

# ЕЛЕКТРОННИЙ ДОКУМЕНТ У ДОКУМЕНТАЦІЙНИХ СИСТЕМАХ УСТАНОВИ ТА ПРОБЛЕМИ АРХІВНОГО ЗБЕРІГАННЯ ЕЛЕКТРОННОГО ДОКУМЕНТА

УДК 930.253:[005.92:004.63](477)

Анна Алексєєнко

## ПУБЛІКАЦІЯ ЕЛЕКТРОННИХ ОБРАЗІВ АРХІВНИХ ДОКУМЕНТІВ: ОСНОВНІ ПРИНЦИПИ (З ДОСВІДУ РОБОТИ ЦДНТА УКРАЇНИ)

У статті розглянуто основні принципи публікації електронних образів архівних документів у мережі Інтернет, визначено особливості підготовки он-лайнних документальних виставок.

*Ключові слова:* інтернетизація, електронна копія документа, он-лайнова виставка, ЦДНТА України.

Архіви є потужними інформаційними системами, що зберігають значні масиви документів і організовують їхнє використання. Розвиток загальнодоступної інформаційної інфраструктури нині відбувається, зокрема, «шляхом створення в електронній формі архівних, бібліотечних, музейних фондів та інших фондів закладів культури»<sup>1</sup>, а також їхнього розміщення у світовій цифровій мережі. Це актуалізує проблеми щодо публікації електронних образів архівних документів, пов'язаних із підготовкою електронних документальних видань, розміщенням он-лайнних виставок, створенням повнотекстових баз даних на Інтернет-порталах.

Проникнення інтернетизації у діяльність архівних установ віддзеркалене у студіях із цієї проблематики. Важливим фундаментальним дослідженням архівних ресурсів України в мережі Інтернет (до 2005 р.) є кандидатська дисертація А. Кисельової<sup>2</sup>. Певний аналіз цього явища проведено у працях Г. Боряка<sup>3</sup>, М. Васильченка<sup>4</sup>, Л. Климчук<sup>5</sup>, Л. Юдіної<sup>6</sup>. У них розглянуто функціонування веб-порталу Держкомархіву (тепер – Укрдержархіву) та веб-сайтів державних архівів областей. Зосереджуючи увагу на їхній структурі й наповненні основних розділів і рубрик, дослідники залишили поза увагою серйозний проблемний пласт, пов'язаний із публікацією електронних образів архівних документів у мережі Інтернет. Ця проблема з плином часу набуває все більшого значення у роботі архівістів.

Стрімкий розвиток інформаційно-комунікаційних технологій, перехід до електронного документообігу, ширше застосування цифрових носіїв інформації змушують замислитися не тільки над питаннями щодо організації роботи з електронними документами та забезпечення

їхнього довічного зберігання, які має вирішити «Концепція планування життєвого циклу електронних документів», а й над використанням копій документів у цифрових форматах.

На увагу заслуговує праця Г. Папакіна, в якій зроблено спробу проаналізувати основні форми представлення історичних джерел в українському сегменті глобального інформаційного простору та класифікувати основні типи інтернет-публікацій<sup>7</sup>. Ґрунтовніше ці питання вивчалися російськими колегами, зокрема Ю. Грум-Гржимайлом та І. Сабенніковою<sup>8</sup>, Г. Залаєвим та Є. Бобровою<sup>9</sup>. Проте рецепції, які вони подають у своїх дослідженнях, не завжди можна здійснити в реаліях українського нормативно-правового поля та рівня інформатизації.

Нижче спробуємо визначити основні принципи, на яких має здійснюватися архівними установами публікація електронних образів документів у мережі Інтернет, зокрема для підготовки он-лайнних виставок.

Інтеграція архівних ресурсів України у світову цифрову мережу розпочалася в Україні з початку ХХІ ст. Уже тоді перед вітчизняними архівістами постала проблема вибору концептуального підходу щодо необхідності розміщення на веб-сайтах документів НАФ. Світова архівна практика подає у цьому питанні дві кардинально протилежні рецепції:

1. Повне подокументне представлення в Інтереті національної архівної спадщини. Успіху в його реалізації досяг Національний архів Австралії. Веб-сайт цього архіву може вважатися одним із найбільш рейтингових<sup>10</sup>. Універсальність австралійської системи полягає в тому, що вона представляє колосальний обсяг інформації на

© Анна Алексєєнко, 2012



рівні подокументної одиниці опису (item level): каталог містить описи 2,6 млн од. зб., що становить 10% архівного фонду країни. Цифрові копії документів можна замовити в режимі он-лайн.

