

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ МІЖНАРОДНИХ ВІДНОСИН
КАФЕДРА КОМП'ЮТЕРНИХ МУЛЬТИМЕДІЙНИХ ВИДАНЬ

ДОПУСТИТИ ДО ЗАХИСТУ
Завідувач випускової кафедри
_____ О.А.Бобарчук
« _____ » _____ 2023р.

КВАЛІФІКАЦІЙНА РОБОТА

(ПОЯСНЮВАЛЬНА ЗАПИСКА)

ВИПУСКНИКА ОСВІТНЬОГО СТУПЕНЯ МАГІСТР
ЗА СПЕЦІАЛЬНІСТЮ 186 «ВИДАВНИЦТВО ТА ПОЛІГРАФІЯ»

Тема: «Технології оцифрування архівних документів»

Виконавець _____ студентка групи ВП-213 Луценко Світлана Костянтинівна
(студент, група, прізвище, ім'я, по батькові)

Керівник _____ к.т.н Бобарчук Олександр Антонович
(науковий ступінь, вчене звання, прізвище, ім'я, по батькові)

Нормоконтролер _____

(підпис)

С.М. Гальченко

(ПІБ)

КИЇВ 2023

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет міжнародних відносин

Кафедра комп'ютерних мультимедійних технологій

Спеціальність 186 «Видавництво та поліграфія»

(шифр, найменування)

ЗАТВЕРДЖУЮ

Завідувач кафедри

О.А.Бобарчук

« _____ » _____ 2023р.

ЗАВДАННЯ

на виконання кваліфікаційної роботи

Луценко Світлани Костянтинівни

(П.І.Б. випусника)

1. Тема роботи «Технології оцифрування архівних документів»

затверджена наказом ректора № 1901/ст від 22.09.2023

2. Термін виконання роботи (проекту): з. 02.10.2023 по 31.12.2023

3. Вихідні дані до роботи (проекту): теоретичні відомості, текстовий та графічні матеріали за тематикою дослідження,

4. Зміст пояснювальної записки: аналіз оцифрування архівних документів, аналіз програмного забезпечення для оцифрування архівних документів, практична реалізація архівів технології оцифрування архівних документів.

5. Перелік обов'язкового ілюстративного матеріалу: презентаційний матеріал..

6. Календарний план-графік

№ пор	Завдання	Термін виконання	Підпис керівника
1	Ознайомитись з постановкою задачі та здійснити аналіз предметної області	02.10.23 – 06.10.23	
2	Визначити поняття та методи оцифрування	07.10.23 – 20.10.23	
3	Проаналізувати програмне для створення архівних документів	21.10.23 – 02.11.23	
4	Дослідити етапи оцифрування архівних документів	03.11.23 – 12.11.23	
5	Обрати види сканерів, які необхідні для сканування документів	13.11.23 – 20.11.23	
6	Створити оцифровані архівні документи для компанії	21.11.23 – 01.12.23	
7	Підготувати презентаційний матеріал	02.12.23 – 08.12.23	

7. Дата видачі завдання: « 02 » жовтня 2023 р.

Керівник кваліфікаційної роботи

_____ (підпис керівника)

Бобарчук О.А.

_____ (ПІБ)

Завдання прийняв до виконання:

_____ (підпис випускника)

Луценко С.К.

_____ (ПІБ)

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи «Технології оцифрування архівних документів» складається з 87 сторінок, містить 39 рисунків та 31 використане джерел.

АРХІВИ, АРХІВНІ ДОКУМЕНТИ, ДОКУМЕНТИ, СКАНУВАННЯ ДОКУМЕНТІВ, ОЦИФРУВАННЯ.

Об'єкт дослідження – процес та технології оцифрування архівних документів.

Предмет дослідження – технології оцифрування архівних документів.

Мета роботи – оптимізація зберігання архівів через технології оцифрування.

Методи дослідження – аналіз сучасних програмних та апаратних рішень для оцифрування архівних документів, проведення практичних експериментів з використання різних технологій оцифрування для архівів.

Програмні засоби: текстовий редактор *Microsoft Word*, програма для сканування, програма для обробки текстової частини документів.

Результати кваліфікаційної роботи рекомендується використовувати під час проведення наукових досліджень та в практичній діяльності фахівців-поліграфістів.

ЗМІСТ

ВСТУП.....	Ошибка! Закладка не определена.
РОЗДІЛ 1. АНАЛІЗ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ	9
1.1 Основні поняття та методи оцифрування.....	9
1.2. Переваги та недоліки оцифрування документів ...	Ошибка! Закладка не определена.
1.3. Історія розвитку оцифрування архівних матеріалів	18
1.4. Оцінка архівних документів.....	26
1.5. Види архівних документів.....	28
Висновки до розділу 1.....	49
РОЗДІЛ 2. АНАЛІЗ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ.....	50
2.1. Аналіз програмного забезпечення для обробки текстової частини.....	50
2.2. Програма для сканування документів.....	55
2.3. Програма для обробки документів	59
2.4. Програма для обробки інформації з документів.....	61
Висновки до розділу 2	66
РОЗДІЛ 3. ПРАКТИЧНА РЕАЛІЗАЦІЯ АРХІВІВ ТЕХНОЛОГІЇ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ	67
3.1. Основні вимоги до обладнання.....	67
3.2. Вимоги до оцифрування (сканування) документів	71
3.3. Процес оцифрування документів	76
Висновки до розділу 3	82
ВИСНОВКИ.....	83
СПИСОК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ ВИКОРИСТАНИХ ДЖЕРЕЛ.....	85

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ, ТЕРМІНІВ

Документ – це письмовий чи друкований текст, який містить інформацію, факти, правила чи інші дані. Це може бути щось написане на папері, електронний файл або будь-який інший носій інформації. Документи використовуються для передачі знань, фіксації правил, реєстрації подій та багатьох інших цілей.

Оцифрування – це процес перетворення аналогової інформації у цифровий формат. Це може включати в себе перетворення аналогових сигналів у цифрові дані або перенесення інформації з фізичних носіїв, таких як папір або плівка, у цифровий вигляд.

Сканування – це процес отримання цифрового зображення або репрезентації об'єкта за допомогою сканера або іншого пристрою для отримання зображення. Сканування часто використовується для перетворення фізичних документів, фотографій чи інших об'єктів у цифровий формат.

Архіви – це місця або системи для зберігання, управління та збереження документів, записів або інших матеріалів, які мають історичне, наукове, культурне або адміністративне значення. Також термін "архів" використовується для позначення колекції документів або записів, які зберігаються разом у відповідному місці.

ВСТУП

Актуальність теми. Сучасний етап розвитку суспільства безперервно визначає нові вимоги до зберігання та обробки інформації, особливо в контексті архівного зберігання. Завдяки стрімкому розвитку технологій, які нас оточують, та зростанню обсягів інформації, яку потрібно зберігати та обробляти, актуальність проблеми оцифрування архівних документів стає невід'ємною частиною сучасної архівістики та інформаційної науки.

Оцифрування архівних документів визнається однією з ключових стратегій збереження та забезпечення доступу до історичної та культурної спадщини. Заміна традиційних, паперових носіїв інформації цифровими форматами відкриває нові можливості для збереження, використання та дифузії архівних документів. У контексті цієї магістерської роботи, дослідження спрямоване на аналіз і вдосконалення технологій оцифрування архівних документів з метою забезпечення їхньої ефективної збереженості та доступності для наступних поколінь.

Процес оцифрування архівних документів є складним і має свої особливості, пов'язані як із технічними, так із науково-методичними аспектами. Вирішення цих завдань вимагає інтеграції сучасних інформаційних технологій, методів обробки та аналізу даних, а також розроблення ефективних стратегій збереження цифрових архівних матеріалів.

Однією з ключових проблем, які вивчаються в даній роботі, є забезпечення якості та автентичності цифрових копій архівних документів. Стандартизація процесів оцифрування, розробка методів верифікації та контролю якості є важливими етапами у створенні стійких та довірених цифрових архівів.

Досягнення успіху в області технологій оцифрування архівних документів може відкрити нові перспективи для наукового співробітництва, культурного обміну та освіти. Ця робота спрямована на вирішення викликів, пов'язаних із забезпеченням якості та доступності архівних документів у цифровому форматі, щоб зберегти та розвивати національну культурну спадщину.

У цьому контексті кваліфікаційна робота висвітлює та аналізує актуальні технологічні та методологічні аспекти оцифрування архівних документів, а також пропонує конкретні рекомендації для оптимізації цього процесу відповідно до сучасних вимог і стандартів збереження інформації.

Об'єкт дослідження: процес та технології оцифрування архівних документів.

Предмет дослідження: технології оцифрування архівних документів.

Метою кваліфікаційної роботи оптимізація зберігання архівів через технології оцифрування.

Завдання:

1. Аналіз сучасних технологій оцифрування.
2. Розробка методів впровадження оцифрування архівів.
3. Визначення стратегій зберігання та доступу до оцифрованих документів.
4. Оцінка ефективності впроваджених технологій.
5. Рекомендації для оптимізації управління архівами через оцифрування.

Методи дослідження: аналіз сучасних програмних та апаратних рішень для оцифрування архівних документів, проведення практичних експериментів з використання різних технологій оцифрування для архівів

Наукова новизна отриманих результатів. Наукова новизна отриманих результатів полягає у синтезі теоретичних і практичних знань, отриманих щодо теми кваліфікаційної роботи. Синтезовані знання будуть практично використані в рамках навчального процесу університету

Апробація отриманих результатів. Результати отриманих в роботі досліджень були оприлюднені на Всеукраїнська науково-практична конференція з міжнародною участю «Мультимедійні технології в освіті та інших сферах діяльності», 2023 рік.

Практичне значення отриманих результатів. Результати досліджень кваліфікаційної роботи рекомендується використовувати в структурі навчального процесу кафедри, можуть бути корисними для студентів спеціальності 186 "Видавництво та поліграфія" щодо підвищення знань за темою "Мультимедіа та комп'ютерні мережі".

РОЗДІЛ 1

АНАЛІЗ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ

1.1. Основні поняття та методи оцифрування документів

Оцифрування документів – це процес конвертації інформації з паперових документів у цифровий формат. Такий підхід є корисним для збереження даних, які вміщені в паперових носіях, а також для зменшення обсягу архівів, полегшення доступу та обробки інформації.

Існують різні методи оцифрування, включаючи сканування та використання технології оптичного розпізнавання тексту (*OCR*). У даному тексті ми розглянемо процес оцифрування документів через застосування сканування та технології *OCR* [1].

Сканування документів

Сканування документів – це процедура отримання електронної версії паперового документа. Для виконання сканування документів використовується спеціальний пристрій, такий як сканер, який перетворює фізичний документ у файл у форматі *PDF*, *JPG* або іншого електронного формату.

Після завершення процесу сканування документ можна зберегти на комп'ютері або в хмарному сховищі. У випадку, якщо документ має кілька сторінок, їх можна об'єднати в один файл.

Хоча сканування дозволяє зберігати дані у електронному форматі, важливо зауважити, що цей метод не надає можливості редагування тексту на документі. Таким чином, для внесення змін до документа потрібно використовувати ручне редагування, що може бути часо- та ресурсозатратним процесом.

Яку якість сканування вибрати

Вибір якості цифрового документа залежить від конкретних завдань, які стоять перед нами. Зазвичай використовують показники роздільної здатності зображення в межах від 200 до 1200 *dpi* (точок на дюйм). Для сканування простих малюнків чи художніх творів, достатньо встановлення значень в 400-500 *dpi*. Однак, якщо маємо

справу з оцифруванням карт чи інших зображень, які потрібно значно збільшити та зберегти всі деталі, рекомендується використовувати максимальну роздільну здатність.

Для оцифрування креслень можна встановити менше значення, наприклад, 300-400 *dpi*, оскільки вони зазвичай добре проглядаються із меншою роздільною здатністю. Для звичайних копій сторінок книг, паспортів чи інших повсякденних документів часто вистачає 300 *dpi*. Щоб подальше редагування зображення, обрізання та зміна передачі кольорів були зручними, використовують стандартні формати, такі як *jpeg*, *pdf*, *tiff*, в яких файли зберігаються сканерами автоматично [1].

Розпізнавання тексту (*OCR*)

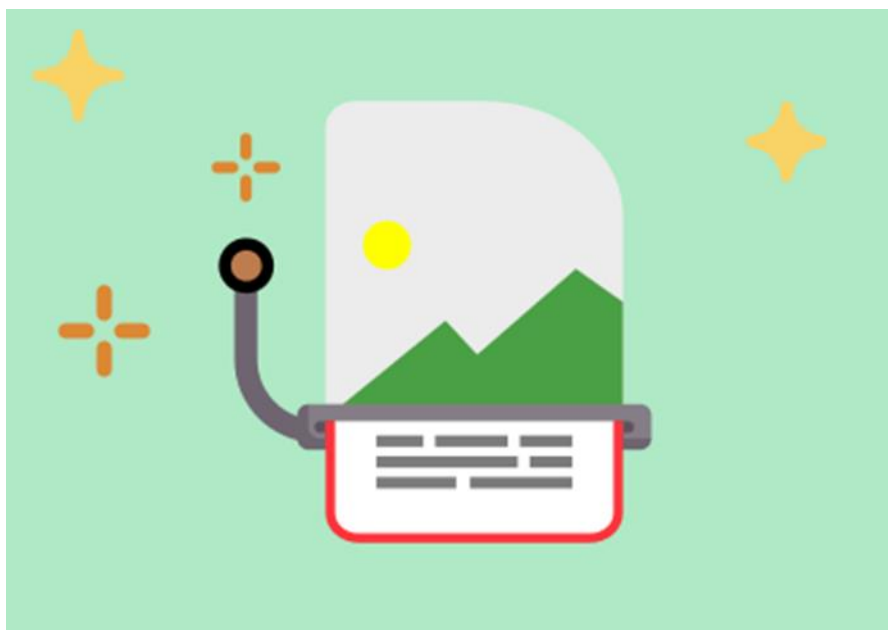


Рис. 1.1. Розпізнавання тексту (*OCR*)

Для редагування тексту у документі можна використовувати технологію OCR. OCR – це процес розпізнавання тексту на сканованому документі та його конвертація у редагований електронний формат. Для проведення OCR процедури використовуються спеціальні програми, такі як Adobe Acrobat Pro, ABBYY FineReader, Readiris та інші, які автоматично розпізнають текст на сканованому документі та перетворюють його у редагований електронний формат.

Цей процес виконується за допомогою звичайного сканера, і отриманий шаблон завантажується у програму розпізнавання тексту. Двигун OCR аналізує зображення відносно його компонентів, розрізняючи між зображенням та текстом. Цей етап є ключовим для подальшого редагування.

Процес OCR може займати певний час та вимагати додаткових ресурсів комп'ютера, особливо при наявності складної структури та графіки в документі. Однак для документів із простим текстом, процес OCR може бути проведений швидко та точно.

OCR також може бути використаний для автоматичного розпізнавання тексту у великих обсягах документів, спрощуючи обробку даних у компаніях з великою кількістю документації. Важливо враховувати, що точність OCR залежить від якості сканування та програми OCR, і рекомендується використовувати високоякісне обладнання та програми для досягнення оптимальних результатів [1].

Розпізнавання рукописного тексту завдяки *ICR*

ICR (Інтелектуальне Розпізнавання Символів) є логічним розвитком технології *OCR*. Це включає докладний аналіз та оцінку результатів сканування, з особливим акцентом на семантичний контекст. У відміню від *OCR*, яке лише розпізнає текст та зображення, *ICR* проводить внутрішній аналіз тексту, визначаючи його доцільність з точки зору змісту. Ця техніка особливо ефективна при розпізнаванні схожих символів, наприклад, 8 або B, і значно підвищує точність цифрового розпізнавання тексту. Навіть якщо оригінал містить вицвілі літери внаслідок віку, *ICR* здатне їх ефективно розпізнавати та оцифровувати.

Однією з важливих областей застосування *ICR* є розпізнавання рукописних текстів, яке часто становить виклик для звичайного програмного забезпечення розпізнавання тексту. З використанням вбудованого *ICR* такі рукописні тексти можуть бути легко та ефективно оцифровані.

Онлайн розпізнавання тексту

Платформи для текстового розпізнавання онлайн, включаючи OCR, пропонують швидке і зручне рішення для користувачів, які не можуть встановлювати

повноцінні програми для розпізнавання тексту. Ці сервіси зазвичай пропонують безкоштовні та платні плани, враховуючи обсяг та складність документів.

Кілька популярних онлайн сервісів можна знайти за запитами, такими як "сканування тексту на зображенні" чи "онлайн OCR". Вони дають можливість завантажувати скановані документи у різних форматах та отримувати розпізнаний текст у вигляді електронних документів, які можна редагувати.

Сервіси для стандартного розпізнавання тексту, такі як Google Docs та Dropbox Paper, зручні для редагування та збереження документів у хмарному сховищі. Вони використовують технологію розпізнавання тексту для автоматичного перетворення зображення тексту у редагований формат.

Проте важливо відзначити, що точність онлайн сервісів для розпізнавання тексту може не досягати рівня повноцінних програм для OCR. Їх можливості та точність можуть бути обмеженими залежно від складності документів та якості сканування. Тому, якщо важлива висока точність та якість результату, рекомендується використовувати повноцінні програми для OCR.

У будь-якому випадку, онлайн сервіси для розпізнавання тексту можуть виявитися дуже корисними, коли необхідно швидко та просто отримати редактований електронний документ із сканованого матеріалу [1].

Існує кілька методів оцифрування документів, які використовуються залежно від характеристик документів та потреб користувача. Деякі з основних методів оцифрування документів включають:

Сканування:

Процес оцифрування може включати різноманітні методи, такі як сканування, фотографування та використання електронних форм. Серед них:

1. Одноманітне сканування: Простий метод, коли документ сканується в один колір без додаткових обробок.
2. Кольорове сканування: Зображення документа зберігається у кольоровому форматі для більшої деталізації.
3. Оптичне розпізнавання тексту (*OCR*):

Розпізнавання тексту: Автоматичне розпізнавання тексту на сканованих документах та його перетворення в редагований текстовий формат.

4. Фотографування:

Використання фотокамери: Замість сканування, можна використовувати цифрову фотокамеру для створення фотографій документів.

5. Електронні форми та документи:

Заповнення електронних форм: Створення електронних версій документів для електронного заповнення та редагування.

