Під час нашого дослідження було розглянуто низку всесвітніх ініціатив у цій галузі, наприклад:

- оцифрування стародруків з музею Плантена–Моретуса компанією Google, до яких надано доступ через Google Books;

- застосунок Time Machine та електронний масив Big Data of the Past, що об'єднують дані інших ініціатив на єдиній пошуковій платформі;

- вебпортал «Szukajwarchiwach.pl», де розміщено польські оцифровані матеріали;

- вебплатформа «Жертви антипольського терору в Радянському Союзі 1934–1938» від польських дослідників, яка стала поштовхом до оприлюднення злочинів радянської системи проти своїх народів.

Найважливішою частиною дослідження міжнародної роботи з оцифрування архівної спадщини є співпраця іноземних архівістів з Україною:

- Europeana активно долучилася до збереження української культури, зробивши доступними онлайн унікальні артефакти з України;

- проєкт SUCHO мобілізував понад 1300 фахівців для цифрової архівації унікальних об'єктів, які опинилися під загрозою;

- завдяки Чеській Республіці та програмі «Ковчег» передано спеціальне обладнання для оцифрування достраждалих під час війни регіонів. Запланована програма «Ковчег-2» націлена вже на популяризацію української архівної спадщини на світовій арені;

- проєкт США та «ALIPH», націлений на підвищення кваліфікації українських архівних співробітників щодо роботи з пошкодженими архівними матеріалами;

- Острозький музей залучений до оцифрування завдяки американському музею, а також Університету Абатства Сент-Джон у США;

- генеалогічна корпорація FamilySearch має широку мережу проєктів з діджиталізації обласних архівів нашої держави.

Нами було доведено, що об'єднання зусиль з іншими країнами дозволить нам зберегти нашу культурну спадщину і зробити її доступною для дослідників з усього світу.

3. Проаналізували переваги співпраці Державного архіву Полтавської області з міжнародною організацією FamilySearch: доступ до сучасних технологій; гарантія якості цифрових копій; забезпечення публічного доступу до архівних документів; додаткове джерело фінансування. Для дослідників та користувачів така ініціатива надає низку нових можливостей пошуку документної інформації: пошук за багатьма критеріями, доступ до цифрових копій онлайн, а також можливість їхнього завантаження, мультимовний інтерфейс.

Крім переваг, подібне оцифрування має і свої певні недоліки, а саме:

- 1) потреба в значному фінансуванні;
- 2) довготривалість;
- 3) необхідність оновлення технічного забезпечення;
- 4) ризик неякісного сканування через незадовільний стан оригіналу.

Потребує такого оновлення й пошуковий портал FamilySearch. Його недоліками є:

- 1) складність пошуку інформації через велику кількість критеріїв;
- 2) неможливість доступу до деяких документів у зв'язку з секретністю;
- 3) помилки в зазначених метаданих;
- 4) недостатній рівень оцифрування світової спадщини;
- 5) складність розпізнавання рукописних текстів документів.

Встановлено, що ця співпраця стала можливою завдяки чітко прописаним обов'язкам компанії, її підрядників та ДАПО. Попри складність трьохсторонніх відносин, ми можемо стверджувати, що їхня взаємодія дає значні результати та вже декілька років підряд визнана найбільш плідною в країні серед аналогічних проєктів.

Задля успішного розвитку системи оцифрування фахівці повинні, якнайперше, зосередитися на усуненні вже виявлених недоліків.

4. Дослідження технічних параметрів діджиталізації архівної спадщини підтвердило важливість якості обладнання для здійснення цього процесу. Тож було обґрунтовано специфіку вибору моделі технічного забезпечення:

- 1) вид фізичного носія інформації визначає тип обладнання;
- 2) для сканування рукописних та друкованих текстових документів варто враховувати їхній розмір та наявність дрібних деталей (чим дрібніший оригінал, тим більший запас роздільної здатності має бути наявний в обраній моделі);
- 3) варто зважати також на швидкість сканування, а також звертати увагу на наявність подальшого оптичного розпізнавання символів.

Після детального вивчення основних форматів збереження відсканованого матеріалу, таких як PDF, JPEG та TIFF, ми наочно побачили різницю якості збережених зображень. Проаналізувавши усі їхні особливості, переваги та недоліки, можна впевнено стверджувати, що попри потребу у великому об'ємі пам'яті, TIFF є найбільш придатним для оцифрування, оскільки саме він гарантує найкращу якість.