Подібний підхід закладений в основу концепції створення «Електронного архіву Національної пам'яті України» – комплексного мультимедійного інформаційно-пошукового мегапроекту, який має залучити не лише електронні копії архівних документів з актуальних тем української історії, а й зображення меморіальних предметів, рукописних книг та рідкісних видань, «причому незалежно від місця зберігання таких цінностей (у тому числі в іноземних архівах, бібліотеках та музеях), адже для Інтернету не існує кордонів»<sup>11</sup>.

Проте сучасні умови технічного, фінансового і кадрового забезпечення архівних установ не дають змоги швидко виконати завдання повного подокументного представлення документів у світовій цифровій мережі<sup>12</sup>.

2. Обмежене розміщення на сайті оцифрованих копій документів або їх фрагментів. Ця позиція характерна, наприклад, для наших російських колег. Обрання її вони мотивують маркетинговим підходом. Так, у рекомендаціях Росархіву з цього приводу зазначено: «оцифрованные копии отдельных документов или их фрагментов рекомендуется помещать на сайте только в качестве иллюстративного или рекламного материала, поскольку копирование документов является платной услугой»<sup>13</sup>.

Архіви України у своїй діяльності обрали третій шлях, який є поєднанням окремих елементів означених вище напрямів: у мережі Інтернет архівні документи публікуються у необмеженій кількості, але з конкретно визначеної тематики у формі он-лайнних виставок. «Основними правилами роботи державних архівів України» (К., 2004) передбачено, що «підготовлені архівом виставки можуть тиражуватися у вигляді комплектів копій документів, тематичних добірок, розміщуватися на сайтах архівних установ та інших юридичних осіб в мережі Інтернет»<sup>14</sup>. При цьому спеціального термінологічного позначення для такої форми діяльності архівних установ в українському нормативно-правовому полі немає. Трапляються різні найменування цього поняття: он-лайнна документальна виставка, виставка on-line, віртуальна виставка Інтернет-публікація, web-архів тощо.

У структурі архівних веб-сайтів для публікації електронних образів документів передбачено обов'язкову рубрику «Документальні вистав-

ки on-line». У пропозиціях Держкомархіву від 22 лютого 2011 р. № 501/0.6 «Щодо покращення інформаційного наповнення веб-сайтів державних архівів» перелічено інформаційні блоки: «автограф української історії; Голодомор; історія державності; радянський тоталітаризм; вибори; персоналії; Друга світова війна; Чорнобиль; пам'ятні дати; катастрофи в архівах; різне; документальні виставки on-line регіональних архівів; бібліографічні виставки Фонду Президентів України»<sup>15</sup>.

На веб-порталі Укрдержархіву он-лайнні виставки розміщено за вказаною систематизацією. Для зручності пошуку створено «Тематичний покажчик виставок та окремих розділів». Він охоплює тільки виставки, представлені на веб-порталі, а також виставки регіональних архівів, на які є посилання у відповідному розділі. Виставкова он-лайн діяльність центральних держархівів залишається не охопленою. Станом на 10 жовтня 2012 р. на їхніх сайтах є 141 документальна виставка (з них 12 – виставки ЦДНТА України).

ЦДНТА України публікує в Інтернет-мережі електронні образи документів із своїх фондів з 2009 р.<sup>16</sup>, від часу появи власного офіційного веб-сайту (адреса: [www.archive.gov.ua](http://www.archive.gov.ua)). Вони розміщуються на правому полі у спеціальному розділі за датою створення у такому порядку: «Плакати часів Великої Вітчизняної війни 1941–1945 рр.» (з 16.09.2009 р.); «ЦДНТА України: Кризь призму історії (1969–2009 рр.)» (з 07.12.2009 р.); «Газотранспортні шляхи Росія – Україна – Європа: проекти газопроводів у ЦДНТА України» (з 17.03.2010 р.); «Бут Євген Миколайович (1941–2010): науковець, наставник, людина» (з 17.09.2010 р.); «85 років від дня народження М. Ф. Кондратьєва – засновника Центрального державного науково-технічного архіву України» (з 22.02.2011 р.); «Об'єкти атомної енергетики у фондах ЦДНТА України» (з 26.04.2011 р.); «Мости та мостові переходи України» (з 16.06.2011 р.); «Картографічні документи з проектною документацією фондів ЦДНТА України» (з 08.11.2011 р.); «150-річчя архітектора О. М. Бекетова» (з 03.03.2012 р.); «Життєвий і творчий шлях архітектора О. І. Заварова» (з 17.03.2012 р.); «Проекти спортивних об'єктів у ЦДНТА України (до Євро-2012)» (з 01.06.2012 р.); «Скарби науково-технічної думки: надходження НТД на державне зберігання до ЦДНТА України за роки незалежності (1991–2012 рр.)» (з 21.08.2012 р.)<sup>17</sup>.