6. Системи управління документами (DMS):

Оцифрування при введенні: Системи *DMS* автоматично оцифрують та класифікують документи при їх завантаженні до системи.

7. Розпізнавання рукописного тексту:

Розпізнавання рукопису: Технології, які дозволяють розпізнавати рукописний текст.

8. Конвертація відео та аудіо з файлів:

Транскрипція аудіо та відео: Автоматична транскрипція аудіо- або відеофайлів у текст.

9. Сканування *QR*-кодів або штрих-кодів:

Ідентифікація за допомогою кодів: Використання *QR*-кодів та штрих-кодів для ідентифікації та зв'язку з електронними версіями документів.

Вибір конкретного методу оцифрування залежить від типу документів, їх характеристик і потреб користувача в подальшому використанні цифрової інформації.

1.2. Переваги та недоліки оцифрування документів

Створення електронних копій бізнес-документів є необхідною частиною поточного документообігу, формування електронного архіву та вирішення різних завдань. У випадку невеликих обсягів це завдання виконується легко співробітниками компанії за допомогою стандартного обладнання. Проте, для сканування специфічних

документів, широкоформатних креслень, книг та інших матеріалів може знадобитися багато часу та/або використання спеціалізованих пристроїв. У таких ситуаціях звертаються за допомогою до професійних служб з сканування документів. Для оцифрування документів можна використовувати різноманітні сканери, але звичайні моделі мають свої обмеження, такі як швидкість роботи, формат та, головне, точність передачі. Тому, коли виникає потреба в оцифруванні важливих документів компанії, часто звертаються за допомогою до фахівців, які володіють спеціалізованим обладнанням та програмним забезпеченням [2]. Це надає можливість:

- сканувати різні джерела інформації будь-якого формату та кольорової палітри, такі як звичайні текстові документи, зображення, креслення, фотографії та інші;
- оцифрувати будь-які початкові джерела, незалежно від типу паперу, його товщини та розміру;
- забезпечувати високу точність передачі, що має особливе значення при роботі з кресленнями та іншими первинними матеріалами, які містять деталі малого розміру;
- оцифрувати документи з великою швидкістю, оскільки професійна техніка працює набагато швидше, ніж звичайні сканери;
- приводити відскановані документи до форми, придатної для подальшої обробки.

Спеціалізоване сканування документів має широкі можливості та приносить значні переваги, такі як виконання заощаджень у часі та зусиллях працівників. Крім цього, воно усуває потребу в встановленні спеціалізованого програмного забезпечення, розділенні та скріпленні документів із підшитих папок та інших рутинних маніпуляціях. Професіонали, які працюють у сфері сканування документів, не лише мають доступ до високоякісної техніки та програм, але й володіють досвідом, що дозволяє виконувати завдання швидко та точно. Таким чином, професійне сканування документів гарантує отримання результату високої якості протягом короткого періоду часу із мінімізацією витрат [2].



Рис. 1.2. Сканування документів

Між переваг оцифрування можна відзначити наступне:

- Збереження коштів на папері та друкарській фарбі, оскільки документи архівуються у цифровому форматі.
- Зменшення використання паперу призводить до зменшення утворення відходів, що сприяє екологічній діяльності.
- Зручність у спільному використанні цифрових документів. Наприклад, при проведенні процедур в іншому місті чи країні документи можна легко відправити електронною поштою, що значно економить час порівняно з традиційним відправленням паперових документів поштою.
- Зручний пошук файлів, які зберігаються у цифровому форматі, на відміну від пошуку серед купи паперових документів.
- Боротьба з неформальною економікою та корупцією в урядових структурах через впровадження оцифровки [3].

Однак оцифрування має свої негативні аспекти:

- Не всі особи можуть легко освоїти цифрові носії, особливо, якщо вони мають похилість або є людьми похилого віку. Таким чином, навіть при впровадженні

організацією цифрових процедур, деяким користувачам може бути не зручно або неприязно до нової системи.

– Цифровий формат не завжди гарантує відсутність помилок. Втрати файлів можуть виникнути через людські помилки чи віруси.

– Процес оцифрування зазвичай не є миттєвим, і для передачі всіх файлів, наприклад, від фізичного носія до цифрового, потрібний певний час.

– При обміні документами чи інформацією у цифровому форматі завжди існує ризик, що дані можуть потрапити в руки несанкціонованих осіб, таких як хакери.

– Низька поширеність технології електронного цифрового підпису.

– У випадку архівних установ для юридичної ваги електронної копії документа.

– Механічне навантаження на палітурку документа.

– Вплив світла та температури під час сканування (зокрема, якщо використовується непридатне обладнання) [5].

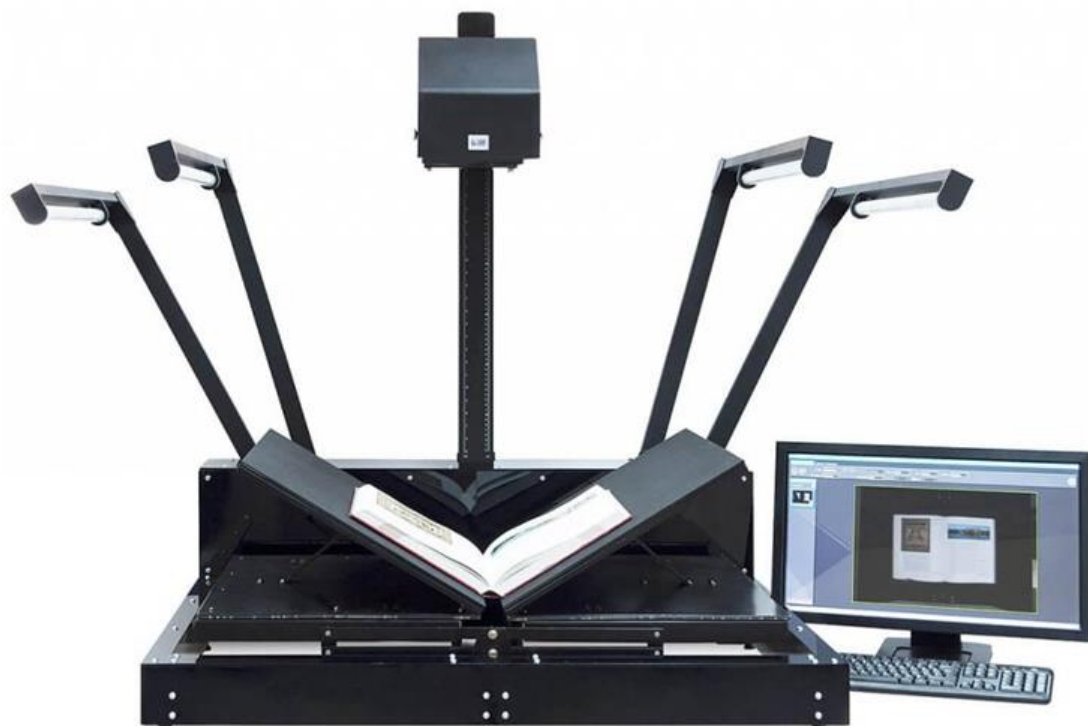


Рис. 1.3. Недоліки оцифрування документів

Головна перевага цифрового архіву полягає в швидкості та зручності процесу пошуку необхідної інформації. При правильній класифікації та структуризації

потрібні документи можна знайти всього за кілька хвилин. Пошук можливий за різними критеріями, такими як категорії, дата складання, автор, або одразу кілька атрибутів. Також можливий пошук за вмістом, що, звісно, є більш зручним порівняно з традиційними папками та стелажми. Для ефективної роботи такої інформаційної мережі особлива увага при ретроконверсії приділяється атрибутуванню, індексуванню цифрових документів, а також формуванню бази даних [4].

Ще однією важливою перевагою оцифрування документів є забезпечення їхнього захисту та безпеки. У відміну від паперових оригіналів, цифрові копії не піддаються фізичному зносу та можуть зберігатися протягом необмеженого періоду часу. Більше того, в цифрових архівах реалізовані надійні механізми захисту вмісту від навмисного або випадкового видалення, а також пошкодження. Інформацію неможливо видалити помилково, завдяки багаторівневому контролю, який обмежує доступ до неї осіб без відповідних прав. Використання технології резервного копіювання бази допомагає максимально захистити та зберегти дані. Наявність доступу для багатьох користувачів стає необхідною перевагою, особливо, коли з архівними документами одночасно працюють різні співробітники в різних відділах або локаціях. У випадку паперових оригіналів уникнути копіювання або сканування для подальшої обробки може бути проблематично, тоді як цифрові копії легко обробляються без цих обмежень.

До речі, це спрощує не лише бізнес-процеси, а й унеможливорює ситуацію безконтрольного копіювання або сканування конфіденційних документів. Налаштувавши права доступу, працювати з документацією можуть лише ті працівники, які мають відповідні повноваження. При цьому доступ можна отримати, не виходячи з робочого місця, а з правами працювати з цифровими документами можна з будь-якої точки світу за умови наявності стабільного Інтернет-підключення.

Економія офісного простору – це додатковий плюс цифрових архівів. Якщо для невеликих компаній це може не бути настільки актуальним, великі підприємства з щоденним документообігом оцінять цю перевагу в повній мірі. Великі сервери із базами даних можуть ефективно замінити численні полиці з папками та стелажі, що

робить оренду офісних або виробничих приміщень менш обтяжливою для бюджету компанії.

При обговоренні переваг оцифрування документів необхідно також враховувати можливість відновлення пошкоджених документів під час ретроконверсії. Якщо паперові копії частково пошкоджені під час зберігання, дані можна відновити (реставрувати) у процесі цифрового перетворення, залучаючи уповноважених працівників підприємства [4].

Інші переваги

– Інтеграція системи документообігу підприємства в єдину інформаційну мережу ІТ дозволяє зберігати різні види документів, такі як кадрові, фінансові, розпорядчі, технічні та проектні.

– Можливості копіювання та миттєвого обміну документами між уповноваженими співробітниками за допомогою електронної пошти або інших засобів зв'язку.

– Одночасна робота з документом для необмеженої кількості користувачів, які мають відповідні права доступу.

– Зберігання у системі сканів паперових оригіналів у форматі електронних документів будь-якого типу.

– Моніторинг дій користувачів: можливість визначити, хто і коли отримував доступ до інформації та вносив зміни. Це забезпечує прозорість та безпеку документообігу в компанії на всіх рівнях [5].

1.3. Історія розвитку оцифрування архівних матеріалів

В рамках архівного законодавства України, зокрема в Законі України "Про Національний архівний фонд та архівні установи" (2001), архівний документ розглядається як рухомий майно, власником якого може бути як держава, так і приватна особа. В українському архівознавстві термін "архівний документ" не є ідентичним поняттю "документ Національного архівного фонду" (НАФ). Останнє визначається як документ, культурна цінність якого визнана за результатами

експертизи, проведеної комісією з фахівців архівної справи і справочинства, представників наукової і творчої громадськості, і який підлягає державному обліку і зберіганню. Документи НАФ можуть перебувати в будь-якій формі власності, передбаченій Конституцією та законами України. Водночас, архівні документи, які були на державному обліку та зберігалися в державних архівних установах до 1991 року, повністю відносяться до НАФ і є власністю держави. Вилучення документів НАФ у власника або уповноваженої ним особи без їхньої згоди забороняється, за винятком випадків, передбачених законами України.

Документи, які накопичилися протягом функціонування державних органів, підприємств, установ, та організацій, а також ті, які передані в державні архівні установи іншими юридичними і фізичними особами без збереження права власності, є власністю держави. Аналогічно, архівні документи, сформовані впродовж діяльності органів місцевого самоврядування, комунальних підприємств, установ, та організацій, а також ті, які передані архівним відділам міських рад іншими юридичними і фізичними особами без збереження права власності, стають власністю територіальних громад.



Рис. 1.4. Історія архівних документів

У випадках, коли архівні документи не мають визначеного власника або їхній власник невідомий, вони переходять у власність держави на підставі рішення суду, яке виноситься за заявою державної архівної установи, за винятком випадків, що передбачені законами України. Для правильного розуміння суті архівних документів важливим є поняття "життєвий цикл документа," яке введено у науковий обіг американським архіваріусом та президентом Товариства американських архівістів Ф. Бруксом у 1940 році. Життєвий цикл документа може мати різні етапи в залежності від особливостей функціонування ділових систем і принципів організації архівної справи в різних країнах. Загалом вважається, що цей цикл включає три етапи: 1) документ використовується як активний регулятор управлінських відносин (приблизно 5 років); 2) документ припиняє виконання своїх функцій і ще не використовується як історичне джерело (до 10 років); 3) документ стає історичним джерелом, його меморіальна функція потребує архівного (вічного) зберігання.

Для забезпечення відповідних умов зберігання архівних документів слід враховувати матеріальну основу кожного документа та, за потреби, проводити комплексні заходи з його реставрації і консервації. Також важливо створювати страхові копії, забезпечувати належний температурно-вологісний режим в архівосховищі та попереджувати несанкціонований доступ та втрати.

У бібліотеках та інших інформаційних установах, де зберігається значний обсяг документів на різних носіях, деякі з них можуть мати ознаки архівних документів. Зокрема, це можуть бути документи, які циркулювали в справочинстві установи для забезпечення її діяльності (наприклад, накази, положення, інструкції, протоколи нарад, доповідні записки, плани, звіти, особові справи, листування тощо). Ці документи, втративши чинність, можуть переходити до архівного підрозділу (архіву) бібліотеки, перебуваючи в другій та третій стадіях свого життєвого циклу.

Найцінніша частина архівних документів, що зберігаються в установі, після закінчення терміну зберігання передається на постійне зберігання до державного або іншого архіву. Ці документи належать до Національного архівного фонду (НАФ), а позначення "Постійно" вказує, що їх слід зберігати довічно під охороною держави. Щодо інших архівних документів, які створюються у діяльності державних органів

та інших установ, згідно з Переліком типових документів із зазначенням строків зберігання, вони класифікуються як тимчасові (від 3 до 5 років) або тривалі (від 10 до 75 років). Після закінчення цих термінів зазвичай відбувається їх утилізація.

Для певних категорій архівних документів, які мають важливе практичне значення на тривалій період, конкретні строки зберігання можуть не визначатися. Наприклад, архівні документи із відміткою "Доки не мине потреба" підкреслюють, що тривалість їх зберігання визначається самостійно організаціями, але не може бути меншою одного-трьох років. Проте важливо зауважити, що для документів, які мають юридичну чинність протягом певного періоду (наприклад, договори, контракти на постачання продукції, облікові документи про товарно-матеріальні цінності і т. д.), такий строк не може бути встановлений [6].

Існують архівні документи, на яких вказано строк зберігання «До ліквідації організації». Такі документи зберігаються в організації безстроково, а при ліквідації організації їх піддають повторній експертизі щодо цінності. Залежно від результатів цієї експертизи документи, які стосуються прав громадян, передаються з описами справ до місцевих архівних установ. Ці установи, створені органами виконавчої влади та місцевого самоврядування, зберігають архівні документи, що не входять до Національного архівного фонду.

З урахуванням умов зберігання, використання та обліку, ці архівні документи фактично розглядаються як архівні, хоча вони не мають офіційного обліку як такі. Це стосується особливої частини бібліотечного фонду, яка включає унікальні, рідкісні, особливо цінні видання, краєзнавчі матеріали, книжкові пам'ятки та інше. Ці документи не доступні для загального користування, і їх не видають читачам через їхню культурну, наукову та історичну цінність. Це особливо стосується фондів національних бібліотек України та інших спеціальних бібліотек, які виконують роль сховища документно-інформаційних ресурсів державного значення. Такі документи не підлягають списанню, за винятком випадків стихійних лих та надзвичайних обставин.

До початку Другої світової війни виявлення цінних архівних документів у фондах бібліотек Української Радянської Соціалістичної Республіки було звичайним

явищем. Утворення цих збірок мало різноманітний характер, відбуваючись як результат систематичного збирання старовинних книг і рукописів громадськими організаціями, культурно-просвітницькими установами, навчальними закладами, науковими товариствами, а також окремими особами-колекціонерами.

Пізніше такі збірки формувалися за волею засновників фондів, а також внаслідок ліквідації або реорганізації численних інституцій радянською владою у період з 1918 по 1940 рік. Наприклад, один із чотирьох відділів Центральної єврейської бібліотеки імені М. Вінчевського в Києві називався "архівним" та зберігав фонди архівних документів, що виникли внаслідок діяльності єврейських політичних партій, таких як Бунд, Поалей-Ціон, Фолкспартей, а також документи організацій і установ, ліквідованих у 1920–1930-х роках. Також значна частина фонду Центральної польської державної бібліотеки в Києві мала статус архіву, і документи з цього фонду не надавались читачам для вивчення вдома. Київська обласна бібліотека імені ВКП(б), тепер Національна бібліотека України імені Ярослава Мудрого, мала велику та цінну колекцію архівних документів, яку було безповоротно втрачено під час нацистської окупації Києва.

З вересня 1992 року в складі Національної бібліотеки України імені В.І. Вернадського (НБУВ) функціонує Інститут архівознавства, який є основним архівним підрозділом у системі НАН України. Основні завдання Інституту включають у себе питання формування, зберігання, обліку та використання документів Архівного фонду НАН України. Крім того, Інститут виконує роль науково-методичного центру для архівних підрозділів НАН України і надає консультаційну та практичну допомогу з питань документування академічних установ, формування архівних фондів і функціонування ділових і архівних підрозділів. Історично Інститут виник як результат перетворень попередніх структур: від березня 1968 року до січня 1984 року він діяв як Центральний науковий архів АН УРСР, а пізніше, до 1992 року, як відділ архівних фондів. В Інституті зберігається майже 300 особових фондів визначених вчених, загалом налічуючи понад 40 000 архівних справ, що включають сотні тисяч архівних документів.



Рис. 1.5. Бібліотека Вернадського

Більшість документів, розташованих у структурних підрозділах Національної бібліотеки України імені В.І. Вернадського (НБУВ), таких як Інститут рукопису НБУВ, Фонд Президентів України (унікальна бібліотечно-архівна колекція), Львівська національна наукова бібліотека України імені Василя Стефаника (відділи рукописів, україніки, рідкісних книг; науковий відділ періодичних видань імені Мар'яна та Іванни Коців; відділення "Палац мистецтв імені Тетяни і Омеляна Антоновичів"; кабінет картографії), а також у Науковій бібліотеці Запорізького відділення Інституту української археографії та джерелознавства імені М.С. Грушевського (відділ рукописів), також мають архівний характер. Архівні фонди присутні також у інших бібліотеках, таких як Харківська державна наукова бібліотека імені В.Г. Короленка, кілька обласних універсальних наукових бібліотек та наукові бібліотеки вищих навчальних закладів III–IV рівнів акредитації [6].