Запропоновано використовувати можливості штучного інтелекту для автоматизації багатьох етапів оцифрування, а також їхнього подальшого пошуку та дослідження.

5. Встановлено, що DСam, як розробка FamilySearch, є спеціалізованим програмним забезпеченням для діджиталізації архівних документів. Досліджено принцип його роботи, а також етапи створення цифрових копій: підготовка обладнання, оформлення метаданих, сканування, зазначення особливостей оцифрованих документів або аркушів, завантаження на жорсткий диск. Подальший шлях інформації лежить через аудиторську перевірку якості зображень та відповідності метаданих. Після цього документи розміщують у пошуковому порталі.

Задля покращення роботи ми пропонуємо внесення низки змін до наявної системи, а саме:

- усунення збоїв створення та редагування списків;
- покращення української версії ПЗ відповідно до виявлених нюансів;
- налагодження системи завантаження та ліквідація збоїв процесу;
- виправлення помилок коду, які спричиняють пошкодження файлів;
- оновлення системи контролю камери задля унеможливлення помилок в її роботі;
- переналаштування DСam на управління наявним сучаснішим обладнанням та пристосування можливостей програми під характеристики камери Nikon 7Z II;
- введення обов'язкової консультації з істориком архіву перед укладанням метаданих списків;
- усунення «проблеми втрати шляху» та, так званої, «помилки нашарування»;
- перегляд обов'язків оператора щодо зміни орієнтації документа;
- виправлення помилки програми, яка стосується автоматичного закриття папок;
- переосмислення автокадрування;
- оновлення інтерфейсу програми та розробка можливостей його персоналізації під конкретні потреби.

Отже, спільні міжнародні проєкти з оцифрування документних матеріалів є основою розвитку сучасної архівної галузі. Така діджиталізація української спадщини гарантує донесення правдивої історії до свідомості світової спільноти та майбутніх поколінь.

Подальше дослідження цієї тематики ми вбачаємо в аналізі ефективності роботи різних технологій оцифрування, а також виявленні користувацьких потреб та розробці нових інструментів для роботи з цифровими копіями.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ ТА ДЖЕРЕЛ

1. Постанова Верховної Ради України 21.03.2023 № 3003-ІХ проєкт Закону України про внесення змін до Закону України «Про культуру» щодо впровадження та ведення електронних реєстрів обліку та управління культурною спадщиною та культурними цінностями України (реєстр. № 8024) прийнято за основу. URL : <https://zakon.rada.gov.ua/laws/show/3003-IX#Text> (дата звернення 30.08.2023).

2. Про авторське права і суміжні права : Закон України від 01.12.2022. № 2811-ІХ. URL : <https://zakon.rada.gov.ua/laws/show/2811-20#Text.2>. (дата звернення: 21.10.2024).

3. Про доступ до архівів репресивних органів комуністичного тоталітарного режиму 1917-1991 років : Закон України від 09.04.2015. № 316-VIII : станом на 1 січ. 2022 р. URL: <https://zakon.rada.gov.ua/laws/show/316-19#Text> (дата звернення: 01.10.2024).

4. Про електронну ідентифікацію та електронні довірчі послуги : Закон України від 05.10.2017. № 2155-VIII. URL : <https://zakon.rada.gov.ua/laws/card/2155-19> (дата звернення: 21.10.2024).

5. Про затвердження Програми оцифрування архівних інформаційних ресурсів на 2022–2025 роки: Наказ від 29.12.2021 №165. URL : https://srpi.archives.gov.ua/tmp/1445_2.pdf?fbclid=IwAR36HtbJf4hpa8fYuNsWxz5s6syHG5i7GeXXLvt5sG6ahZr_zX7AQMZplA (дата звернення 30.08.2024).

6. Про захист персональних даних : Закон України від 01.06.2010. № 2297-VI. URL : <https://zakon.rada.gov.ua/laws/card/2297-17> (дата звернення: 21.10.2024).

7. Про інформацію : Закон України від 02.10.1992. № 2657-XII. URL : <https://zakon.rada.gov.ua/laws/card/2657-12> (дата звернення: 21.10.2024).

8. Про Національний архівний фонд та архівні установи : Закон України від 24.12.1993. № 3814-XII. URL : <https://zakon.rada.gov.ua/laws/card/3814-12> (дата звернення: 21.10.2024).

9. Про ратифікацію Конвенції про захист культурних цінностей у разі збройного конфлікту 1954 року. *Офіційний вебпортал парламенту України*. URL : https://zakon.rada.gov.ua/laws/show/995_001-99#Text (дата звернення: 23.10.2024).