Серед он-лайнних виставок ЦДНТА України переважають виставки документів із фондів



особового походження, приурочені до ювілеїв фондоутворювачів – видатних архітекторів, науковців, інженерів. За архівною проектною документацією підготовлено виставки, що ілюструють розвиток в Україні важливих галузей: будівництво, енергетика та транспорт. Цікавим напрацюванням є он-лайнова виставка «Картографічні документи з проектною документації фондів ЦДНТА України», яка розкриває видову та тематичну різноманітність архівної картографічної документації. При її підготовці виконано маркування – у програмі Photoshop на електронну копію документа поміщено логотип архіву – як один із засобів захисту архівної документної інформації від необмеженого копіювання без посилання на джерело<sup>18</sup>.

Створюючи он-лайнову виставку, працівники відділу використання інформації ЦДНТА України (саме цей відділ відповідає за виставкову діяльність та інформаційний супровід архівного веб-сайту) використовують як загальну методику проведення тематичної документальної виставки, так і апробовують рецепції, запропоновані методичними рекомендаціями «Підготовка та розміщення он-лайнових документальних виставок», розроблених у 2012 р.

Теоретичне осмислення створення он-лайнової документальної виставки та порівняння її з традиційною виставковою експозицією показує, що вона поєднує у собі елементи архівної виставки і технотронної публікації. При підготовці таких виставок необхідно враховувати їхні особливості, обумовлені способом оприлюднення інформації документів – публікація електронних образів документів у мережі Інтернет.

Основними перевагами он-лайнових документальних виставок є: охоплення широкого кола користувачів; можливість віддаленого доступу до документної інформації (необмеженість просторових кордонів); незалежність експонування електронних образів документів від часу роботи архівної установи; відсутність необхідності використовувати оригінали документів; можливість подання кольорового факсимільного зображення їх; можливість акцентування уваги на окремому фрагменті документа; одночасне експонування різних видів документів (писемних та аудіовізуальних); можливість представлення значних за розмірами форматів документів (притаманних для креслеників науково-технічної документації); відсутність потреби у спеціальному виставковому обладнанні; низька собівартість підготовки та організації експонування копій документів.

Електронне експонування за допомогою веб-ресурсів є перспективним щодо публікації науково-технічної документації – профільної для ЦДНТА України. Так, воно нівелює низку проблем, пов'язаних із значними габаритами креслеників при натурному експонуванні графічної документації, зокрема потреб у придбанні спеціального виставкового обладнання.

Суттєвими недоліками, що ускладнюють підготовку та експонування он-лайнових документальних виставок є: проблема захисту електронних копій архівних документів від несанкціонованого копіювання; труднощі здійснення контролю за використанням електронних копій архівних документів; необхідність регулярного оновлення конфігурації веб-сайту через швидкий розвиток інформаційно-комунікаційних технологій; відсутність екскурсійного супроводу або додаткового пояснення (при необхідності) від фахівця-архівіста; неможливість швидкого пошуку зображення документа без створення додаткового довідково-пошукового апарату.

Документ в Інтернет-публікації отримує найвищий ступінь публічності: стає вільним, доступним для інтерпретацій, джерелознавчого аналізу та історичного дослідження. Починаються складні взаємовідносини документа з громадською свідомістю.