Архівні документи є значущою частиною фонду Державної наукової архівної бібліотеки (ДНАБ), створеної в січні 1972 року, яка функціонує як юридична особа та підпорядкована Державній архівній службі України. Особливо великий обсяг складають друковані матеріали, що колись належали Празькому українському архіву. Цей термін вжито для опису зібраних у Празі міжвоєнним періодом документальних колекцій української еміграції, які включають Український національний музей-архів

при Українському інституті громадознавства, Український історичний кабінет та Музей визвольної боротьби України. Після Другої світової війни ці та інші трофейні колекції, пов'язані з Україною, були передані в УРСР та розподілені між кількома центральними державними архівами, КДБ УРСР та ДНАБ. ДНАБ зберігає багато унікальних друкованих документів, які існують в єдиному екземплярі або відсутні в інших бібліотеках України.

Серед них – штабні мапи першої половини ХХ століття з оперативною дислокацією військових підрозділів та інші цінні матеріали, такі як радянські та білогвардійські листівки і плакати, видання періоду Української революції 1917–1921 років, друкована продукція української еміграції, матеріали УПА, а також нормативні документи та методичні розробки архівних установ. Також серед збережених документів є бібліотечна колекція Рейхскомісаріату України, що включає понад 1,500 видань, а також книги та періодика іноземними мовами.

Українська бібліотека імені Симона Петлюри в Парижі, а також інші бібліотеки української діаспори, монастирські та університетські книгозбірні в Канаді, США, Великій Британії, Польщі, Німеччині, Австрії та інших країнах є сховищами цінних архівних колекцій україніки.

Всі видання, які знаходяться у Книжковій палаті України, також суттєво володіють ознаками архівних документів. Ця наукова установа виступає в якості головного сховища всіх видів друкованої продукції, виданої в Україні. Завдяки широкому спектру функцій, що покладаються на неї, вона веде комплектування та зберігання повного і недоторканого фонду Державного архіву друку [6].

Історія розвитку оцифрування архівних матеріалів тісно пов'язана з технологічними та інформаційними змінами, які відбувалися протягом останніх десятиліть. Оцифрування архівів має кілька етапів розвитку:

- 1. Ранній етап (до 1990-х років):** У цей період оцифрування архівів в основному обмежувалося перекладанням аналогових документів в цифровий формат за допомогою сканерів і цифрових камер. Проте, обсяги інформації були обмеженими через обмежену потужність та доступність зберігання цифрової інформації.

2. **Розвиток технологій (1990-2000-ті роки):** З розвитком комп'ютерів та збільшенням їх потужності, а також із зростанням обсягу зберігання даних, оцифрування отримало нові можливості. З'явилися спеціалізовані програми для обробки та зберігання цифрових архівів.

3. **Ера Інтернету (2000-2010-ті роки):** З поширенням Інтернету зростає потреба у доступі до архівів з різних частин світу. Оцифрування стало покликане полегшити обмін інформацією та документами між різними інституціями та організаціями. Багато архівів і бібліотек розпочали активний процес оцифрування своїх колекцій.



Рис. 1.6. Історія цифрових архівів

4. **Сучасність (з 2010-х років):** Завдяки стрімкому розвитку технологій штучного інтелекту, машинного навчання та обробки природної мови, оцифрування архівів стало більш автоматизованим та ефективним. Автоматичне розпізнавання символів (OCR) дозволяє перетворювати скановані документи в редагований текст, що полегшує пошук та аналіз інформації.

5. Використання відкритих стандартів: Зростання усвідомленості щодо необхідності дотримання відкритих стандартів для зберігання та обміну архівів сприяє інтероперабельності між різними інституціями та забезпечує довгостроковий доступ до цифрових архівів.

В сучасному світі цифрування архівів виступає ключовим компонентом збереження культурної спадщини та забезпечення доступу до інформації. Розвиваючись, технології постійно удосконалюють методи цифрування, роблячи цей процес більш доступним і ефективним.

1.4. Оцінка архівних документів

Рішення щодо необхідності архівування цифрових об'єктів приймаються на підставі оцінки їхньої неперехідної цінності. Оцінювання є лише однією з стратегій, оскільки архівісти також використовують методи статистичної вибірки та аналізу ризиків для визначення матеріалів, які слід зберегти. Статистична вибірка ґрунтується на статистичних методах для створення представницької вибірки та відповідного відбору паперових та цифрових документів для архівування або дигіталізації. Метод аналізу ризиків передбачає оцінку потенційних труднощів, які можуть виникнути при роботі з цифровими ресурсами, а також їхніх можливих наслідків, таких як ризик застаріння формату файлів та носіїв, втрата, пов'язана з збереженням цифрових ресурсів тощо. Зазвичай аналіз ризиків виконується разом із іншими критеріями відбору, такими як оцінка витрат.

До 1980-х років все більше архівістів почали звертати увагу на те, що створення цифрових об'єктів та будь-яких електронних записів проводилося без урахування вимог до їх тривалого збереження. У 1990-х роках стало очевидним, що існує проблема старіння програмного забезпечення і, відповідно, неможливості надійного архівування створеної за його допомогою інформації. З впровадженням електронного архівування з'явилися перші спроби регламентації процесу. Наприклад, була створена робоча група проекту *InterPARES*, яка займалася розробкою моделі для відбору електронних об'єктів, що підлягають архівуванню [7].

Макро-оцінка включає в себе відбір електронних записів для архівування, враховуючи їх важливість та цінність в контексті, створеного організаціями чи акторами протягом певного часового проміжку. *InterPARES* модель визначає чотири основні етапи оцінки контексту електронних записів:

– **Складання та аналіз інформації про записи та їх контексти:** Рішення про архівування приймається, враховуючи різні аспекти, такі як юридично-адміністративний, фінансовий, процедурний, документальний та технологічний контексти. Це означає врахування правових рамок, фінансових структур, процедур створення записів, їхньої структури та технічних аспектів створення.

– **Оцінка ваги:** Архівісти збирають дані про цінність та автентичність електронних матеріалів. Цінність визначається з урахуванням того, наскільки матеріали служать інтересам авторів та суспільства. Якщо існує сумнів щодо автентичності записів, їхнє архівування може бути піддане сумнівам.

– **Визначення можливості збереження:** Етап визначення можливості збереження включає визначення змісту матеріалів та технічних елементів, які необхідно зберегти з точки зору справжності. Експерти обговорюють ці вимоги щодо збереження з організаціями, відповідальними за архівування записів.

– **Ухвалення рішення про оцінку:** Остаточне рішення щодо оцінки приймається на основі розгляду всіх аспектів, включаючи контекст, цінність та можливості збереження.

Мікро-оцінка визначає важливість конкретних документів, зокрема, оцінюється їхнє відповідність політиці збору та зберігання утримувача. Основним критерієм є дотримання цих документів установленної політики. Якщо документи не відповідають цим критеріям, запит на їхнє електронне архівування може бути відхилений, навіть якщо вони мають велику цінність. Наприклад, Earth Resources Observation and Science (EROS) відмовились надавати доступ до набору даних космічної програми Аполлон, оскільки вони не відповідали місії збереження фільмів та фотографій, зроблених із Землі.

Другий за важливістю критерій – це цінність цифрових ресурсів, яка може бути первинною (адміністративною, фіскальною, юридичною) та вторинною (зручність

використання та доступність). Третій критерій – вартість, яка може вплинути на вибір між різними версіями одного матеріалу. Архівісти також враховують можливість архівування, оскільки деякі зберігачі можуть відмовитися приймати цифрові ресурси через обмеження форматів файлів або носіїв, що впливає на довгострокове збереження [7].



Рис. 1.7. Оцінка архівних документів

1.5. Види архівних документів

Архівні документи можна класифікувати за різними критеріями, такими як форма, зміст, термін зберігання та інші. Ось кілька загальних видів архівних документів:

1. За формою:

Текстові документи. Системи підготовки текстових документів класифікуються за їхнім призначенням. Програми для підготовки текстів розроблені для створення різних видів письмової документації, такої як статті, листи, технічні описи та інші. Завдяки комп'ютеру та спеціалізованому програмному забезпеченню можна легко створювати документи, які традиційно підготовлялися на друкарських машинках.

Це дозволяє багаторазово вносити зміни до окремих фрагментів тексту без повторного введення всього документа. Користувач може змінювати шрифти, вставляти малюнки, створені за допомогою комп'ютера, та легко друкувати потрібну кількість примірників документу за допомогою принтера.

Програми для підготовки текстів автоматично виконують такі функції, як автоматичне форматування документів, перевірка правопису та інші. Це робить процес підготовки текстових матеріалів на комп'ютері не тільки швидшим і ефективнішим порівняно з друкарською машинкою, але і відкриває нові можливості, які раніше були недоступні.

Сучасні програми для підготовки текстових документів мають різноманітні характеристики та функціональні можливості. Вони відрізняються у способах введення і редагування тексту, форматуванні та виводі на друк, а також у рівні складності їх вивчення та використання. Зазвичай програми для підготовки текстів можна розподілити на три основні класи:

Сучасні програми для створення текстових документів суттєво відрізняються за своїми характеристиками, можливостями редагування і введення тексту, форматування та виводу на друк, а також за складністю їх вивчення і використання. Серед цих програм можна виділити три основні категорії:

- Редактори текстів,
- Редактори документів або текстові процесори,
- Настільні видавництва.

Текстові редактори представляють собою системи обробки тексту, які для внутрішнього представлення тексту використовують лише стандартні коди, такі як кінець рядка, перевід каретки та кінець сторінки. Типовий текстовий редактор з мінімальним набором функцій дозволяє вводити, редагувати і зберігати будь-який символний текст. Зазвичай такі редактори використовуються для створення текстів програм, і вони забезпечують виконання різноманітних функцій [8]:

- Введення тексту з клавіатури;
- Редагування тексту;
- Видалення, копіювання або переміщення частин тексту;
- Вставка фрагментів тексту з інших файлів;
- Пошук усіх входжень заданої групи символів (контексту);
- Заміна одних груп символів (контексту) іншими;
- Збереження набраного тексту у файлі;

– Друк тексту на різних типах принтерів.

До групи програм для створення текстових документів відносяться такі програми, як *Norton Editor* від *Peter Norton Computing Inc.*, *SideKick* від *Borland*, *Brief* від *Solution Systems*, а також багатофункціональний редактор *Multi-Edit*, розроблений *American Cybernetic Inc*, і редактор Блокнот (*Notepad*), який включений до операційних систем *Windows 9x*. Серед цього класу програм також визначаються редактори Турбо середовищ – інтегровані засоби для створення, компіляції, відлагодження і виконання програм на мовах програмування, таких як Бейсік, Паскаль, Сі, Асемблер. Невід'ємною частиною цих програм є редактори з вищезазначеними можливостями для обробки текстів програм.

Нові текстові процесори зменшили область використання редакторів текстів, але не повністю їх витіснили. Редактори текстів застосовуються тоді, коли:

1. Документ має суто інформативний характер і не потребує спеціального форматування.

2. Документ містить довідкову інформацію, яка повинна займати мінімальний обсяг на носії з обмеженою ємністю, наприклад, файли на інсталяційних лазерних дисках із інструкціями по встановленню програмного забезпечення.

3. Документ має дуже великий обсяг, і використання редакторів текстів дозволяє забезпечити мінімальний обсяг файлу для такого документу [8].

Текстовий документ та його об'єкти



Рис. 1.8. Текстовий документ

Графічні документи.

Графічні документи включають в себе різноманітні матеріали, що мають графічну або візуальну форму. Це можуть бути малюнки, схеми, діаграми, картини, фотографії, карти, графіки, візуалізації та інші зображення. Ось деякі типові приклади графічних документів:

1. Малюнки та Ілюстрації:

Художні роботи та малюнки: Матеріали, що мають мистецьке значення.

Технічні схеми та креслення: Малюнки, які використовуються для пояснення конструкцій, деталей, алгоритмів тощо.

2. Фотографії:

Зображення подій та місць: Фотографії, які фіксують моменти часу та події.

Наукові фотографії: Матеріали, що використовуються для ілюстрації наукових досліджень, експериментів тощо.

3. Діаграми та Графіки:

Статистичні графіки: Інформаційні графіки, які відображають числові дані.

Організаційні схеми: Діаграми, які представляють структуру організацій, процесів або взаємозв'язків.

4. Карти:

Географічні карти: Матеріали, що відображають географічні регіони, місцевості, маршрути тощо.

Тематичні карти: Карти, які вказують на розподіл певних характеристик, таких як населення, клімат, економіка тощо.

5. Презентації та Візуалізації:

Слайди та презентації: Візуальні матеріали, що використовуються для передачі інформації під час презентацій.

Інфографіка: Компактні графічні представлення даних або інформації.

Графічні документи є важливим інструментом для передачі інформації, і вони можуть бути використані в різних галузях, включаючи мистецтво, науку, бізнес, освіту та багато інших. Вони допомагають візуалізувати складні концепції, роблять інформацію більш доступною та зрозумілою для широкої аудиторії.

1. Креслення деталі представляє собою документ, який має зображення конкретної деталі та інші дані, необхідні для її виготовлення і контролю.
2. Складальне креслення (СК) – це документ, що включає в себе зображення виробу та інші важливі дані, необхідні для зборки (виробництва) та контролю виробу.
3. Креслення загального вигляду (ВЗ) – це документ, який визначає конструкцію виробу, взаємодію його основних частин і пояснює принцип його роботи.
4. Схема – це документ, на якому умовними зображеннями показані складові частини виробу і зв'язки між ними.
5. Монтаже креслення – це документ, який включає просте зображення виробу, предмети, які використовуються при монтажі, повне або часткове зображення пристрою, до якого виріб кріпиться (включаючи конструкцію фундаменту), установчі та приєднувальні розміри, а також технічні вимоги до процесу монтажу виробу.

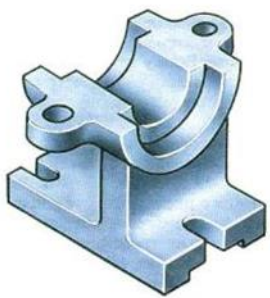
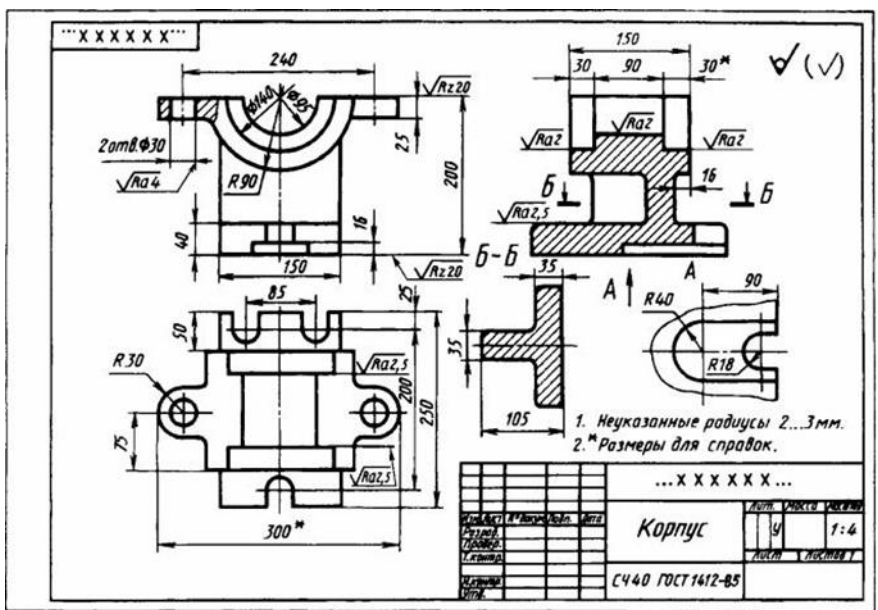


Рис. 1.9. Графічний документ

Фотографії та ілюстрації

Фотографії та ілюстрації є важливою категорією архівних документів, які відображають візуальну інформацію. Ці документи можуть представляти

різноманітні аспекти історії, культури, науки та інших сфер. Ось деякі розглядані аспекти фотографій та ілюстрацій в архівах:

1. Документація подій: Фотографії можуть фіксувати важливі події, такі як публічні заходи, церемонії, конференції, демонстрації і т.д. Ці зображення можуть слугувати як важливий історичний матеріал.

2. Портрети: Фотографії осіб, які можуть включати політичних лідерів, вчених, артистів чи інших видатних особистостей. Ці портрети можуть бути важливим джерелом для вивчення історії та культури.

3. Фотодокументація об'єктів та місць: Зображення будівель, місцевостей, архітектурних споруд, пам'яток тощо. Це може бути важливим для відтворення змін у середовищі та розвитку інфраструктури.

4. Ілюстрації для навчальних та наукових цілей: Малюнки, діаграми, схеми, які використовуються для пояснення наукових концепцій або ілюстрації навчального матеріалу.

5. Мистецькі роботи: Ілюстрації та мистецькі твори, які відображають творчість і культурне надбання.

6. Рекламні матеріали: Фотографії та ілюстрації, використовувані для реклами продуктів, послуг чи подій.

Зберігання фотографій та ілюстрацій в архівах дозволяє зберігати важливі візуальні аспекти минулого для майбутніх поколінь та дослідників. Це також може використовуватися для створення виставок, вивчення еволюції культурного середовища та розкриття різних аспектів історії.



Рис. 1.10. Фотографії та ілюстрації

Аудіо та відеозаписи

Аудіо та відеозаписи є важливими формами архівних документів, оскільки вони зберігають звукову і зображувальну інформацію. Ці документи можуть бути цінним ресурсом для розуміння історії, культури, науки та інших аспектів суспільства. Ось деякі роздуми про аудіо та відеозаписи в архівах:

1. Оральна історія: Аудіозаписи інтерв'ю, свідчень та розповідей можуть зберігати особисті досвіди та свідчення людей, що мали важливий вплив на події або суспільство.

2. Документація подій: Відеозаписи можуть фіксувати важливі події, такі як конференції, концерти, спортивні змагання, політичні заходи та інші.

3. Записи лекцій і презентацій: Аудіозаписи лекцій відомих осіб або вчених можуть бути корисним ресурсом для навчальних та дослідницьких цілей.