10. Про ратифікацію Конвенції про заходи, спрямовані на заборону і запобігання незаконному ввезенню, вивезенню та передачі права власності на культурні цінності. *Офіційний вебпортал парламенту України*. URL : <https://zakon.rada.gov.ua/laws/card/5396-11> (дата звернення: 23.10.2024).

11. Про ратифікацію Конвенції про охорону всесвітньої культурної і природної спадщини. *Офіційний вебпортал парламенту України*. URL : https://zakon.rada.gov.ua/laws/show/995_089#Text (дата звернення: 23.10.2024).

12. Про ратифікацію Конвенції про охорону нематеріальної культурної спадщини. *Офіційний вебпортал парламенту України*. URL : https://zakon.rada.gov.ua/laws/show/995_d69#Text (дата звернення: 23.10.2024).

13. Про ратифікацію Конвенції про охорону та заохочення розмаїття форм культурного самовираження. *Офіційний вебпортал парламенту України*. URL : https://zakon.rada.gov.ua/laws/show/952_008 (дата звернення: 23.10.2024).

14. Програма оцифрування архівних інформаційних ресурсів на 2022-2025 роки. *Офіційний сайт Державної архівної служби України*. URL : https://srpi.archives.gov.ua/tmp/1445_2.pdf (дата звернення: 27.10.2024).

15. Про схвалення Стратегії захисту документної спадщини як запоруки збереження національної ідентичності та державності на період до 2027 року та затвердження операційного плану заходів з її реалізації у 2025-2027 роках. *Офіційний вебпортал парламенту України*. URL :

<https://zakon.rada.gov.ua/laws/show/1349-2024-%D1%80#n17> (дата звернення: 25.12.2024).

16. Цивільний кодекс України від 13.04.2012. № 4651-VI. URL : <https://zakon.rada.gov.ua/laws/show/435-15Text>. (дата звернення: 21.10.2024).

17. Анатолій Хромов: про оцифрування, збереження й відкритість архівів. *Ukrainer*. URL : <https://www.ukrainer.net/anatoliy-khromov/> (дата звернення: 21.08.2024).

18. Відтепер FamilySearch доступний українською мовою. *Укрінформ*. URL : <https://www.ukrinform.ua/rubric-regions/3450183-vidteper-familysearch-dostupnij-ukrainskou-movou.html> (дата звернення: 09.12.2024).

19. Гранітна гора. *Вікіпедія*. URL: https://uk.wikipedia.org/wiki/Гранітна_гора (дата звернення: 04.11.2024).

20. Грінберг Л. Ф. Оцифрування документної спадщини архівів і бібліотек під час війни. *Інформація та соціум: збірник матеріалів VIII Міжнародної науково-практичної конференції* (м. Вінниця, 02 червня 2023 р.). Вінниця: ДонНУ імені Василя Стуса. 2023, С. 78–80.

21. Державна архівна служба України. *Facebook*. URL : <https://www.facebook.com/archives.ua> (дата звернення: 12.11.2024).

22. Державна архівна служба України. *Державна архівна служба України. Офіційний вебпортал органу виконавчої влади*. URL: <https://archives.gov.ua/ua/2023/05/30/новини-оцифрування-документів-держ-30/> (дата звернення: 28.08.2024).

23. Державний архів Київської області. *Facebook*. URL : <https://www.facebook.com/dako.gov.ua> (дата звернення: 12.11.2024).

24. Державний архів Полтавської області. *Офіційний сайт Державного архіву Полтавської області*. URL: <https://poltava.archives.gov.ua/> (дата звернення: 04.11.2022).

25. Долик М.М. Інформаційна система оцифрування документів для збереження історико-культурної спадщини України. Збірник тез доповідей

VIII Міжнар. наук.-тех. конф. молодих учених та студентів «Актуальні задачі сучасних технологій». Тернопіль: ТНТУ, 2019. Том 2. С. 32 – 33.

26. Долик М.М. Інформаційна система оцифрування та зберігання архівних документів. Тернопіль : ТНТУ, 2019. 141 с.

27. Досі пахнуть згарищем: як триває реставрація архівних документів, постраждалих під час пожежі у Кам'янці-Подільському. *Новини Хмельницького «Є»*. URL : https://ye.ua/syspilstvo/66820_Dosi_pahnut_zgarischem_yak_trivaye_restavraciya_arhivnih_dokumentiv_postrazhdalih_pid_chas_pozhezhi_u_Kam_yanci_Podilskomu.html (дата звернення: 27.11.2024).

