

ОЦИФРУВАННЯ ДОКУМЕНТІВ ЯК ІННОВАЦІЙНИЙ СПОСІБ ЗБЕРЕЖЕННЯ ІСТОРИКО-КУЛЬТУРНОЇ СПАДЩИНИ

Світлана Литвинська – к.ф.н., доц.

Національний авіаційний університет (м. Київ, Україна)

DIGITIZATION OF DOCUMENTS AS AN INNOVATIVE WAY OF PRESERVING HISTORICAL AND CULTURAL HERITAGE

Svitlana Lytvynska – PhD (Philology), Associate Prof.

National Aviation University (Kyiv, Ukraine)

The article deals with the problems of retrospective digitization of documents. The websites of archival institutions mission in archival heritage promotion are analyzed.

Ключові слова: digitizing, information, cultural heritage, archive, media.

Завдання зберігати інформацію на всіх етапах розвитку суспільства було одним із пріоритетних, його вирішення має забезпечити для майбутніх поколінь як збереженість знань, накопичених попередніми поколіннями, так і нову інформацію. Обсяги інформації, що подається у цифровій формі, збільшуються не тільки за рахунок існування нової інформації, яка вже має електронну форму, але й за рахунок переведення у цифрову форму раніше створених інформаційних ресурсів, зокрема які містять матеріали історико-культурної спадщини. Подання такої ретроспективної інформації у цифровому вигляді дозволило вирішити низку проблем зберігання інформації, створити нові можливості для доступу та її оброблення.

Для створення копій використовується оцифрування – спосіб одержання і зберігання зображень у машинопрочитуваній формі. За допомогою сканера або цифрової фотокамери текст і зображення на паперовому носії перетворюється на подвійний цифровий код. Одержані дані зберігаються на магнітних або оптичних дисках. Потім за допомогою спеціальної програми оптичного розпізнавання знаків коду оцифрований документ перетворюється на текст, який можна редагувати за допомогою програми оброблення текстів [1, с. 120].

Сьогодні процес оцифрування має два підходи:

– обов'язковий: отримання копій сторінок у вигляді графічних (зазвичай растрових) зображень, яке здійснюється шляхом сканування або фотографування з подальшою обробленням і збереженням в одному з форматів графічних файлів. У цьому випадку повністю зберігається оригінальна верстка книги, і виключаються будь-які помилки, проте неможливий пошук або витягування фрагментів тексту;

– опціональний: розпізнавання тексту (технологія «оптичного розпізнавання символів» – OCR) з подальшим збереженням розпізнаного тексту в одному з форматів електронних книг. В цьому випадку стає можливий повнотекстовий пошук у книзі та індексація великих масивів

© Світлана Литвинська

електронних книг. Проте важко відтворити оригінальну верстку, зображення, схеми і формули. Також робляться помилки під час розпізнавання інформації.

Останнім часом (особливо з появою формату DjVu) все частіше застосовується змішаний підхід: текст книги розпізнається в автоматичному режимі й накладається на оригінальні растрові зображення сторінок, що дозволяє поєднати переваги обох підходів. Залежно від способу оцифрування використовується таке обладнання: планшетні сканери (чорно-білі та кольорові); планетарні сканери; цифрові фото- і відеокамери [2, с.13]. Оцифрування документів є достатньо складним процесом, що вимагає наявності високопродуктивної техніки і значних витрат часу.

Результатом виконаної роботи з оцифрування документів є:

- по-перше, вилучення оригіналів документів з процесу обслуговування, що сприятиме їхньому збереженню;
- по-друге, забезпечення для великої кількості користувачів одночасного швидкого доступу до документів.

Ідея створення цифрового архіву обумовлена серйозними проблемами, які виникають в архівній справі:

1. Збереження матеріалів. Основним фактором є можливість зробити копію архіву, причому оригінали для цієї операції використовують лише один раз. Надалі таку копію можна тиражувати, транспортувати, репродукувати, використовувати в поліграфії і пресі, вивчати або демонструвати, і всі ці операції відбуваються без нищення оригіналу.

2. Зменшення кількості обслуговуючого персоналу. Накопичення матеріалів вимагає збільшувати персонал, який обслуговує архів. Цифровий же архів дає змогу, не збільшуючи персоналу, працювати не з оригіналами, а лише з оцифрованими копіями.

3. Стала система архівування. Замінюючи систему опису, створюють конвертор, який переводить описи вже існуючих в архіві матеріалів на нову систему.

4. Доступність. Як правило, ті чи інші матеріали архіву не завжди можуть бути доступні для використання (реставраційні роботи, обмеження доступу до певних фондів, використання матеріалу в експозиціях або в наукових студиях). Робота з копією вирішує цю проблему, при цьому різні науковці можуть мати доступ до однієї й тієї ж потрібної інформації.

5. Географічна локалізація. Доступність архіву через інтернет дає можливість сотням науковців працювати з архівом на своєму робочому місці. Стає можливим організувати презентації, демонстрації архіву в інших містах, відкривати експозиції в Інтернеті.

6. Час наукових досліджень. Створюючи складну базу даних для оцифрування архіву, варто реалізувати методи обробки цієї інформації. Як показав досвід, наукові дослідження, виконані в різних архівах упродовж кількох місяців, тепер, за допомогою бази даних, виконують за лічені хвилини [3].