4. Музичні записи: Записи виконання музичних творів, концертів та інших музичних подій.

5. Документація мовної та звукової культури: Записи мов, діалектів, розмов та інших аспектів звукової культури.

6. Спогади та інтерв'ю: Аудіозаписи спогадів індивідів, які можуть мати історичне значення.

7. Документація наукових досліджень: Відеозаписи експериментів, обговорень та інших наукових заходів.

8. Архіви теле- та радіоemisій: Зберігання та доступ до аудіо та відеоматеріалів, які були транслювані по телебаченню чи радіо.

Зберігання аудіо та відеозаписів в архівах дозволяє зберегти важливі аспекти звукового та візуального спадку для майбутніх поколінь і дослідників. Архіви можуть використовуватися для вивчення розвитку культури, аналізу мовних змін, дослідження історичних подій та багато іншого.



Рис. 1.11. Аудіо та відеозаписи

6. За вмістом

Особисті документи

Особисті офіційні документи використовуються громадянами для вираження своїх пропозицій, прохань, скарг, а також для передачі повноважень іншій особі щодо отримання різних документів, грошей, товарів та інших послуг. Вони також закріплюють права особи перед державою, установами (підприємствами) та іншими особами.

Ця категорія документів виконує різні мовні та управлінські функції. Серед них інформаційна, комунікативна та кумулятивна, а також спеціальні управлінські функції, такі як регулювання, правове закріплення та статистична функція. Особисті офіційні документи підтверджують ідентифікацію особи та її права, обов'язки і соціальний статус. Вони часто містять особливу інформацію, таку як повне прізвище, ім'я та по батькові. Деякі з них також включають фотографію власника, наприклад, паспорт, студентський квиток, службове посвідчення чи перепустка.

Для висловлення звернень до керівництва підприємства або державних установ створюються особисті документи, які складаються самим автором, який є укладачем таких документів. Серед цих документів можна виділити:

- Заява;

- Резюме;
- Автобіографія;
- Характеристика;
- Розписка;
- Доручення.

Хоча текст цих документів може бути різним і формується автором, обов'язкові реквізити, такі як назва виду документа, дата та підпис, повинні бути включені [10].

Заява є документом, у якому міститься прохання чи пропозиція від одного чи кількох осіб і адресована установі або посадовій особі. Структура заяви включає реквізити, такі як адреса (назва організації або посадової особи), відомості про заявника, назву виду документа, текст, перелік додатків, підпис автора та дату написання заяви.

Резюме представляє собою документ, який містить короткі біографічні відомості і представляється особою під час пошуку роботи на комерційному підприємстві. У резюме вказуються реквізити, такі як назва виду документа, дата, прізвище, ім'я, по батькові, посада, поштова адреса претендента, телефон, текст, підпис.

Автобіографія представляє собою документ, в якому особа, що його складає, надає опис свого життя і діяльності. Відмінністю від типового формуляру автобіографія не обмежена стандартами і може бути створена від руки на аркуші чистого паперу або спеціальному бланку, зазвичай при вступі на роботу або у інших контекстах.

У автобіографіях обов'язково вказують основні деталі, такі як прізвище, ім'я, по батькові, дату і місце народження, інформацію про освіту (повні найменування навчальних закладів), армійську службу, трудову діяльність (зазначені у хронологічному порядку), громадську роботу, отримані нагороди, сімейний стан та склад родини, останнє місце роботи і посаду, домашню адресу, дату написання та особистий підпис.

Розписка – це документ, що підтверджує передачу та отримання різних матеріальних цінностей, грошей, товарів іншого роду, який підтверджується

підписом одержувача. Реквізити розписки включають назву документа, посаду, прізвище, ім'я та по батькові видачі розписки, інформацію про одержувача, конкретний предмет, за який видається розписка, підставу для отримання цих цінностей, дату та підпис одержувача.

Доручення – це документ, що надає конкретній особі право виконувати певні дії від імені підприємства чи окремого громадянина.

Заповіт – це волевиявлення особи (заповідача) щодо розпорядження своїм майном у випадку смерті.

Посвідка – це документ, який свідчить про факти діяльності чи повноваження громадян.

Посвідчення – це документ, який також свідчить про факти діяльності або повноваження громадян.

Посвідчення особи – це кадровий документ, який підтверджує факт роботи конкретної особи в певній установі, і має реквізити, такі як назва установи, вид документа, дані про одержувача (прізвище, ім'я, по батькові), строк дії, підпис керівника, гербова печатка, дата видачі, а також фотографія одержувача [10].

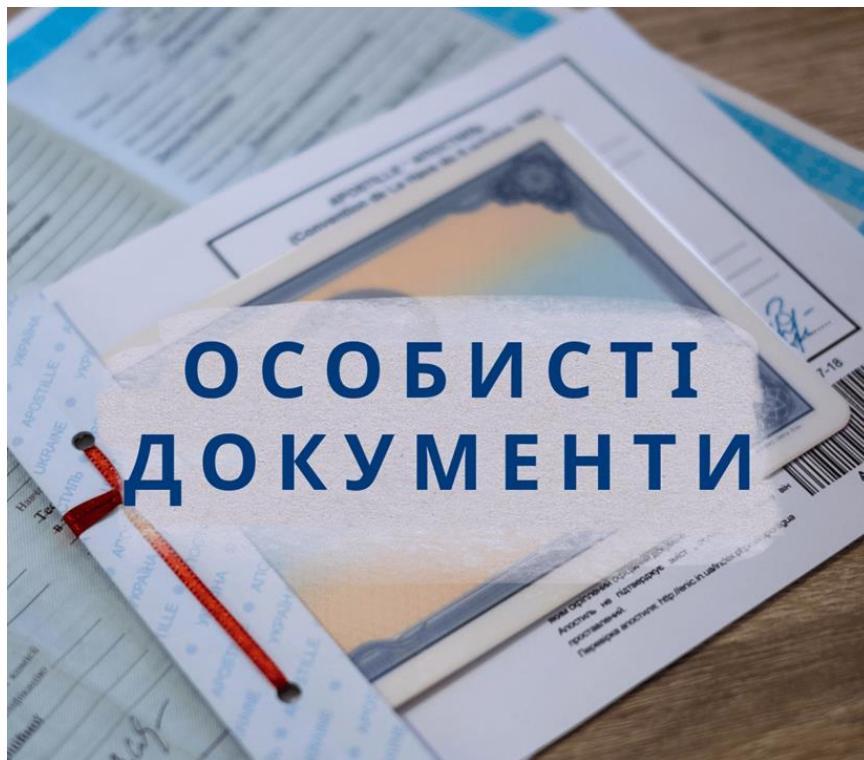


Рис. 1.12. Особисті документи

Документи організацій

Документи організацій – це категорія архівних документів, яка охоплює різноманітні матеріали, створені та утримувані різними видами організацій, включаючи підприємства, установи, урядові агентства, громадські організації та інші.

Ці документи можуть включати такі елементи:

1. Статутні та організаційні документи:

Статути та установчі документи: Офіційні документи, що визначають цілі, структуру та правила діяльності організації.

Положення та правила: Документи, які регулюють внутрішню організацію та функціонування організації.

2. Документи управління:

Протоколи засідань та звіти: Записи вирішень та обговорень, що відбуваються на засіданнях органів управління.

Звіти керівництва: Інформація про результати роботи та фінансовий стан організації.

3. Фінансові документи:

Бухгалтерські звіти: Фінансова звітність, яка відображає фінансовий стан організації.

Рахунки та фінансові документи: Документи, пов'язані з фінансовою діяльністю, включаючи рахунки, квитанції тощо.

4. Документи персоналу:

Кадрові документи: Заяви, трудові контракти, кадрові картки тощо.

Політики та процедури персоналу: Документи, що визначають правила та вимоги стосовно управління персоналом.

5. Кореспонденція:

Листи, електронні повідомлення, меморандуми: Документи, що стосуються комунікації в межах організації та зовнішньої комунікації.

6. Юридичні документи:

Договори та угоди: Офіційні документи, що регулюють правові зобов'язання організації.

Судові рішення: Інформація про справи, в яких організація брала участь.

Ці документи грають ключову роль у внутрішньому управлінні, документуванні діяльності та забезпеченні відповідності правовим та фінансовим стандартам. Вони також можуть бути важливим джерелом інформації для досліджень історії та функціонування організацій у різних сферах.



Рис. 1.13. Документи організацій

Документи юридичної природи

Документи юридичної природи включають в себе різноманітні правові матеріали, що регулюють діяльність осіб, організацій та установ. Ці документи мають юридичний статус і можуть включати різні види договорів, рішень суду, правових актів та інших документів. Ось деякі типові приклади документів юридичної природи:

1. Договори та Угоди:

Трудові контракти: Угоди між роботодавцем і працівником, що визначають умови праці.

Комерційні договори: Угоди між підприємствами або фізичними особами, що регулюють комерційні відносини.

Орендні та лізингові угоди: Документи, що встановлюють умови надання в оренду або лізингу майна чи послуг.

2. Судові Рішення та Рішення Арбітражу:

Рішення суду: Офіційні документи, які визначають вироки та вирішення судових справ.

Рішення арбітражу: Вирішення спорів за допомогою арбітражного суду або арбітражного рішення.

3. Законодавчі та Нормативні Акти:

Закони: Офіційні документи, прийняті національними чи регіональними законодавчими органами.

Розпорядження та регуляторні акти: Документи, видані урядовими органами, що визначають правила та регуляції.

4. Документи про Інтелектуальну Власність:

Патенти: Документи, які надають права на використання та захист винаходів.

Авторські права: Документи, що визначають права на літературні, художні та наукові твори.

5. Нотаріальні акти:

Нотаріальні посвідчення: Офіційні документи, які отримують правову вагу внаслідок посвідчення нотаріуса.

6. Документи стосовно Захисту Персональних Даних:

Політики конфіденційності: Документи, що визначають правила збору, зберігання та обробки персональних даних.

Ці документи важливі для визначення правового статусу та обов'язків сторін, а також для вирішення правових питань та спорів. Зберігання юридичних документів у архівах дозволяє забезпечити доступ до юридичної інформації та забезпечити її юридичну обов'язковість.



Рис. 1.14. Документи юридичної природи

7. За терміном зберігання

Тимчасовий документ

"Тимчасовий документ" може вказувати на різні види документів, які мають тимчасовий характер або тимчасову чинність. Залежно від контексту він може включати в себе різні інтерпретації. Ось кілька можливих розумінь:

1. Тимчасовий правовий документ: Може вказувати на документ із тимчасовою дією або тимчасовим статусом, такий як тимчасова угода, тимчасовий контракт або тимчасове рішення.

2. Тимчасовий документ у сфері медицини: Може вказувати на тимчасовий медичний документ, такий як тимчасова медична картка або тимчасовий рецепт.

3. Тимчасовий документ у сфері освіти: Може вказувати на тимчасовий академічний документ, наприклад, тимчасовий студентський квиток або тимчасовий академічний запис.

4. Тимчасовий документ у сфері бізнесу: Може вказувати на тимчасовий бізнес-документ, такий як тимчасовий контракт, тимчасовий ліцензійний договір або тимчасове партнерство.



Рис. 1.15. Тимчасові документи

Документи сталих архівів

Документи сталих архівів – це матеріали, які зберігаються на постійній основі через їхню важливість для історії, культури, науки, або через їхню унікальність та цінність. Ці документи мають велике історичне значення та призначені для зберігання на тривалий термін. Ось деякі приклади документів сталих архівів:

1. Історичні Документи:

Договори міжнародного значення: Наприклад, мирні угоди, дипломатичні угоди, документи, що регулюють стосунки між державами.

Декларації та Конституції: Офіційні документи, що визначають правові основи функціонування держави чи організації.

Історичні листи та кореспонденціям від венечных особистостей: Наприклад, листи від політичних лідерів, видатних вчених, митців.

2. Архівні Фотографії та Відеоматеріали:

Зображення та записи важливих подій: Фотографії та відеозаписи, які фіксують ключові моменти історії.

Документація культурної спадщини: Зображення та відеозаписи архітектурних пам'яток, традицій, мистецтва тощо.

3. Наукові та Технічні Документи:

Наукові дослідження та відкриття: Публікації та документація важливих наукових відкриттів.

Технічні креслення та патенти: Документація, що описує нові технології та винаходи.

4. Документи Культурної Ідентичності:

Літературні та художні твори: Манускрипти, рукописи, твори мистецтва.

Традиції та фольклор: Збірки народних оповідань, пісень, звичаїв.

5. Фільми та Звукові Записи:

Історичні кінопроекти та документальні фільми: Зберігання фільмів та записів, що відображають історію і культуру.

Аудіозаписи культурних подій: Записи виступів, музичних концертів, радіопередач.

Ці документи є важливим елементом збереження історичної пам'яті та надають можливість майбутнім поколінням досліджувати та розуміти минуле. Зазвичай їх зберігають в спеціальних архівах та бібліотеках для забезпечення їхньої консервації та доступу.



Рис. 1.16. Документи сталих архівів

4. Відкриті документи

Термін "відкриті документи" може мати кілька трактувань, залежно від контексту. Нижче подано декілька можливих розумінь:

1. Документи у відкритому доступі: Це може вказувати на документи, які загальнодоступні для перегляду та використання, зазвичай через відкриті архіви, бібліотеки або інтернет-ресурси. Такі документи можуть бути доступні всім без обмежень або з обмеженим доступом для перегляду.

2. Документи, які ще не закриті або не завершені: Це може вказувати на робочі матеріали, які ще не завершені, але вже відкриті для перегляду або рецензування. Такі документи можуть бути частиною робочих процесів або проектів.

3. Відкриті формати файлів: В іншому контексті це може вказувати на формати файлів, які відкриті для використання і редагування різними програмами, не обмеженими пропрієтарними обмеженнями.

4. Відкриті дані: Також це може вказувати на ініціативу з публікації відкритих даних, де урядові або організації роблять дані доступними для громадськості з метою збільшення прозорості та інновацій.



Рис. 1.17. Відкриті документи

Закриті документи

Закриті документи – це ті, до яких доступ обмежений або обмежений усередині організації чи спеціального архіву. Це може стосуватися чутливої інформації, яка має конфіденційний чи обмежений характер. Такі документи можуть бути класифіковані як внутрішні, таємні, конфіденційні або засекречені, залежно від характеру інформації та вимог безпеки. Ось деякі приклади закритих документів:

1. Таємниці оборони:

Військові плани та стратегії: Документи, що містять інформацію про військові операції та оборонні стратегії.

2. Державна безпека:

Інформація про розвідку та контррозвідку: Документи, що містять інформацію про діяльність розвідувальних служб та заходи безпеки держави.

3. Комерційна та Патентна Інформація:

Комерційні відомості та торгові таємниці: Інформація про стратегії компаній, фінансові плани, винаходи, які можуть бути обмежені з питань конкуренції.

4. Медична конфіденційність:

Медичні записи та інформація про пацієнтів: Документи, що містять особисту та конфіденційну медичну інформацію.

5. Юридична приватність:

Документи юридичних справ: Конфіденційна юридична інформація, яка може бути обмежена з питань приватності клієнтів чи бізнес-партнерів.

6. Особиста інформація:

Персональні досьє: Інформація про особисте життя чи професійну історію конкретних осіб.

Зберігання та обробка закритих документів вимагає дотримання строгих процедур безпеки та контролю доступу, щоб забезпечити конфіденційність та уникнути несанкціонованого розголошення чутливої інформації.



Рис. 1.18. Закриті документи

5. За функцією

Фінансові документи

Обліково-фінансові документи виконують різноманітні функції, виступаючи не лише як носії інформації для фінансового обліку, але й як інструмент, що сприяє господарській діяльності підприємства. Економічна значущість документів проявляється в їхній ролі джерела для контролю та аналізу. Іншими словами, документи виконують контрольно-аналітичну функцію, забезпечуючи відстеження правильності проведення операцій і контроль за дотриманням платіжної, розрахункової та фінансової дисципліни. Вони також допомагають контролювати ефективно та раціональне використання матеріальних, трудових і фінансових ресурсів. Документи є інструментом для розгортання різнобічного аналізу господарської діяльності, включаючи використання земельних ресурсів, управління основними та оборотними активами, а також виявлення причин негативних явищ, таких як зниження виробництва чи урожайності [11].

Правильно оформлені первинні документи служать письмовим доказом проведених господарських операцій і можуть використовуватися судовими органами

при розгляді справ з економічних питань. Судові органи признають документи юридично значущими, якщо вони були складені своєчасно та належним чином.

Документи виявляють вирішальне значення для фінансового контролю, аудиту та документальних ревізій. Окрім паперових форм, широко застосовуються також електронні носії облікової інформації, такі як магнітні диски, стрічки, дискети і інші [11].



Рис. 1.19. Фінансові документи

Наукові документи

Наукові документи представляють собою спеціальну категорію документів, які містять інформацію про наукові дослідження, відкриття, аналіз та іншу наукову інформацію. Ці документи є основою для розвитку науки та сприяють обміну знаннями серед науковців та спеціалістів у різних галузях. Ось деякі типові наукові документи:

1. Наукові Статті:

Статті у наукових журналах: Детальні описи наукових досліджень, експериментів та відкриттів, що підлягають публікації та рецензуванню.

2. Наукові Дослідження та Репортажі:

Дослідження та звіти: Документи, що включають опис методології, результатів та висновків наукових досліджень.

3. Дисертації та Тези:

Докторські дисертації та магістерські тези: Об'ємні наукові твори, що представляють результати довготривалих наукових досліджень.

4. Конференційні Звіти:

Звіти з наукових конференцій та симпозіумів: Документи, що включають презентації та обговорення наукових досліджень.

5. Наукові Книги:

Монографії та наукові посібники: Книги, що висвітлюють результати досліджень та докладно розглядають певні теми.

6. Технічні Звіти та Документація:

Технічні звіти і документація про винаходи: Інформація щодо нових технологій, винаходів та патентів.

7. Академічні Публікації:

Академічні книги та журнали: Видання, які представляють різні аспекти наукових вивчень та досліджень.

Ці документи грають ключову роль у розвитку науки та розумінні світу, допомагаючи науковцям обмінюватися знаннями та вдосконалювати наукові методи та теорії. Багато з цих матеріалів піддаються рецензуванню та публікації, щоб забезпечити високий рівень якості та достовірності наукової інформації.