28. Електронний архів. *Вікіпедія*. URL : https://uk.wikipedia.org/wiki/Електронний_архів (дата звернення: 12.08.2024).

29. Кам'янецький архів: 10 років коми. попереду ще 100. *Газета «Подільський»*. URL : <https://podolyanin.com.ua/history/7745/> (дата звернення: 27.11.2024).

30. Кириленко О.В. Сучасне спеціалізоване обладнання для оцифрування архівних документів. *Інформаційні технології і системи в документознавчій сфері* : зб. матеріалів ІХ Всеукраїнської наукової студентської конференції. Вінниця : ДонНУ імені Василя Стуса, 2024. С. 99 – 101.

31. Копіювання документів у архівних установах України: методичні рекомендації. *Держ. архів. служба України, УНДІАСД*; уклад.: Л.В. Дідух. Київ, 2016. 68 с.

32. Курило В., Комова М. Можливості застосування систем штучного інтелекту та машинного навчання як інструментів для оцифрування документів. *Інформація, комунікація, суспільство 2024* : матеріали 13-ї Міжнар. наук. конф. Львів, 2024. С. 133–136.

33. Лист Папи Бенедикта XIV. Рим. 1744 рік. *Державний архів Волинської області*. Ф. 312. Оп. 1.Спр. 185. Арк. 1.

34. Литвинська С. Оцифрування документів як інноваційний спосіб збереження історико-культурної спадщини. *Історико-культурна спадщина*:

збереження, доступ, використання : матер. Міжнар. наук.-практ. конф., м. Київ, 7–9 квітня 2015 р. Київ : Талком, 2015. С. 276–279.

35. Людський дотик їх нищать: як у Хмельницькому оцифровують архівні документи. *Ye.ua*. URL : https://ye.ua/syspilstvo/66518_Lyidskiy_dotik_yih_nischit__yak_u_Hmelnickomu_ocifrovuyit_arhivni_dokumenty.html (дата звернення: 12.11.2024).

36. Менська О.А. Оцифрування об'єктів культурної спадщини як спосіб формування нового цифрового культурного простору України. *Сучасні аспекти модернізації науки: стан, проблеми, тенденції розвитку: матеріали XXXV Міжнародної науково-практичної конференції / за ред. І.В. Жукової, Є.О. Романенка*. м. Стамбул (Туреччина): ГО «ВАДНД», 07 серпня 2023 р. С. 59-68.

37. Міжнародна діяльність. *Державний архів Хмельницької області*. URL : <https://dahmo.gov.ua/view.php?cat=68> (дата звернення: 12.11.2024).

38. Мусіна Г. Курило Н.О. Оцифрування як засіб зберігання архівних документів в умовах війни. *Актуальні виклики розвитку інформаційної діяльності в глобальному вимірі* : матеріали Всеукраїнської науково-практичної конференції. Полтава: Вид-во ДЗ «ЛНУ імені Тараса Шевченка», 2024. С. 146-149.

39. Обладнання та матеріали. Архівні інформаційні системи. URL: <https://arinsy.com/equipment-and-materials/> (дата звернення: 07.12.2024)

40. Онищенко О.С. Цифровізація – стратегічний шлях розвитку бібліотечної сфери. *Бібліотечний вісник*, 2021. № 4. С. 3–9. URL : http://nbuv.gov.ua/UJRN/bv_2021_4_3 (дата звернення 01.07.2024).

41. Описи та справи. *Державний архів Київської області*. URL : <https://opisi.dako.gov.ua/site/cfk> (дата звернення: 12.11.2024).

42. Оцифрування. Вікіпедія. URL : <https://uk.wikipedia.org/wiki> (дата звернення: 12.08.2024).

43. Оцифрування складних документів. *Архівна спадщина України*. URL: <https://www.arhives.in.ua/rodinni-arxivi/ocifruvannya-skladnix-dokumentiv-ta-kolekcij-zobrazhen/> (дата звернення: 12.08.2024).

44. Про роботу Укрдержархіву, архівних установ і спеціальних установ страхового фонду документації у 2023 році та пріоритети на 2024 рік. *Державна архівна служба України*. URL : https://www.kmu.gov.ua/storage/app/sites/1/17-civik-2018/zvit2023/zvit_archiv_2023.pdf (дата звернення: 12.11.2024).