Перший і вирішальний крок у створенні цифрового архіву – визначення потреб і пріоритетів. Різні чинники можуть бути вирішальними під час переведення всього архіву в цифровий формат:

- нагальна потреба обмежити доступ до оригіналів;
- загроза їхнього руйнування;
- надходження великої кількості матеріалу і збільшення кількості обслуговуючого персоналу;
- матеріал занадто різноплановий і дуже часто змінюється система його опису;
- створюється каталог за різними архівами і система опису в них має свої розбіжності;
- велика кількість науковців не має можливості одночасно працювати з оригіналом;
- істотні обмеження, що їх створює географічна локалізація.

Слід зазначити, що робота з оцифрування паперових видань ведеться за двома напрямками:

- створення графічних копій кожної сторінки книги, на яких можна побачити текстуру паперу, використання книжкових прикрас та малюнків, використання стародавніх алфавітів у манускриптах, або друківаних книгах, які вже не використовуються у сучасному книговидавництві.
- переведення паперових видань у текстовий формат.

На сучасному етапі для архівного зберігання інформації в цифровій формі переважно використовують магнітні стрічки. Накопичено значний досвід зберігання аудіо- і відеоінформації на магнітних носіях. Існують зразки магнітних стрічок з аналоговими аудіозаписами, на яких інформація зберігається більше 50 років. За останні десятиліття значно удосконалено системи магнітного запису на стрічковому носії.

Відсутність надійних носіїв для цифрового запису визначила в останні роки інтерес до зберігання інформації на мікрофільмах в аналоговій формі. Пропонується технологія архівного зберігання, за якою створюється цифровий файл, що записується в аналоговій формі на мікрофільм як резервна копія для довготривалого зберігання. У разі втрати цифрової копії вона може бути відновлена при скануванні мікрофільму. Термін зберігання галогенідосрібних мікрофільмів може складати десятки років. Однак ця технологія не дозволяє записувати мультимедійну інформацію, забезпечувати високі швидкості відтворення даних. Також проводяться роботи з оцифрування мікроплівок.

Серед розроблених на сьогодні технологій найбільше задовольняють вимоги щодо створення носіїв для довготривалого зберігання оптичні способи запису і зберігання інформації. Найдовший термін зберігання серед оптичних дисків мають оптичні диски на скляних підкладках з одношаровим покриттям типу WORM. Термін зберігання записаної інформації на них може становити, на думку розроблювачів, 100 років, тоді як такий термін для стандартних компакт-дисків – не більше 20–30 років

З оцифрованих документів наповнюється база даних (БД), яка розробляється за допомогою різного програмного забезпечення. Також оцифровані матеріали можна подати до Internet. Наразі провідні українські архіви мають свої веб-сайти, де розміщують оцифровані матеріали. Це сприяє популяризації важливих документів, спрощує дослідникам доступ для опрацювання інформації. Проте не всі архіви достатньо повно розміщують на своїх веб-сайтах оцифровані документи. Проаналізувавши наповнення веб-сайтів шести українських обласних архівів (Державного архіву Чернівецької області, Державного архіву Черкаської області, Державного архіву Хмельницької області, Державного архіву Харківської області, Державного архіву Херсонської області, Державного архіву Тернопільської області) було з'ясовано, що лише на веб-сайтах трьох архівів є оцифровані колекції документів. Також скрізь відсутня реклама історичних документів, не вказано обсяг оцифрованих документів у квадратних метрах. Однак позитивним є той факт, що на кожному веб-сайті архіву (крім Державного архіву Черкаської області) репрезентовані документальні виставки он-лайн (тематичні, ювілейні, персональні тощо). В їхніх експозиціях оприлюднюються важливі, унікальні документи. Тематика цих виставок постійно оновлюється, а кількість збільшується. Наприклад, яскраво простежується така позитивна динаміка в роботі Державного архіву Чернівецької області. Так, на веб-сайті зазначеного архіву у 2011 р. була репрезентована лише 1 виставка, у 2012 р. – 2 виставки, у 2013 р. – 4 виставки, у 2014 р. – 9 виставок [3]. У січні 2015 р. вже з'явилась перша виставка, присвячена румунському поетові Міхаю Емінеску. На веб-сайті Державного архіву Тернопільської області на початок року розміщено 2 виставки он-лайн [4].

Отже, оцифрування сприяє популяризації документів соціально-культурної спадщини, забезпечивши при цьому їхнє збереження і можливість швидкого ознайомлення з ними.

Список використаних джерел та літератури

1. Акеройд Д. Управление развитием электронных библиотек/ Д. Акеройд // Науч. и технич. б-ки. – 2009. – № 2. – С. 119–127.
2. Бенц С. Огляд питань планування процесу переведення документів у цифрову форму / С. Бенц // Вісн. Кн. палати. – 2012. – № 2. – С. 13–15.
3. Відкритий електронний архів громадянського суспільства [Електронний ресурс]. – К.: Міжнародний фонд «Відродження»: ІАЦ «Громадянський простір», 2011. – Спосіб доступу: <http://www.e-archive.org.ua/about.jsp>. – Назва з екрана.
4. Державний архів Тернопільської області: офіційний сайт [Електронний ресурс]. – Режим доступу : <http://te.archives.gov.ua/>. – Заголовок з екрана.
5. Державний архів Чернівецької області: офіційний сайт [Електронний ресурс]. – Режим доступу : <http://cv.archives.gov.ua/>. – Заголовок з екрана.