Рис. 1.20. Наукові документи

Висновки до розділу 1

У першому розділі кваліфікаційної роботи було розглянуто основні поняття, які було досліджено та проаналізовано згідно даної теми.

Дізналися, які є переваги та недоліки оцифрування документів одні з них це збереження коштів на папері та друкарській фарбі, оскільки документи архівуються у цифровому форматі, зменшення використання паперу призводить до зменшення утворення відходів, що сприяє екологічній діяльності.

Ще є різні архівні документи, які використовуються для оцифрування документів.

РОЗДІЛ 2

АНАЛІЗ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ

2.1. Аналіз програмного забезпечення для обробки текстової частини

Microsoft Word є професійним текстовим редактором, спроектованим для оформлення, компоновання та структурування текстового вмісту в розширеному програмному середовищі. Це невід'ємний інструмент для створення різноманітних документів, таких як доповіді, лекції, дипломи, реферати та інший контент, що вимагає введення тексту з клавіатури.

Різнманітні теми і шаблони надають можливість стилізувати та якісно оформити ваш документ, а бібліотека макетів забезпечує відповідне форматування в межах визначених стандартів. Особливу увагу слід приділити вбудованому плагіну для перевірки орфографії тексту на будь-якій мові, а також виділенню помилок у синтаксисі, пунктуації та граматиці термінів, фраз і словосполучень у вашому документі [12].

Продукт також має модуль для обробки поштової кореспонденції, дозволяючи вибрати одержувача та тип документа (конверти, наклейки, листи або каталог). Вбудовані режими перегляду документів, такі як режим читання, розмітка вмісту та веб-документ, роблять роботу з документами більш зручною та ефективною.

Microsoft Word надає розширений набір можливостей і привілеїв, які дозволяють користувачам максимально розширити та налаштувати функціонал продукту:

- Створення макросів, які можуть значно розширити функціональний спектр продукту. Вони можуть бути активовані натисканням кнопок миші або гарячою комбінацією клавіш.

- Індивідуальна настроювання орієнтації сторінки, розміру, розривів, відступів та інтервалів. Включає можливість додавання колонок, номерів рядків,

переносів, малюнків, фігур, знімків екрану, зображень з Інтернету, значків, приміток та гіперпосилань.

- Інтелектуальна обробка змісту, виноска та посилань на інші джерела, яка автоматично враховує існуючі документи та формує фінальну версію посилання на активному аркуші.
- Вбудовані засоби для спільної роботи та рецензування файлів, де автори та рецензенти можуть додавати примітки, коригувати вміст та захищати аркуш.
- Завдання індивідуального масштабування документів *Word* з будь-яким рівнем деталізації.
- Інтерактивний посібник, спрямований на полегшення первинного ознайомлення з програмою, використовуючи інтегрований анімований довідник по основних операціях, що часто виконуються в *Word* [12].



Рис. 2.1. Редактор *Word*

Microsoft Word має своє коріння в текстовому процесорі *Bravo*, який має оригінальний графічний інтерфейс і був розроблений в дослідницькому центрі «*Xerox PARC*». Творець *Bravo*, Чарльз Симоні (*Charles Simonyi*), покинув *PARC* у 1981 році, а пізніше того ж року перейшов до роботи з Броді, з яким він спільно працював над процесором *Bravo*.

Перший реліз *Word* для *MS-DOS* відбувся в кінці 1983 року і став об'єктом невеликого інтересу через наявність конкуруючого продукту – *WordPerfect*. Проте версія для *Macintosh*, випущена в 1985 році, отримала широкий розповсюдження. За два роки «*Word 3.01* для *Macintosh*» покращив свої позиції (версія 3.0 була нестабільною і швидко замінена).

Хоча *MS-DOS* була текстовою операційною системою, *Word* для *DOS* став першим текстовим процесором для *IBM PC*, який міг відображати розмітку тексту (наприклад, напівжирний або курсивний текст) під час редагування. Хоча це не був абсолютно повноцінний *WYSIWYG*-редактор, він все ж був значним кроком вперед порівняно з іншими текстовими процесорами, такими як *WordStar* і *WordPerfect*, які використовували простий текстовий екран з кодами розмітки.

Однак у програмному забезпеченні для *DOS* загальноприйняті були власні комбінації "гарячих клавіш" для кожної команди, які часто були складні для запам'ятовування (наприклад, в *Word* для *DOS*, для збереження файлу використовувалася комбінація *ESC-T-S*). Більшість секретарів були звиклі користуватися тільки *WordPerfect*, тому компанії неохоче переходили на конкуруючі продукти, які мали обмежені переваги.

Перша версія *Word* для *Windows*, випущена у 1989 році за ціною 500 доларів США, відображала обраний Microsoft напрямок розвитку. Вона взяла багато елементів від *Macintosh* і використовувала стандартні клавіатурні скорочення, такі як *CTRL-S* для збереження файлу. Після випуску *Windows 3.0* наступного року продажі зросли, оскільки *Word 1.0* працював краще з *Windows 3.0*, ніж з попередніми версіями *Windows/386* і *Windows/286*. Головний конкурент, *WordPerfect*, не зміг швидко випустити робочу версію під *Windows*, що стало вирішальною помилкою. Версія 2.0 встановила *WinWord* на позицію лідера на ринку.

У *Word* для *Macintosh* завжди було визнано відсутність справжніх конкурентів, навіть не зважаючи на наявність програм, таких як *Nisus*, яка дозволяла виділяти декілька незв'язних шматків тексту (ця можливість з'явилася лише у *Word 2002* під час виходу *Office XP*). Багато користувачів не вбачали кардинальних відмінностей

між ранніми версіями, такими як 3.01 (випущена в 1987 році) і версією 5.0 (випущена в 1991 році).

Проте версія 6.0 для *Macintosh*, представлена у 1994 році, була зустрінута деякою скептичністю. Це була перша версія, яка майже повністю збігалася за кодом ядра з версією для *Windows*. Щоб уникнути плутанини, версія під *Windows* після 2.0 також отримала номер 6.0.

Подальші версії поступово розширювали можливості, переходячи за межі звичайного текстового процесора. Інструменти малювання дозволили виконувати прості операції верстки, такі як вставка графіки в документ, хоча для цих завдань зазвичай використовувалися спеціалізовані програми для верстки. Введення об'єктів, порівняння версій документа, підтримка багатьох мов та інші функції були додані в наступні роки.

На сьогодні *Microsoft Word* є найпопулярнішим текстовим процесором, що призвело до того, що його закритий формат документа став стандартом. Файлове розширення «*.doc*» на платформі *IBM PC* тепер вважається синонімом двійкового формату *Word 97 – 2000*. Багато конкуруючих програм підтримують сумісність з цим форматом, завдяки наявності фільтрів експорту та імпорту. Велика частина інформації, необхідної для роботи з цим форматом, отримується шляхом зворотного інжинірингу, оскільки значна частина не є відкритою для загального доступу. Структура документа може відрізнятися в різних версіях *Word*, і відмінності можуть бути досить суттєвими. Форматування, яке виглядає нормально в останній версії, може виглядати по-іншому в старих версіях програми через відсутність зворотної сумісності.

Microsoft опублікувала специфікацію двійкового формату файлів *Word 97-2007* у лютому 2008 року. У *Word 2003* використовується власний *XML*-формат, що використовує публічно документовану схему, відому як *WordprocessingML*, доступну для всіх редакцій *Word 2003*. *Apache Jakarta POI*, написана на *Java*, цілком має надання можливості читання та запису в бінарний формат файлів *Microsoft Word*.

У пакеті *Office 2007 Microsoft* впроваджує формат файлів *Office Open XML* для всіх своїх офісних додатків, який був стандартизований *ISO* у квітні 2008 року.

Як і інші програми в *Microsoft Office*, *Word* може розширювати свої можливості за допомогою вбудованої макромови. З початку використовувався *WordBasic*, але починаючи з версії *Word 97*, використовується *VBA – Visual Basic for Applications*. Проте це може створювати ризики для вбудованих в документи вірусів, відомих як "макровіруси". Наприклад, епідемія хробака *Melissa* стала наочним прикладом цього. У зв'язку з цим багато експертів рекомендують завжди встановлювати високий рівень налаштувань безпеки при використанні *Word* та використовувати антивірусне програмне забезпечення. Перший вірус, який заражав документи *Microsoft Word*, був *DMV*, створений у грудні 1994 року Дж. Макнамара для демонстрації можливостей створення макровірусів. Першим вірусом, що викликав першу світову епідемію макровірусів в липні-серпні 1995 року, був *Concept*.

У серпні 2009 року суд штату Техас заборонив продаж програми *Word* на території США, оскільки *Microsoft* незаконно використовувала метод читання *XML*-файлів, патент на який належав канадській компанії *i4i* [13].

Формати файлів

Хоча розширення «*.doc*» використане в багатьох різних версіях *Word*, насправді мова йде про чотири різних формати файлів:

- *Word for DOS*
- *Word for Windows 1 та 2; Word for Mac 4 та 5*
- *Word 6 та Word 95; Word 6 for Mac*
- *Word 97, 2000, 2002, 2003 та 2007; Word 98, 2001, X та 2004 for Mac*

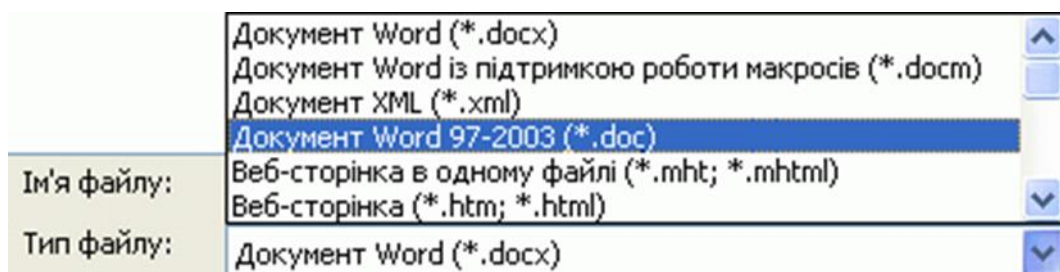


Рис. 2.2. Формати файлів

Розширення "*docx*" вказує на формат *Office Open XML* і використовується в програмах *Word 2007* для *Windows* та *Word 2008* для *Mac*. *Word* не може працювати з документами у форматі *OpenDocument* без конвертації.

Microsoft не гарантує правильне відображення документів на різних комп'ютерах, навіть якщо вони використовують однакові версії *Microsoft Word*. Це означає, що документ, створений в одному місці, може виглядати по-різному у іншого користувача.

У *Microsoft Word* можна встановити три типи паролів: для відкриття, для змін та для внесення приміток і виправлень (для версій 2003 і старших). У нових версіях *MS Word 2007/2010* цей тип паролю називається "Пароль для обмеження форматування і редагування" [13].

Незалежно від версій *MS Word*, "пароль дозволу запису" і "пароль на внесення приміток і виправлень" можна видалити з документа миттєво. Ці паролі призначені не тільки для захисту документу, але й для забезпечення спільної роботи над ним. Основний захист документу від несанкціонованого доступу забезпечує "пароль для відкривання документа".

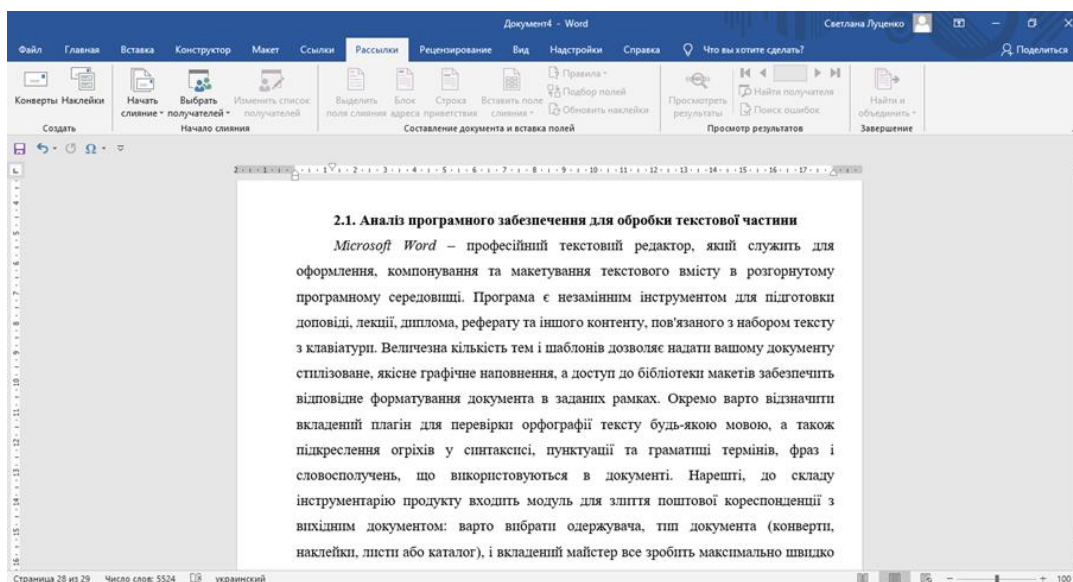


Рис. 2.3. Текстовий редактор *Microsoft Word*

2.2. Програма для сканування документів

ABBYY Scan Station надає можливість виконати сканування сотень сторінок та зберегти результат за кілька хвилин, забезпечуючи широкий спектр можливостей для виправлення зображень. Програму можна використовувати як для сканування невеликих обсягів документів звичайним сканером, так і для перетворення обширних паперових архівів у електронний формат в режимі потоку. Крім того, рішення *ABBYY* дозволяє підприємствам з великою мережею філій налагоджувати розподілене сканування документів з подальшою централізованою обробкою, причому станції сканування можуть розташовуватися в різних містах.

Програму можна використовувати як особисту станцію сканування або в поєднанні з професійними сканерами. Вона зберігає відскановані документи та фотографії у єдиному вікні, що дозволяє використовувати функції експорту у формати *PDF*, *TIFF*, *JPEG*, *PNG* та інші. Стандартні засоби сканування зберігають кожен сторінку як окремий файл, а *ABBYY Scan Station* дозволяє налаштувати процес сканування багатосторінкових документів, книг або журналів. Програма також включає інструменти корекції матеріалу, що сканується, дозволяючи виконувати різноманітні маніпуляції: обрізання, обертання, затирання даних, заливання та інші [15].

Список можливостей та переваг *ABBYY Scan Station* включає:

- Підтримку домашніх БФП та сканерів;
- Роботу із промисловим обладнанням;
- Наявність лікування *ABBYY Scan Station*, що усуває потребу в використанні "crack";
- Підтримку роботи із багатосторінковими документами;
- Збереження результатів у форматах *PDF*, *PNG*, *JPG*, *TIFF* та інших популярних;
- Інструменти базової корекції зображень;
- Створення профілів із налаштуваннями сканування.

Переваги для користувача: При оцінці зручності для користувачів та компаній, програма *ABBYY Scan Station* відзначається такими можливостями:

- Ручне індексування пакетів та документів;
- Автоматична обробка документів;
- Розподіл пакетів на окремі документи та видалення листів-розділювачів;
- Видалення експортованих пакетів;
- Єдиний інтерфейс для роботи з різними типами сканерів в компанії;
- Сумісність із багатьма типами сканерів, незалежно від їхнього рівня.

Щоб полегшити роботу з програмою, результати обробки можна зручно зберігати на комп'ютері або у папці на *FTP*. Це особливо актуально при використанні програми для обробки великої кількості документів за допомогою багатофункціональних сканерів або для щоденного сканування надходження документів в офіс [16].

ABBYY Scan Station використовує інтелектуальний підхід до сканування, роблячи це пакетами. Після сканування програма автоматично розпізнає та розділяє відскановані пакети на окремі документи відповідно до встановлених параметрів, таких як кількість сторінок, порожні сторінки-розділювачі або наявність штрих-кодів. Налаштування обробки зберігаються у профілях, які називаються "типами пакетів". Кожен профіль включає у себе налаштування сканування, обробки зображень, складання та експорту, що дозволяє легко організувати роботу з різними типами документів, такими як договори, рахунки-фактури та журнальні статті.

Таким чином, для кожного типу документів ми налаштовуємо власний профіль:

Профіль 1: договори скануються у чорно-білому вигляді, групуються в документи за допомогою реєстраційних штрих-кодів та зберігаються в системі документообігу;

Профіль 2: рахунки-фактури скануються у чорно-білому вигляді без розділення на окремі документи та відсилаються на розпізнавання у систему *ABBYY FlexiCapture*;

Профіль 3: журнальні статті скануються у кольорі з роздільною здатністю 600 dpi та зберігаються у *SharePoint*. Завдяки цим профілям підготовка програми для обробки документів кожного типу стає доступною за одним натисканням миші.

ABBYY Scan Station включає розгалужений набір інструментів для редагування зображень, таких як обрізка, поворот, заливка, усунення спотворень та інвертування. Це дозволяє оптимізувати обробку документів з метою поліпшення якості розпізнавання (у випадку подальшого застосування розпізнавання) або приведення їх у відповідний формат для зберігання в архіві. Наприклад, використання заливки для частини зображення однотонним кольором дозволяє захвати конфіденційну інформацію, якщо вона міститься в документі [14].

ABBYY Scan Station підтримує більшість моделей документних сканерів та інтерфейси *TWAIN*, *ISIS* та *WIA*. Після сканування програма зберігає відскановані зображення у форматах *PDF*, *TIFF*, *JPEG*, *JPEG2000*, *BMP*, *PNG*, *PCX*, *DCX* та дозволяє їх експортувати до папок на диску.

ABBYY Scan Station також надає єдиний інтерфейс для корпоративних сканерів у випадку, коли в компанії використовуються різні моделі сканерів. Встановлення програми на всіх робочих місцях забезпечить обробку паперових документів з однаковими налаштуваннями, сприяючи створенню уніфікованого електронного архіву.