45. Програма ЮНЕСКО «Пам'ять світу». *Центральний державний архів громадських об'єднань та україніки*. URL : <https://tsdahou.archives.gov.ua/programa-yunesko-pamyat-svitu/> (дата звернення: 23.10.2024).

46. Пророцтво (1490 р.) іспанського ченця Севільського монастиря Франческо, знайдене вченим Гербертом Шротерхом і опубліковане в 1834 р. *Державний архів Рівненської області*. Ф. Р–30. Оп. 1. Спр. 30. Арк. 1.

47. Публічний звіт директора Державного архіву Полтавської області за 2023 рік. *Офіційний сайт Державного архіву Полтавської області*. URL: <https://poltava.archives.gov.ua/diialnist-arkhivu/plany-zvity-informatsiini-rovidomlennia> (дата звернення: 12.11.2024).

48. Радченко Д.І. Міжнародне нормативно-правове регулювання процесу оцифрування архівної спадщини. *Документно-інформаційні комунікації в умовах глобалізації: стан, проблеми та перспективи* : матеріали ІХ Міжнар. наук.-практ. конф., 21 листоп. 2024 р. / редкол. : І.Г. Передерій [та ін.]. Полтава, 2024. С. 225 – 231.

49. Радченко Н. Досвід співпраці архівних установ та суспільства у процесах формування історичної пам'яті. *Актуальні виклики розвитку інформаційної діяльності в глобальному вимірі* : матеріали Всеукраїнської науково-практичної конференції. Полтава : Вид-во ДЗ «ЛНУ імені Тараса Шевченка», 2024. С. 149-159.

50. Ревко Н. Актуальні архіви: як оцифровують історичні документи. *Ms.detector.media*. URL : <https://ms.detector.media/onlain-media/post/16203/2016-03-07-aktualni-arkhivy-yak-otsyfrovuyut-istorychni-dokumenty/> (дата звернення: 21.08.2024).

51. Романовський Р. В. Оцифрування архівних документів: досвід державних архівів Польщі. *Архіви України* : наук.-практичний журнал. Київ : Державна архівна служба України, № 3 (320), 2019. С. 113–127.

52. Сіденко Д.А. Оптимізація процесу оцифровки словників 60-90-х років ХХ століття. Харків : ХНУРЕ, 2021. 67 с.

53. Слово року 2019. *Мислово*. URL: <http://myslovo.com/wp-content/uploads/2021/11/Слово-року-2019.pdf> (дата звернення: 19.08.2024).

54. США запускають проєкт із захисту українських державних архівів. *Укрінформ*. URL : <https://www.ukrinform.ua/rubric-culture/3831195-ssa-zapuskaut-proekt-iz-zahistu-ukrainskih-derzavnih-arhiviv.html> (дата звернення: 12.11.2024).

55. У Острозькому музеї книги і друкарства оцифрують старовинні рукописи за сприяння США. *Читомо*. URL : <https://chytomo.com/ustrozkomu-muzei-knyhy-i-drukarstva-otsyfruiut-starovynni-rukopysy-zaspryiannia-ssha/> (дата звернення: 04.11.2024).

56. У Чехії створили програму «Ковчег» – для збереження українських стародруків. *Радіо Свобода*. URL : <https://www.radiosvoboda.org/a/news-chehija-zberezhennya-ukrajinskyh-starodrukiv/32809733.html> (дата звернення: 12.11.2024).

57. Україна належить до світових лідерів за темпами оцифрування архівних документів, – Мін'юст. *Урядовий портал*. URL : <https://www.kmu.gov.ua/news/ukraina-nalezhyt-do-svitovykh-lideriv-za-tempamy-otsyfruvannia-arkhivnykh-dokumentiv-miniust> (дата звернення: 23.11.2024).

58. Цирульнев Ю.Б. Практичні аспекти проєктів з оцифрування документів в державних установах України. *Мультимедійні технології в*

освіті та інших сферах діяльності: науково-практична конференція з міжнародною участю. 2 листопада 2021 р. Київ : НАУ, 2022. С.121-124.

59. Шевченко М.О. Краєзнавчі документи: пріоритетний об'єкт оцифрування. *Вісник Книжкової палати*. 2020. № 1. С. 34–38.

60. Шевченко М.О. Оцифрування фондів бібліотек України. Харків : ХДАК, 2020. 267 с.

61. Шиян О. 34 кілометри документів: Державний архів Львівської області оцифровує сховище. *ZAXID.NET*. URL : https://zaxid.net/34_kilometri_dokumentiv_otsifruvannya_v_arhivi_lvivskoyi_oblasti_n1576759 (дата звернення: 21.08.2024).