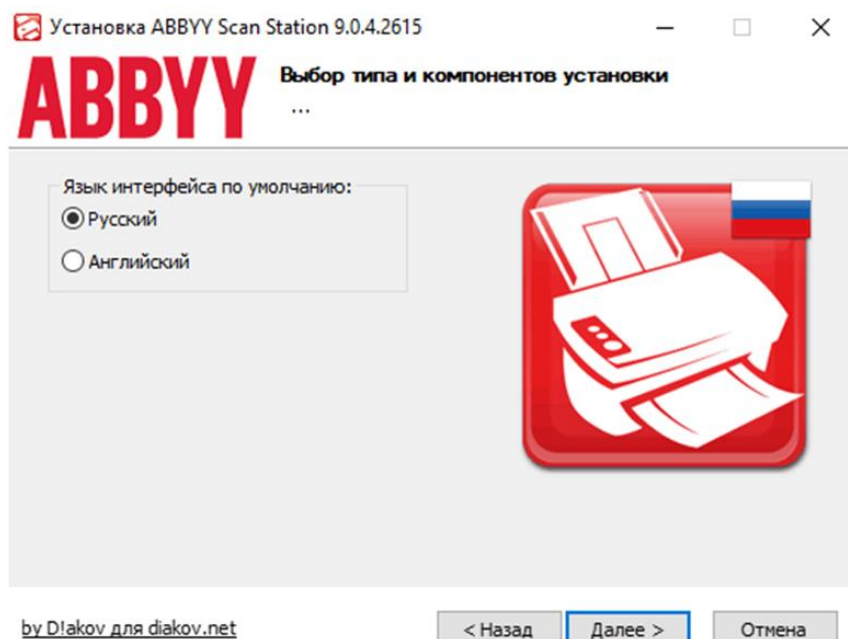


Рис. 2.4. Програма для сканування документів

2.3. Програма для обробки документів

ABBYY Recognition Server є серверним рішенням високої продуктивності для обробки документів, отриманих із факсу, сканера, МФУ або фото, та їх подальшого введення в системи документообігу та зберігання.

Основні переваги цього рішення включають високу продуктивність та безлімітну масштабованість. *ABBYY Recognition Server* здатний обробляти будь-який обсяг документів і успішно використовується в проектах з розпізнавання обсягами понад 1 мільйон сторінок щомісячно.

Програма відзначається простотою встановлення, адміністрування та використання. Після одноразової настройки сценаріїв обробки документів вона автоматично виконує всі завдання. Якщо для обробки документів потрібна участь оператора, програма надає всі необхідні інструменти через зручний та зрозумілий інтерфейс.

ABBYY Recognition Server є зручним рішенням для обробки документів у всій організації. Вона розпізнає зображення, отримані у 190 мовах, точно відтворює структуру вихідного документа та зберігає їх в інформаційну систему підприємства.

ABBYY Recognition Server пропонує гнучкі можливості інтеграції з різними інформаційними системами. Його рішення включає різноманітні інструменти для зручної інтеграції з різними інформаційними системами, додатками, мережевими сканерами та МФУ. Це охоплює *COM-based API* та *Web Service API*, скриптування, а також можливості кастомізації обробки документів за допомогою *XML*-описів.

ABBYY Recognition Server використовує індексування документів, надаючи метадані кожному документу, які відображають його основний зміст, такі як номер документа та дата розпізнавання. Це дозволяє швидко та зручно заповнювати картки документа в електронних архівах та системах електронного документообігу.

Розширені можливості створення архівів документів *ABBYY Recognition Server* включають перетворення зображень у редаговані формати, що полегшує роботу з документами та дозволяє створювати повнотекстовий індекс для пошукових сервісів.

Програма також автоматично створює текстові версії всіх документів, як тих, що надходять до архіву, так і зберігаються в ньому.

ABBYY Recognition Server безпосередньо інтегрується з електронними архівами на базі *Microsoft Office SharePoint*, спрощуючи сценарії обробки документів, що зберігаються в *MS SharePoint*.

Ще однією перевагою є спрощена інтеграція з інформаційними системами, включаючи можливість імпортування груп користувачів з *Active Directory* та налаштування прав для користувачів, а також імпорт списків полів для індексації з зовнішніх файлів або баз даних.

У новому випуску з'явилися додаткові сценарії обробки документів. Особливий режим роботи спрямований на технічні креслення, де графічні об'єкти залишаються незмінними, а текст на кресленнях, будь він вертикальним чи горизонтальним, розпізнається.

Покращені можливості роботи з *PDF*-документами дозволяють зберігати всі характеристики вихідного *PDF*-файлу при імпорті. Це включає закладки, метадані, додатки (зображення, текстові документи і т.д.), а також додається шар із розпізнаним текстом у новий файл [17].

Тепер користувачі можуть швидко переглядати документи у веб-браузері завдяки можливості створювати адаптовані версії документів з мінімальним розміром, придатним для комфортного читання з екрана.

ABBYY Recognition Server став більш відмовостійким і забезпечує швидше, зручніше та простіше відновлення системи.

Впровадження рішення допомагає легко знаходити документи, відповідати корпоративним і державним стандартам, забезпечує єдиний формат для зберігання, ефективно наповнює та нормалізує електронний архів. *ABBYY Recognition Server* автоматично сканує і розпізнає паперові документи, зберігаючи їх в електронному архіві.

Розширені можливості введення документів в режимі "Канцелярія" дозволяють легко сканувати та реєструвати документи в системі з кількох клацань миші.

ABBYY Recognition Server може бути налаштований як сервіс розпізнавання і конвертації документів для робочих груп, співробітників всієї організації або окремих співробітників [18].

Переваги впровадження включають:

- Легкий пошук необхідних документів.
- Гарантія відповідності документів корпоративним і державним стандартам.
- Збереження документів у єдиному форматі.
- Ефективне наповнення та нормалізація електронного архіву.
- Автоматизоване сканування та розпізнавання паперових документів з їх збереженням в електронному архіві за допомогою *ABBYY Recognition Server*.

збереженням в електронному архіві за допомогою *ABBYY Recognition Server*.

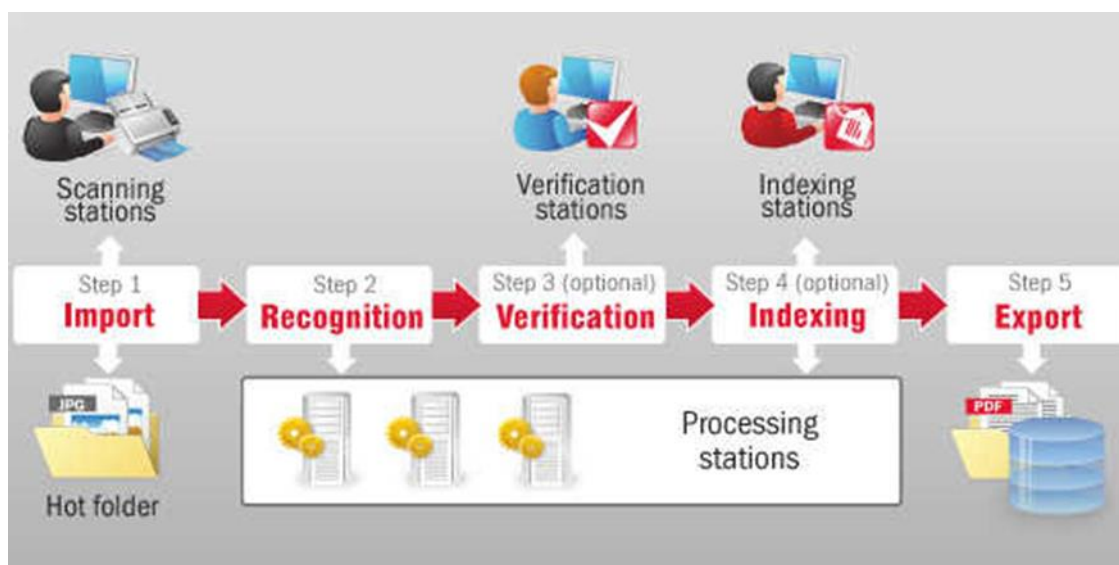


Рис. 2.5. Програма для обробки документів

2.4. Програма для обробки інформації з документів

ABBYY FlexiCapture – це рішення для автоматизованого введення даних та обробки документів. Цей продукт спрощує вилучення інформації з паперових документів та впроваджує збереження даних в інформаційній системі підприємства. *ABBYY FlexiCapture* дозволяє різним організаціям, включаючи великі корпорації, урядові структури та навчальні заклади, автоматизувати процес введення даних до інформаційних систем, що призводить до зниження витрат та підвищення якості обслуговування клієнтів.

ABBYY FlexiCapture служить як єдиний центр обробки всіх видів документів, дозволяючи ефективно вводити різноманітні документи в систему, уникаючи ручного втручання.

Процес автоматизації дозволяє ідентифікувати та обробляти різноманітні типи документів автоматично, без потреби в додаткових модулях чи програмному забезпеченні. *ABBYY FlexiCapture* може самостійно класифікувати надходячі документи, звільняючи користувача від необхідності попереднього сортування або ручного додавання сторінок-роздільників.

Система також відзначається розширюваністю та масштабованістю, що робить її ефективним рішенням для організацій різних розмірів та потреб.

Гнучкість та розширюваність є ключовими характеристиками *ABBYY FlexiCapture*, пропонуючи два варіанти установки: локальний і мережевий, які взаємодіють між собою. Це дозволяє малим і середнім підприємствам швидко розпочати автоматизацію введення документів з невеликими витратами та легко масштабувати проект зі зростанням компанії. Інтуїтивне графічне середовище не вимагає програмістських навичок для використання системи, що сприяє значному скороченню витрат на налаштування та впровадження.

ABBYY FlexiCapture пропонує швидку окупність та повернення інвестицій, адже система дозволяє уникнути багатьох ручних операцій, таких як попереднє сортування, клавіатурний введення та ручна перевірка. Завдяки скороченню витрат на ручне введення даних користувачі можуть швидко відчути вигоди від використання *ABBYY FlexiCapture* і відновити вкладені кошти у систему.

Простота використання системи *ABBYY FlexiCapture* визначається можливістю користувача швидко розпочати роботу завдяки самонавчанню та широкому спектру інструментів для інтеграції. Технологія автонавчання в цій системі спрощує процес створення гнучких описів (*FlexiLayout*) та класифікаторів документів, роблячи значний внесок у скорочення часу налаштування системи для оптимальної роботи з різними видами документів [19].

ABBYY FlexiCapture демонструє підвищену відмовостійкість та ефективний розподіл навантаження, підтримуючи Microsoft Cluster Server. Компоненти сервера

можуть бути встановлені на відмовостійких кластерах та кластерах з балансуванням навантаження.

Web-станція введення документів в *ABBYY FlexiCapture* є ідеальним рішенням для фронт-офісу, де можна повністю автоматизувати обробку вхідних клієнтських документів в режимі реального часу. Оператор може сканувати та верифікувати документи через зручний *web*-інтерфейс. Застосунок може використовуватися на запит з будь-якого робочого місця, що забезпечує ефективну автоматизацію введення великого обсягу документів у компаніях з численними віддаленими офісами та точками обслуговування клієнтів.

Web-станція верифікації даних в *ABBYY FlexiCapture* реалізована для взаємодії робочих місць операторів із сервером системи через Інтернет. Завдяки *web*-інтерфейсу станції верифікації даних можливе проведення віддаленої перевірки документів.

В *ABBYY FlexiCapture* доступно два методи роботи станції сканування: через віддалену інтернет-установку за допомогою технології *ClickOnce* або через веб-інтерфейс у браузері, що сприяє більш гнучкому використанню функцій сканування.

Інтерфейс станції верифікації може бути настроєний за допомогою спеціальних інструментів та скриптів програми. Це дозволяє враховувати особливості процесу верифікації та адаптувати інтерфейс та процеси обробки до конкретних вимог проекту [19].

Покращена архітектура системи дозволяє *ABBYY FlexiCapture* обробляти більші обсяги сторінок завдяки ефективному використанню продуктивності кількох десятків процесорів на сервері обробки.

Основні переваги використання *FlexiCapture* включають зниження витрат на обробку документів за рахунок автоматизованої класифікації, розпізнавання та експорту даних з будь-яких джерел і типів документів до інформаційних систем. Це призводить до скорочення часу ухвалення рішень, оскільки швидка та точна обробка даних сприяє прискоренню прийняття управлінських рішень. Крім того, система забезпечує контроль на всіх етапах та рівнях, надаючи розширені можливості

моніторингу та звітності для керування потоком документів та оптимізації роботи співробітників з урахуванням пріоритетів компанії.

FlexiCapture володіє інтелектуальним вилученням даних, що базується на машинному навчанні та обробці тексту природною мовою. Це дозволяє платформі автоматично розпізнавати, класифікувати, витягувати, перевіряти та передавати дані в інформаційні системи з будь-яких текстових повідомлень та документів. Крім того, інтелектуальна класифікація дозволяє спрощувати управління інформацією в організації, зменшуючи час обробки даних та полегшуючи процес прийняття рішень.

FlexiCapture володіє легкою масштабованістю та високою продуктивністю, що робить його ідеальним рішенням для компаній, які потребують швидкої обробки великих обсягів даних. Платформа може ефективно обробляти понад три мільйони сторінок на день [20].

Система забезпечує зручний моніторинг та звітність через інтуїтивно зрозумілий інтерфейс, який дозволяє оперативно отримувати інформацію про стан та продуктивність системи. *FlexiCapture* також пропонує різноманітні варіанти звітів, таких як помилки, кількість пакетів у процесі обробки, етапи, списки активних станцій та операторів тощо.

Функція автонавчання дозволяє ефективно зменшити час впровадження та знизити витрати на обслуговування системи. Система вдосконалює свою роботу, збираючи зворотний зв'язок від користувачів, що сприяє зниженню витрат на технічну підтримку. Це особливо актуально для реалізації обширних проєктів, де потрібно оптимально обробляти великі обсяги даних.

Веб-станції *FlexiCapture*, побудовані на базі *HTML5*, забезпечують кросбраузерну підтримку для таких відомих браузерів, як *Chrome*, *Firefox*, *Safari*, *IE*, *Opera* та інші. Це дозволяє компаніям з розгалуженою структурою філій розподіляти виробничі процеси між підрозділами, незалежно від того, чи використовують робочі станції локально, віддалено або через браузер.

Система надає підтримку *SLA (Service Level Agreement)* і включає моніторинг *SLA* як одну з ключових функцій. Ці нові можливості дозволяють оптимізувати роботу системи, встановлювати та контролювати швидкість обробки для

найважливіших бізнес-процесів. Крім того, система дозволяє встановлювати бажаний час обробки даних та рівень важливості для різних типів пакетів документів.

Можливість введення даних за допомогою мобільних пристроїв стає більш доступною завдяки мобільному клієнту *FlexiCapture*. Цей інструмент дозволяє надсилати та обробляти фотографії документів, а сповіщення про успішне завантаження та обробку даних допомагають відстежувати ці процеси.



Рис. 2.6. Програма для обробки інформації з документів

Висновки до розділу 2

В цьому розділі було розглянуто програми які необхідні для повноцінного оцифрування архівних документів, адже щоб полегшити роботу з програмою, результати обробки можна зручно зберігати на комп'ютері або у папці на *FTP*. Це особливо актуально при використанні програми для обробки великої кількості документів за допомогою багатофункціональних сканерів або для щоденного сканування надходження документів в офіс.

Серед програм було обрано програму *Microsoft Word*, серед програм для сканування документів *ABBYY Scan Station*, програма для обробки документів *ABBYY Recognition Server*, програма для обробки інформації з документів *ABBYY FlexiCapture*.

РОЗДІЛ 3

ПРАКТИЧНА РЕАЛІЗАЦІЯ АРХІВІВ ЧЕРЕЗ ТЕХНОЛОГІЇ ОЦИФРУВАННЯ АРХІВНИХ ДОКУМЕНТІВ

3.1. Основні вимоги до обладнання

Всі методи цифрування можна класифікувати як контактні та безконтактні. Застосування контактного сканування є допустимим у випадках, коли: а) поверхня для цифрування є пласкою або може бути розгладжена без пошкодження об'єкта; б) контакт із скляною поверхнею не завдає шкоди об'єкту; в) світло пристрою не впливає негативно на об'єкт; г) для спеціальних видів об'єктів, які вимагають найвищого роздільного здатності (більше 1000 *dpi*, наприклад, слайдів).

У всіх інших випадках рекомендується використовувати безконтактне сканування. Освітлювачі для безконтактного сканування повинні бути вибрані так, щоб забезпечити якісне освітлення об'єкта, при цьому уникнувши випромінювання інфрачервоної та ультрафіолетової радіації, що може пошкодити документ [21].

Створення електронних копій архівних документів може бути виконане за допомогою такого обладнання як:

- професійного сканувального обладнання;
- книжкових планетарних безконтактних сканерів;
- цифрових фотоапаратів.

Вибір конкретного обладнання залежить від:

- складу архівних фондів;
- максимальних розмірів документів (сканери доцільні для форматів А2, А3, А4, А5 та менших, а фотоапарати для документів формату А1 – А0 у відсутності сканерів цього формату);
- матеріальної основи та формату носія документа;
- фінансових можливостей архівної установи.

Основні критерії вибору сканерів для створення електронних копій архівних документів включають:

1. Безпека та збереженість оригіналу:

Важливість забезпечення безпеки та недоторканості оригіналів під час процесу сканування.

2. Якість електронної копії:

Необхідність забезпечення високої якості створених електронних копій для збереження точності та чіткості інформації.

3. Розмір столу сканувального пристрою:

Обрання сканера з розміром столу, що відповідає максимальному розміру оригіналів, для уникнення фрагментарного сканування та подальшої комп'ютерної "склейки" зображень.

4. Інші технічні характеристики:

Врахування інших технічних параметрів обладнання, які можуть впливати на ефективність та результативність сканування.

Оптимальним рішенням при виборі сканерів для архівних документів є книжкові планетарні (безконтактні) сканери формату не менше А2. Вони повинні бути оснащені лампами холодного світла або *LED*-світлодіодними освітлювачами в книжковою коліскою. Це дозволяє ефективно сканувати різноманітні архівні документи, такі як книги, малюнки, старі матеріали та атласи, забезпечуючи високу якість і недоторканість оригіналів [21].

Залежно від методу оцифрування, архівні установи можуть використовувати різне обладнання:

1. Книжкові сканери на основі цифрових фотоапаратів:

Рекомендуються для оцифрування документів з вигасаючими текстами з подальшою реставрацією цифрової копії. Ефективні при використанні високоякісних цифрових дзеркальних фотоапаратів.

2. Професійні слайд-сканери:

Використовуються для оцифрування фотоплівок та слайдів з роздільною здатністю не менше 4000 *dpi* та оптичною щільністю не менше 3,3D.

3. Планшетні або плівкові сканери:

Використовуються для оцифрування фотодокументів з оптичною щільністю $D \geq 3$ та роздільною здатністю не менше 2400 *dpi*.

4. Цифрові фотокамери:

Найбезпечніший спосіб створення електронних копій архівних документів на сьогоднішній день.

При скануванні негативів та слайдів використовуються шаблони, що постачаються із сканером. Для позитивів використовують тільки планшетний сканер без шаблонів та адаптер для прозорих матеріалів.

З використанням книжкового сканера, спеціально призначеного для оцифрування архівних документів, рекомендується проводити наступні види оцифрування:

- Друкованих документів з паперовою основою.
- Документів, які часто копіюються.

Для книжкового сканера на основі цифрових фотоапаратів рекомендується використовувати для:

- Документів, які не можна безпечно притискати.
- Документів із згасаючими текстами для подальшої цифрової реставрації отриманої копії [21].

Планшетний сканер підходить для:

- Фотовідбитків, до яких не ставляються жорсткі вимоги щодо мінімізації шкідливого впливу копіювальної техніки.

Для слайд-сканера слід використовувати для:

- Негативних та позитивних плівок і слайдів з фотоматеріалами.

Дозволено також використовувати комбінації різного обладнання для розв'язання завдань з оцифрування документів різних форматів.

Вимоги до обладнання для оцифрування можуть варіюватися в залежності від конкретного завдання та потреб користувача. Однак основні компоненти, які зазвичай використовуються для оцифрування, включають:

1. Камери або сканери:

- Камери:
- Висока роздільна здатність для отримання якісних зображень.
- Можливість автоматичного фокусування.
- Здатність працювати в умовах низької освітленості.
- Сканери:
- Оптична роздільна здатність для точного відтворення документів або зображень.

- Можливість автоматичної обрізки та корекції кольору.

2. Комп'ютер або мобільний пристрій:

- Якісний комп'ютер або мобільний пристрій для обробки та зберігання оцифрованих даних.

- Великий обсяг пам'яті для зберігання великих файлів.

- Процесор із достатньою продуктивністю для роботи з великою кількістю даних.

3. Програмне забезпечення для оцифрування:

- Графічний редактор або спеціалізоване програмне забезпечення для обробки та редагування оцифрованих зображень або документів.

- Програми для розпізнавання тексту (OCR) для перетворення текстових документів у редагований формат.

4. Засоби зберігання:

- Зовнішні жорсткі диски або хмарні сервіси для зберігання оцифрованих файлів.

- Правильно організована система каталогів для легкого пошуку та управління файлами.

5. Засоби забезпечення безпеки:

- Антивірусне програмне забезпечення для захисту від можливих загроз безпеки.

- Засоби шифрування для захисту конфіденційної інформації.

6. Мережеві можливості:

– Якщо необхідно, можливість підключення до мережі для забезпечення обміну оцифрованими даними.



Рис. 3.1. Вимоги до обладнання

3.2. Вимоги до оцифрування (сканування) документів

Що означає відсканувати документ?

Зважаючи на процес отримання електронних сканів, важливо розглянути пристрій та принцип роботи сканера більш детально. Сканер – це технічний пристрій, призначений для зчитування зображення, розташованого на плоскому носії, такому як папір, та перетворення його в цифровий формат. Існують різні типи сканерів, такі як слайд-сканери, офісні та широкоформатні, але їхній принцип роботи схожий – використання світлочутливих елементів для перетворення світла, відображеного на документі, в цифровий формат.

Необхідно враховувати, що конвертація в підсумковий документ (текстовий, растровий або векторний) відбувається за допомогою спеціального програмного забезпечення на етапі обробки отриманого відбитка. Таким чином, результатом сканування є готовий файл, який представляє собою оброблені дані.

Важливо також відзначити, що сканер може використовуватися на проміжному етапі процесу, відомому як ксерокопіювання. Під час ксерокопіювання сканер служить елементом первинної обробки, отримуючи світлокопію документа та готуючи пакет для друку на принтері. Таким чином, сканування може використовуватися як окремий процес отримання електронного відображення, або як етап у вигляді ксерокопіювання, в залежності від потреб користувача [22].

Який вигляд має сканер і які у нього характеристики?

З розвитком процесу оцифрування документів розширювалася різноманітність моделей сканерів. Від офісних до промислових, що здатні обробляти потокові документи або великі формати, сканери стали невід'ємними помічниками у різних галузях, від освіти та бухгалтерії до інженерних бюро та дизайнерських контор.

Сканер, будучи окремим пристроєм, що підключається до комп'ютера, став важливим елементом для якісної обробки документів, що відображається у високих вимогах до апаратури. Різновидами сканерів є вбудовані та окремі пристрої, які поєднують функції принтера з механізмами подачі документів і пакетної обробки. МФУ (багатофункціональні пристрої) зазвичай включають сканер і можуть працювати як у режимі передачі на комп'ютер, так і непрямо, надсилаючи дані на друк без збереження їх у файл.

Щодо характеристик сканерів, слід враховувати такі параметри:

- Кольоровість (чорно-біле або кольорове сканування).
- Формат та розмір приймача (робота зі стандартами А4, А3, А0, А1).
- Роздільна здатність (вимірюється в *PPI* та *DPI*, тобто пікселі на дюйм та точки на дюйм відповідно).
- Швидкість обробки (час, потрібний для сканування одного документа).

Залежно від призначення можна виділити промислові та офісні сканери для різних завдань. Також існують ручні сканери, які використовуються там, де

неможливо розташувати документ горизонтально, або потрібна швидка обробка даних, наприклад, сканери штрих-кодів у торгових точках.

Навіщо потрібне сканування зараз?

Сканування є необхідним етапом в документообігу та важливим процесом при створенні електронних архівів. Навіть при поширенні електронних документів, паперова продукція залишається необхідною, і маса матеріалів продовжує зберігатися у формі друкованих аркушів. Бібліотеки та архіви намагаються резервувати свої колекції у вигляді електронних файлів на різних носіях, але в бізнесі та науковій сфері потреба у скануванні документів є великою.

Для бізнесу та науки сканування включає в себе оцифрування документів різних типів, включаючи креслення та плани. Сканування креслень стає популярною послугою, особливо в містах, де існує значна кількість технічних вищих навчальних закладів та конструкторських фірм [22].

Бізнесові та наукові вимоги до сканування можуть включати потокове сканування (обробка кількох сторінок одночасно), розпізнавання тексту під час сканування, або перетворення векторного формату. Широкоформатне сканування може вирішувати різноманітні завдання завдяки високій роздільній здатності та швидкості обробки.

Щодо вимог до сканерів, важливі аспекти – це швидкість, якість та можливість обробки різних розмірів документів. Від програмного забезпечення також залежить якість результату, і різні класи сканерів можуть мати різні можливості в цьому плані. Промислові сканери можуть автоматизовано обробляти велику кількість сторінок, в той час як офісні моделі можуть вимагати додаткового програмного забезпечення для розширених функцій.

Однаково для будь-яких мет, завдань чи вимог, архівний документ піддається оцифруванню лише один раз. У випадку, якщо документ вже має електронну версію, всі дії, пов'язані з виконанням замовлення, проводяться на основі його робочої копії. Кожен архівний документ, незалежно від його складності, включаючи обкладинки, шмуцтитули, аркуші з інформацією служби, зворотні сторінки тощо, підлягає повному оцифруванню.

У випадку "порожніх" аркушів (без тексту, графіки та іншої інформації, не пронумерованих), які входять до складу одиниці зберігання, є можливість не проводити оцифрування. Проте аркуші з засвідчувальним написом, які не підлягають оцифруванню, повинні бути відзначені у віртуальній копії з обов'язковим поясненням причин.

Нерозброшуровані справи оцифровуються у вигляді розвороту, а окремі документи – поаркушево, включаючи зворотний бік. Поаркушеве оцифрування нерозброшурованих справ можливе лише у випадку, коли формат аркушів у висоту перевищує розмір скануючого столу, і справи слід розміщувати горизонтально.

Зображення розвороту або аркуша повинно включати краї палітурки справи або краї аркуша, які мають йти вздовж межі зображення. Тому важливо відступити межу сканування на 0,5–1 см від краю документа, щоб забезпечити видимість його краю на зображенні.

Зазвичай справи, які підлягають оцифруванню, не розшиваються і можуть бути розшиті лише у виняткових випадках за згодою керівництва архіву, якщо копіювання розшитої справи є повністю неможливим.

Розшивання справ може бути обгрунтовано різними причинами, такими як забезпечення збереження документів (такі, що туго зшиті або зброшуровані, можуть пошкодитися або деформуватися при розгортанні на 180° та використанні притискного скла) або неможливістю повністю скопіювати інформацію, оскільки частина не вміщується у сканувальну область.

Рішення про розшивання документів приймається лише при наявності можливості палітурування після сканування розшитих справ. Після завершення роботи справа обов'язково повторно розшивається.

Якщо в справі є документи різних розмірів, для підкладання під менші документи рекомендується використовувати аркуші паперу, щоб наступний документ більшого розміру не впливав на зображення.

Небажаним є фрагментарне оцифрування великоформатних оригіналів із наступною комп'ютерною "склеюю" зображення. Для оцифрування

великоформатних документів рекомендується використовувати спеціалізовані сканери формату А1–А0 або цифровий фотоапарат.

Якщо архівний документ має тривимірну форму (наприклад, книжкове видання, сувій, рулон тощо), крім оцифрування обкладинки, шмуцтитула та сторінок за допомогою сканера, необхідно створити цифрову фотографію об'єкта, щоб відобразити його фізичний розмір, форму та особливості (наприклад, книжковий корінець, три обрізи книги, застібки) [21].

Якщо документ зберігається у складеному вигляді, його слід оцифрувати у всіх можливих варіантах складення, якщо ці варіанти мають самостійну інформаційну цінність. Для документів, таких як телетайпна або телеграфна стрічка, книга в азбуці Брайля, філіграні, рельєфні зображення та інші деталі, які видно лише під певним освітленням або кутом, сканування проводиться за допомогою спеціального обладнання.

Якщо документ має історичну упаковку, таку як конверт, папка, обкладинка, чи кофр, необхідно створити електронну копію цієї упаковки за допомогою цифрового фотоапарата.

При оцифруванні складних об'єктів, таких як фотоальбоми, де кожна сторінка містить кілька архівних документів (фотографій), які неможливо демонтувати, спочатку створюється електронна копія цілої сторінки, а потім, за допомогою спеціальних установок сканувального обладнання, скануються кожний документ (фотографія) окремо.

У випадку зміни зовнішнього вигляду оригіналу архівного документа (наприклад, внаслідок реставрації), документ повторно оцифровується, а неактуальна копія замінюється новою.

При оцифруванні бажано розміщувати на поверхні сканера дві лінійки, що демонструють фізичні розміри документа. Використання притискного скла сканерів небажано в разі, якщо товщина справи перевищує глибину книжкової коліски сканера, справа зшита таким чином, що при розгортанні на 180° виникає "хвиля", що може пошкодити документи, або при скануванні ілюмінованих або пергаментних рукописів і книг.

Перед розпочатком процесу створення електронних копій документів рекомендується провести регулярне налаштування сканеру за допомогою набору для калібрування (калібрувальні таблиці) або спеціальних тест-об'єктів (кольорових або сірих шкал) для подальшого контролю кольоровості, контрастності та чіткості електронного зображення. Також слід провести калібрування монітора комп'ютера [21].

Загальні вимоги до створення електронних майстер-копій за допомогою фотоапаратів принципово не відрізняються від вимог до сканування документів. Цифрова фотозйомка повинна здійснюватися з максимальною можливою роздільною здатністю фотокамери (не менше 150 *dpi*) у кольоровому режимі. Лінійки слід розташовувати поруч з документом для демонстрації його реальних фізичних розмірів.

При фотозйомці слід уникати додаткового освітлення. Основні вимоги до освітлення включають достатню яскравість, рівномірність, відсутність різких тіней і блисків. Ідеальне освітлення забезпечує чотири однакові тіні (промені), коли об'єкт оцифрування розміщений у центрі робочої зони. Спалах слід уникати, щоб уникнути блисків на знімку.

Основні параметри фотокамери слід налаштовувати в залежності від розмірів і кольорів документів.

3.3. Процес оцифрування документів

Сканування

Щоб увійти в середовище сканування, користувачеві необхідно ввести його попередньо зареєстровані дані у відповідні поля вікна входу в обліковий запис оператора.

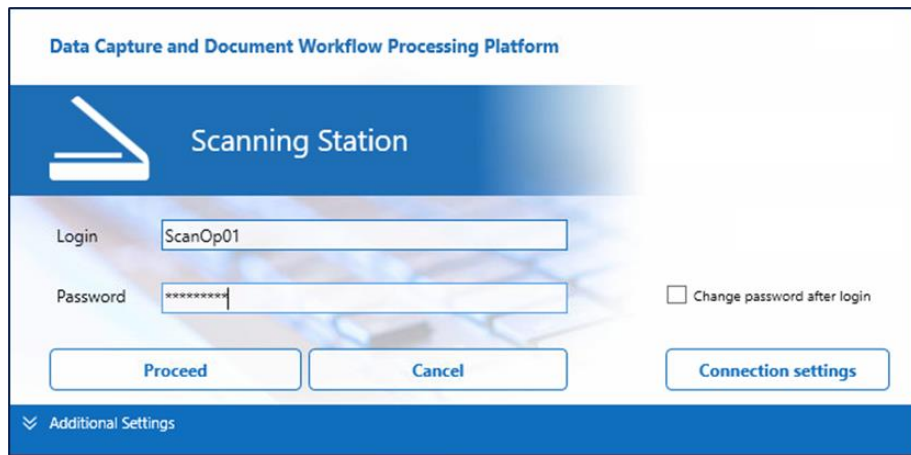


Рис. 3.2. Реєстрація користувача

Процес сканування починається зі створення пакету, який буде наповнюватись графічними образами. Для цього необхідно натиснути на кнопку «*Create Batch*».

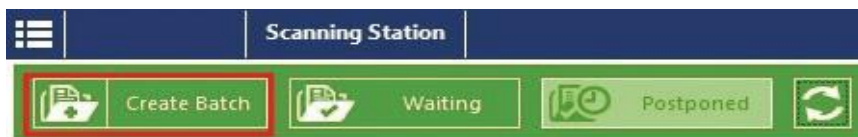


Рис. 3.3. Процес сканування

У вікні, що з'явилося необхідно обрати відповідний попередньо налаштований робочий процес.

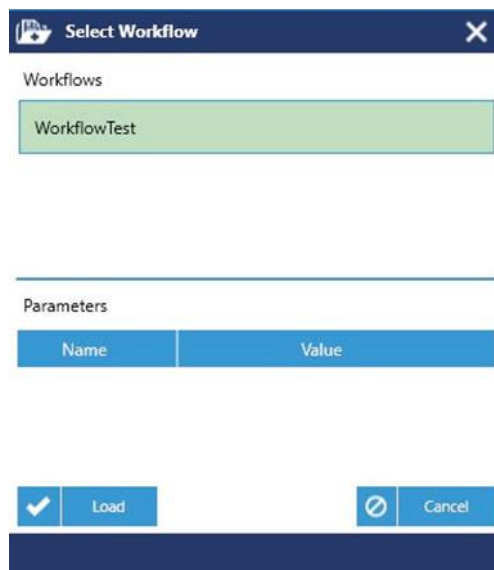


Рис. 3.4. Налаштування робочого процесу

У наступному вікні буде запропоновано обрати необхідний сканер.

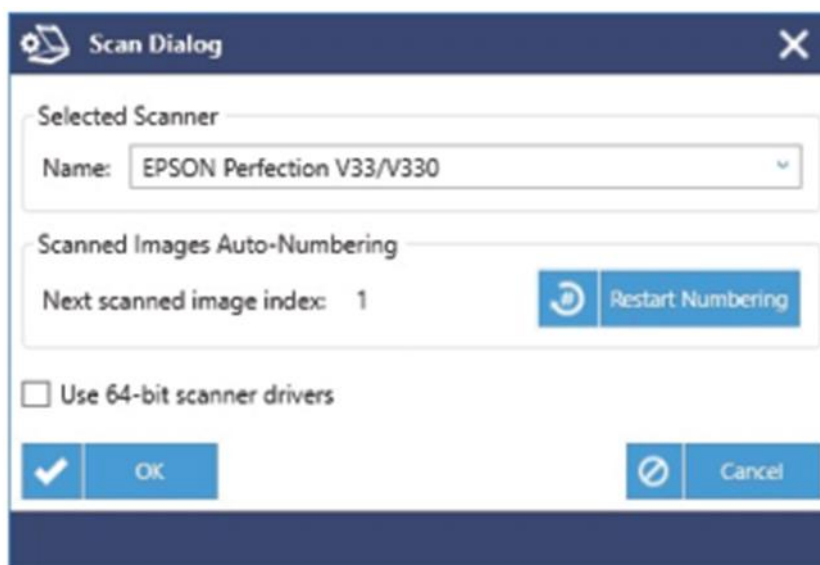


Рис. 3.5. Вибір сканера

Відскановані графічні образи формуватимуть пакет документів.



Рис. 3.6. Формування пакету документів

Реєстрація

В тому випадку, коли необхідно оцифрувати готові зображення, використовуються тікети. Тікет (англ. «Білет») – це файл, який містить в собі необхідну для подальшого оцифрування інформацію про вхідні зображення.

Отримання тікетів забезпечується спеціально розробленим рішенням «*Ticket Creator*», що формує структуровані зв'язки між вхідними зображеннями та модулями Програмного продукту.

Після звернення станцією до створених тікетів починається реєстрація даних зображень в системі робочих процесів, передбачених Програмним продуктом.