62. Шпарик О. Концептуальні засади цифрової трансформації освіти: європейський та американський дискурс. *Український Педагогічний журнал*. 2020. № 4. С. 65–76. <https://doi.org/10.32405/2411-1317-2021-4-65-76>.

63. ЮНЕСКО виділяє \$1,5млн на цифровізацію культурної спадщини України. *Міністерство культури та інформаційної політики: офіційний вебсайт*. URL : <https://mkp.gov.ua/news/9246.html> (дата звернення 30.08.2024).

64. Як оцифрувати документи: сканування та розпізнавання тексту. *Lube*. URL : <https://lube.ua/news/yak-otsifruvati-dokumenty-skanuvannya-ta-rozpiznavannya-tekstu> (дата звернення: 12.08.2024).

65. About us. *Time Machine Europe*. URL: <https://www.timemachine.eu/about-us/> (date of access: 11.11.2024).

66. Andhe S., Kawadkar N., Rani D. BK. Convolutional Neural Networks Based on Text Recognition. *SSRN Electron*. 2022.

67. DCam-ukrainian. *FamilySearch* URL : <https://cms-b-assets.familysearch.org/8a/ae/bb458a9d47a28c843c37edf2b693/dcam-9-0-ukrainian.pdf> (date of access: 21.10.2024).

68. Digitization. Oxford English Dictionary. URL: <https://www.oed.com/search/dictionary/?scope=Entries&q=digitization> (date of access: 19.08.2024).
69. Discover Europe's digital cultural heritage. *Europeana*. URL: <https://www.europeana.eu/en> (date of access: 11.11.2024).
70. Dots per inch. *Wikipedia*. URL : https://uk.wikipedia.org/wiki/Dots_per_inch (date of access: 30.12.2024).
71. FamilySearch. *Wikipedia*. URL: <https://uk.wikipedia.org/wiki/FamilySearch> (date of access: 04.11.2024).
72. FamilySearch's 2023 Year in Review. *FamilySearch*. URL : <https://www.familysearch.org/en/blog/2023-familysearch-year-in-review> (date of access: 04.11.2024).
73. Google оцифрує понад 100 тисяч стародруків. *Читомо*. URL: <https://chytomo.com/google-otsyfruie-ponad-100-tys-starodrukiv/> (дата звернення: 04.11.2024).
74. Lu Z. Multi-Objective Evolutionary Design of Deep Convolutional Neural Networks for Image Classification. *IEEE Trans. Evolution Computation*, 2020. P. 1.
75. Mast Garrett. A wee bit of cleanup: updating scottish locality pages on the FamilySearch research Wiki. Undergraduate Honors Theses. URL : https://scholarsarchive.byu.edu/studentpub_uht/398 (date of access: 14.11.2024).
76. Meng F., Ghena B. Research on Text Recognition Methods Based on Artificial Intelligence and Machine Learning. *Advances Comput. Communication*. Vol. 4, No. 5, 2023. P. 340– 344.
77. Nikon D850. *Hotline.ua*. URL : <https://hotline.ua/ua/av-fotoapparaty/nikon-d850-series/> (date of access: 01.01.2025).
78. Nikon Z7 II. *Hotline.ua*. URL : <https://hotline.ua/ua/av-fotoapparaty/nikon-z7-ii-kit-24-70mm/?tab=about> (date of access: 01.01.2025).
79. Ofiary antypolskiego terroru w Związku Sowieckim 1934-1938. *Ofiary terroru*. URL : <http://www.ofiaryterroru.pl/> (date of access: 11.11.2024).

80. Program digitalizacji dóbr kultury oraz gromadzenia, przechowywania i udostępniania obiektów cyfrowych w Polsce 2009–2020. URL : <https://www.nck.pl/badania/raporty/program-digitalizacji-dobr-kultury-orazgromadzenia-przechowywania-i-udostepniania> (date of access: 04.11.2024).

81. Tarandintsy, Lubny, Poltava, Ukraine, Soviet Union. ID № 112988209. *FamilySearch*. URL : <https://www.familysearch.org/ark:/61903/3:1:3QJH-VQGY-JFWZ?view=explore> (date of access: 30.12.2024 p.).

82. What to Expect from FamilySearch in 2024. *FamilySearch*. URL : <https://www.familysearch.org/en/blog/whats-new-in-familysearch-2024> (date of access: 04.11.2024).