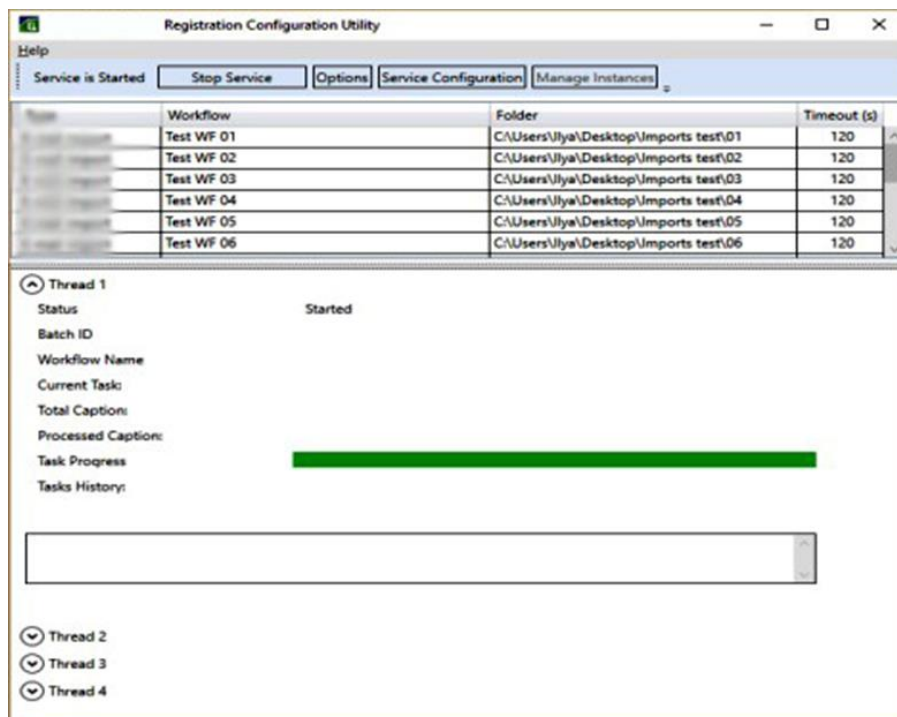


Рис. 3.7. Створення тікетів

Графічна обробка та ретроконверсія

Оскільки процеси попередньо налаштованої автоматичної графічної обробки зображень та їх розпізнавання відбуваються у фоновому режимі на відповідних станціях, участь операторів в даних процесах відсутня.

Верифікація та валідація

Аналогічно станціям сканування, станції верифікації також потребують входу зареєстрованого оператора.

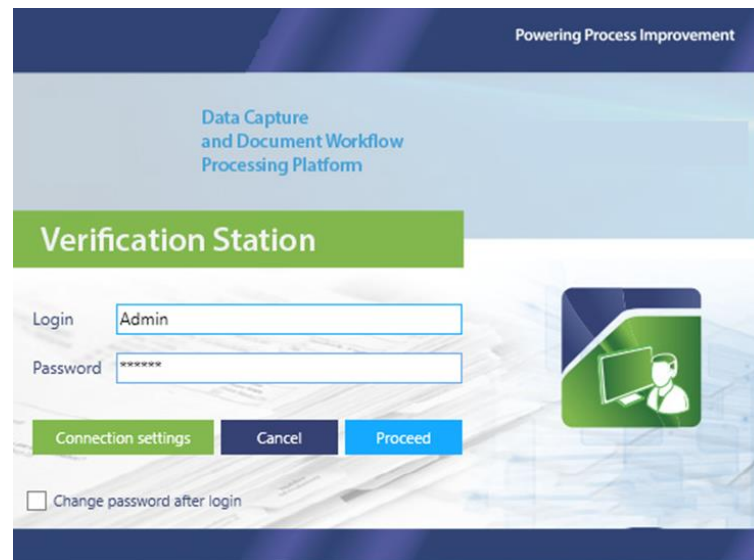


Рис. 3.8. Верифікація

Після вводу облікових даних оператор потрапляє в середовище станції верифікації.

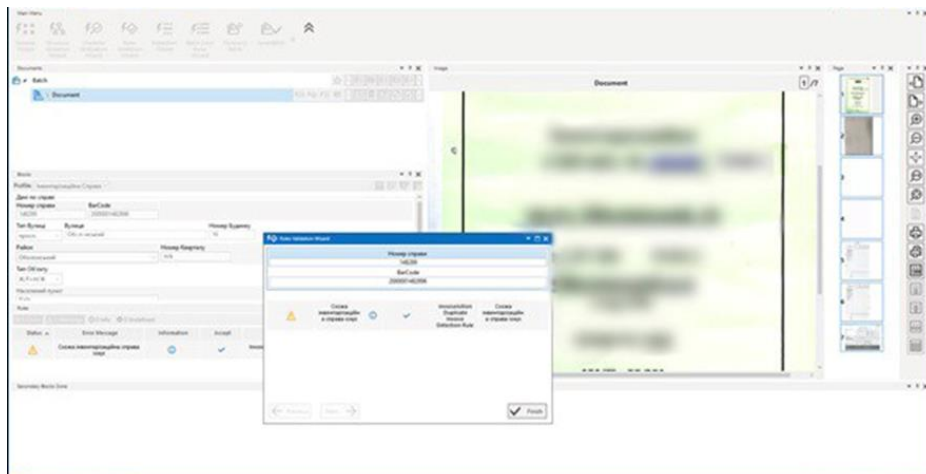


Рис. 3.9. Середовище станції верифікації

Для перевірки правильності заповнення попередньо налаштованих індексних полів необхідно порівняти інформацію, зображену на графічному образі із вилученими текстовими даними. Для цього в лівій частині робочого середовища

розташоване вікно з назвами відповідних полів. Для зручності виконання верифікації, невпевнено розпізнані символи виділені червоним кольором. Також, жовтим кольором виділено ті індексні поля, що не відповідають попередньо налаштованим правилам.

Рис. 3.10. Перевірка правильності

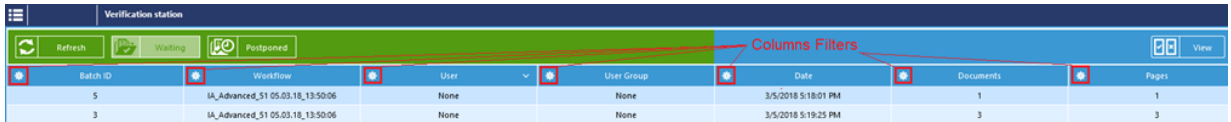
На станції валідації оператор підтверджує (за необхідності) правильність заповнення перевірених або введених вручну індексних даних.

Структура інтерфейсу станції валідації практично аналогічна структурі станції верифікації. Оператор валідації має доступ до зміни значень, які є винятками з правил. Виконання цих змін здійснюється через панель інструментів середовища станції.

Item ID	QTY	Description	Rate	Amount
12386-15	26	Senior Program Manager	172.80	4,522.80
13785-01	1	SPAREPART and Mailing	798.00	798.00

Рис. 3.11. Структура інтерфейсу

Коли валідація закінчена, оператор може відправити перевірені дані на експорт або перейти у вікно моніторингу і перенаправити (за необхідністю) той чи інший пакет даних на повторну перевірку. Для зручності навігації присутні фільтри по різним параметрам.



Batch ID	Workflow	User	User Group	Date	Documents	Pages
5	IA_Advanced_51 05.03.18, 13:50:06	None	None	3/5/2018 5:18:01 PM	1	1
3	IA_Advanced_51 05.03.18, 13:50:06	None	None	3/5/2018 5:19:25 PM	3	3

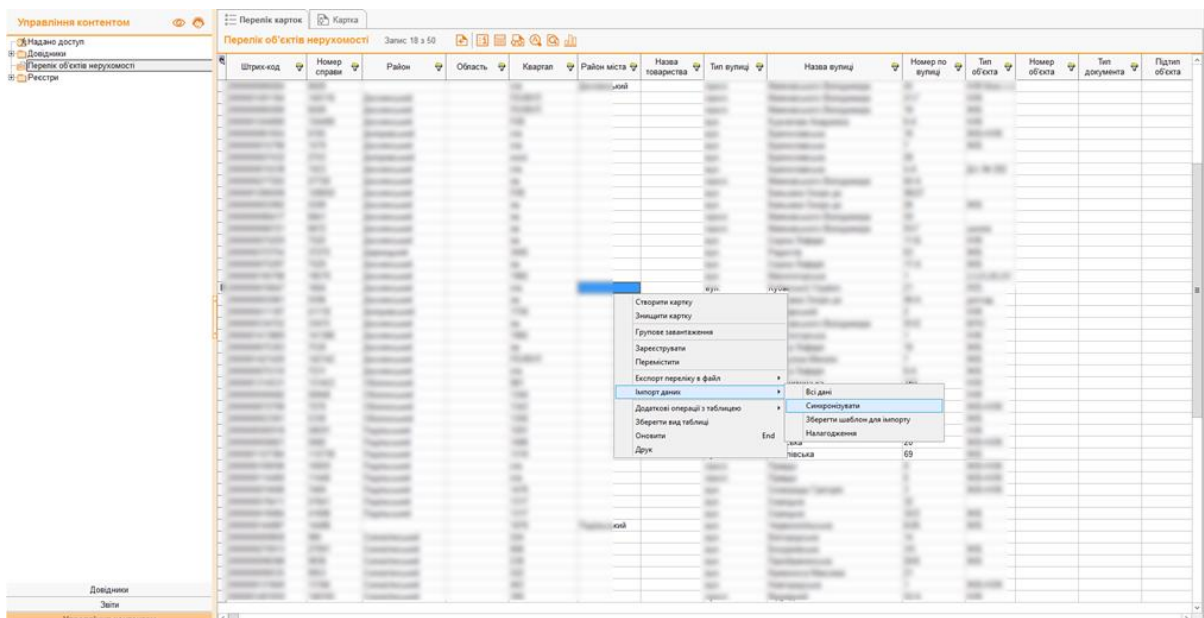
Рис. 3.12. Закінчена валідація

Експорт

Процес експорту відбувається на відповідній станції у фоновому режимі, тому участь оператора не потребується.

Формування електронної бази даних

Формування електронної бази даних електронними інформаційними ресурсами відбувається за допомогою налаштованого імпорту в середовищі системи електронного документообігу.



Штрих-код	Номер справи	Район	Область	Кадрат	Район міста	Назва товариства	Тип вулиці	Назва вулиці	Номер по вулиці	Тип об'єкта	Номер об'єкта	Тип документа	Підтип об'єкта

Рис. 3.13. Формування електронної бази

Висновок до розділу 3

В цьому розділі ми розглянули процес оцифрування архівних документів, а саме сканування документів, обробка документів, обробка інформації та як набирати та редагувати текст.

Р

о

з Було розглянуто, які є вимоги до оцифрування (сканування) документів. Які є види сканерів та які варто використовувати при оцифруванні документів.

л

я

н

у

т

о

р

о

б

о

т

у

M

i

c

r

o

s

o

f

t

ВИСНОВКИ

Під час виконання кваліфікаційної роботи розглянули цілі, що були поставлені в меті роботи, а саме оцифрування архівних документів, що відповідає даній темі.

Обґрунтували вибір програмного забезпечення, які використовували під час оцифрування документів. Було описано вимоги до оцифрування (сканування) документів.

Обробили інформацію в документах по темі кваліфікаційної роботи «Технології оцифрування архівних документів». Та проводили їх верифікацію.

При розробці кваліфікаційної роботи було закріплено знання та навички роботи з

Але було детально дослідження технології оцифрування архівних документів, їхні переваги та виклики. Результати аналізу показали, що впровадження цих технологій в архівну практику може значно полегшити доступ до інформації, збереження та управління архівними документами.

Однією з ключових переваг є підвищення ефективності роботи з архівами, адже оцифровані документи дозволяють швидший і зручніший пошук, перегляд та аналіз інформації. Також важливо враховувати аспекти збереження та захисту даних, які стають актуальними при використанні цих технологій.

Аналіз показав, що впровадження технологій оцифрування архівних документів вимагає уважного планування, розробки ефективних стратегій та врахування специфіки конкретної архівної системи. Також було виявлено важливість залучення фахівців із сфери інформаційних технологій та архівісти для успішної реалізації проектів оцифрування.

У цій роботі надано рекомендації щодо ефективного впровадження технологій оцифрування архівних документів з урахуванням сучасних тенденцій та вимог до

s

o

f

t

збереження інформації. Дослідження також вказує на потенційні переваги для наукових та дослідницьких проектів, що базуються на архівних даних.

Узагальнюючи вищевикладене, впровадження технологій оцифрування архівних документів варто розглядати як перспективний інструмент для поліпшення роботи з архівами, прискорення наукових досліджень та забезпечення більш ефективного використання інформаційних ресурсів.

СПИСОК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Оцифрування документів [Електронний ресурс]. – Режим доступу: <https://lube.ua/news/yak-otsifruvati-dokumenti-skanuvannya-ta-rozpiznavannya-tekstu>. (дата звернення 20.10.23). – Назва з екрана.
2. Переваги сканування документів [Електронний ресурс]. – Режим доступу: <https://orbrm.ru/articles/preimuschestva-professionalnogo-skanirovaniya-dokumentov/>. (дата звернення 20.10.23). – Назва з екрана.
3. Переваги та недоліків оцифрування документів [Електронний ресурс]. – Режим доступу: <https://uk.economy-pedia.com/11041145-digitization>. (дата звернення 20.10.23). – Назва з екрана.
4. Перевага оцифрованого архіву [Електронний ресурс]. – Режим доступу: <https://blogic.ru/blog/retrokonversiya-preimushchestva-otsifrovki-bumazhnogo-arkhiva/>. (дата звернення 25.10.23). – Назва з екрана.
5. Мінуси оцифрування документів [Електронний ресурс]. – Режим доступу: <https://undiasd.archives.gov.ua/doc/mr-vpluvtex.pdf>. (дата звернення 25.10.23). – Назва з екрана.
6. Історія оцифрування архівних документів [Електронний ресурс]. – Режим доступу: <https://ube.nlu.org.ua/article/%D0%90%D1%80%D1%85%D1%96%D0%B2%D0%BD%D0%B8%D0%B9%20%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82>. (дата звернення 02.11.23). – Назва з екрана.
7. Оцінка архівних документів [Електронний ресурс]. – Режим доступу: https://ru.wikipedia.org/wiki/%D0%AD%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%BD%D0%BE%D0%B5_%D0%B0%D1%80%D1%85%D0%B8%D0%B2%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5. (дата звернення 02.11.23). – Назва з екрана.
8. Текстовий документ [Електронний ресурс]. – Режим доступу: <https://ukreferat.com/chapters/komputerny-nauki/zasobi-pidgotovki-tekstovih-dokumentiv-referat.html>. (дата звернення 02.11.23). – Назва з екрана.

9. Графічний документ [Електронний ресурс]. – Режим доступу: <https://studfile.net/preview/5009784/page:15/>. (дата звернення 02.11.23). – Назва з екрана.
10. Особисті документи [Електронний ресурс]. – Режим доступу: http://dspace.wunu.edu.ua/bitstream/316497/21206/1/%D0%A9%D0%A3%D0%A0%D0%9A%D0%9E%20_.%D0%90..pdf. (дата звернення 04.11.23). – Назва з екрана.
11. Фінансові документ [Електронний ресурс]. – Режим доступу: https://oduv.edu.ua/wp-content/uploads/2016/09/lek-4-2016-mov_dil_par.pdf. (дата звернення 04.11.23). – Назва з екрана.
12. Текстовий редактор [Електронний ресурс]. – Режим доступу: <https://officetpackage.info/microsoft-word/>. (дата звернення 08.11.23). – Назва з екрана.
13. *Microsoft Word* [Електронний ресурс]. – Режим доступу: https://www.wikidata.uk-ua.nina.az/MS_Word.html/. (дата звернення 08.11.23). – Назва з екрана.
14. *ABBY Scan Station* [Електронний ресурс]. – Режим доступу: <https://diakov.net/13158-abbyu-scan-station-9042615.html>. (дата звернення 08.11.23). – Назва з екрана.
15. Програма для сканування документів [Електронний ресурс]. – Режим доступу: <https://pcprogs.net/1346-abbyu-scan-station.html>. (дата звернення 15.11.23). – Назва з екрана.
16. Програма для сканування документів [Електронний ресурс]. – Режим доступу: <https://nua.in.ua/zhurnaly/tehnologii/raspoznavanie-skanirovannykh-dokumentov-abbyu-scan-station/>. (дата звернення 15.11.23). – Назва з екрана.
17. Програма для обробки документів [Електронний ресурс]. – Режим доступу: https://www.id-mt.ru/produkty/abbyu/abbyu_recognition_server/. (дата звернення 15.11.23). – Назва з екрана.
18. Програма для обробки документів [Електронний ресурс]. – Режим доступу: <https://www.itech.net.ua/ua/abbyu-recognition-server>. (дата звернення 20.11.23). – Назва з екрана.

19. Програма для обробки інформації з документів [Електронний ресурс]. – Режим доступу: <https://www.neosoft.pro/catalog/production?vendor=abbyu&product=abbyu-flexicapture>. (дата звернення 24.11.23). – Назва з екрана.
20. Програма для обробки інформації з документів [Електронний ресурс]. – Режим доступу: https://www.id-mt.ru/produkty/abbyu/abbyu_flexicapture/. (дата звернення 24.11.23). – Назва з екрана.
21. Вимоги до обладнання [Електронний ресурс]. – Режим доступу: https://undiasd.archives.gov.ua/doc/mr_copu_docs.pdf. (дата звернення 28.11.23). – Назва з екрана.
22. Вимоги до обладнання [Електронний ресурс]. – Режим доступу: <https://fileprint.com.ua/news/skanirovanie-dokumentov-chto-eto-takoe-i-gde-eto-votrebovano/>. (дата звернення 28.11.23). – Назва з екрана.
23. Тардаскіна Т.М. Електронна комерція: Навчальний посібник / Тардаскіна Т.М. Стрельчук Є.М., Терешко Ю.В. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 244 с..
24. Про електронну комерцію: Закон України від 03.09.2015 р. № 675-VII [Текст]. – Відомості Верховної Ради України. – 2015. – № 45 – С. 410.
25. Крапивенко А. В. Технології мультимедіа і сприйняття відчуттів. - М.: Біном. Лабораторія знань, 2010. – 271с.
26. Плєскач В.Л. Електронна комерція: [підручник] / В.Л. Плєскач, Т.Г. Затонацька. – К.: Знання, 2007. – 535 с.
27. Петрушенко, В. М. (2010). Цифрові архіви: концепції та технології. Наукові праці НДІ історії України НАН України, 103-109.
28. Грищенко, Т. (2017). Оцифрування та електронний архів: питання технології та зберігання. Вісник Книжкової палати, 11, 25-31.
29. Бурлакова, І. (2014). Електронна архівна справа та цифрові технології у зберіганні та використанні архівних документів. Наукові записки Національного університету «Острозька академія». Серія "Філологічна", 33, 64-68.

30. Кравець, О. (2015). Технології цифрової обробки архівних документів. Інформаційні технології в освіті, науці та виробництві, 1, 177-183.
31. Соловйова, Л. (2019). Цифрові технології збереження архівних документів. Науковий вісник Херсонського державного університету. Серія: Історія, 35, 149-152.