Через необмеженість просторово-часової демонстрації образів документів в Інтернет-середовищі потрібно виважено підходити до відбору документів для експонування. Наприклад, вибір науково-технічної документації обумовлений такими застереженнями: не допустити перевантаження віртуального інфомасиву незрозумілими неспеціалісту вузькопрофільними відомостями; надати експозиції цілісно-систематичного характеру; показати науково-технічні розробки з більшою питомою вагою в контексті розвитку галузі. При правильному підборі документів позитивну роль для візуального сприйняття експозиції відіграє така властивість НТД як наочність. До виставки включається НТД, яка зосереджує в собі цінну інформацію, спроможну відтворити матеріальні образи (фотодокументи, карти, архітектурні та ескізи малюнки, кресленики тощо).

При підготовці і доборі документів для публікації у мережі Інтернет слід керуватися такими принципами:

– принцип доцільності полягає у проведенні експертизи цінності щодо необхідності розміщення кожного документа у світовій інформаційній системі з метою запобігти перенасиченості (перевантаженню) інформацією Інтернет-простору;





– принцип достатності полягає у врахуванні усіх документів, необхідних для повноти розкриття теми, а також у недопущенні зайвої деталізації при відборі документної інформації (як-то масив вузькопрофільної інформації із звітів про НДР, яка не розрахована на широкого користувача);

– принцип законності полягає у дотриманні вимог законодавства про Національний архівний фонд щодо захисту конфіденційної інформації, про інтелектуальну власність і авторське право;

– принцип достовірності полягає у забороні модифікації документної інформації при редагуванні електронного образу документа, у точності і документальній обґрунтованості подання відомостей інформаційного супроводу;

– принцип уніфікації полягає у виробленні єдиних внутрішньогалузевих методів описування, представлення і передавання інформації у мережі Інтернет (так званого власного стилю);

– принцип ідентифікації полягає у створенні повноцінного інформаційного супроводу до кожного документа, призначеного посилити комунікативну роль документа, зробити його максимально доступним, зрозумілим для користувача (при цьому заголовок документа є його важливою ідентифікуючою ознакою, що персоніфікує, передусім, його зміст, вид, дату, місце складення, автора та адресність; а посилання на стратиграфію документа – назву архіву, назву і номер фонду, назву і номер справи, сторінку справи – підтверджує його наявність як фізичної реалії, воно запрошує до перевірки достовірності, автентичності його надання користувачеві через документальну публікацію).

Поданий перелік не є вичерпним, але його можна вважати базовим для керування при підготовці он-лайнних виставок.

<sup>1</sup> Про Основні засади розвитку інформаційного суспільства в Україні на 2007–2015 роки [Електронний ресурс]: Закон України від 9 січ. 2007 р. № 537-V. – Режим доступу: <http://zakon.rada.gov.ua/cgi-bin/laws/main.cgi?nreg=537-16>. – Назва з екрана.

<sup>2</sup> Кисельова А. А. Архівні ресурси України в глобальній інформаційній мережі: види, доступ, перспективи: автореф. дис.... канд. іст. наук: 07.00.10 / Держкомархів України, Укр. НДІ архів. справи та документознавства. – К., 2005. – 16 с.; Кисельова А. А. Українські архівні ресурси в Інтернеті // Рукописна та книжкова спадщина України: археограф. дослідж. унік. архів. та бібл. фондів. – К., 2004. – Вип. 9. – С. 238–247.

<sup>3</sup> Боряк Г. В. Десять років інформатизації архівної справи. Проблеми. Добутки. Перспективи [Електронний ресурс]. – Режим доступу: <http://www.archives.gov.ua/Publicat/Researches/Doslidz-Boryak.php>. – Назва з екрана; Боряк Г. В. НАІС і проблеми доступу до архівної

інформації в документальних архівних та інформаційних системах // Національна архівна інформаційна система «Архівна та рукописна україніка» і комп'ютеризація архівної справи в Україні: зб. наук. пр. – К., 1996. – Вип. 1. – С. 19–40; Боряк Г. В. Пріоритетні проекти та перспективні напрями діяльності архівних установ в інформаційній сфері [Електронний ресурс]: матеріали до виступу Голови Державного комітету архівів України на розширеному засіданні колегії Держкомархіву, присвяченому підсумкам роботи державних архівних установ у 2006 році та їх завданням на 2007 рік (13–14 лютого 2007 р., м. Івано-Франківськ). – Режим доступу: <http://www.archives.gov.ua/Problems/>. – Назва з екрана.

<sup>4</sup> Васильченко В. В. Електронні ресурси обласних державних архівів України в мережі Інтернет: стан і перспективи розвитку в сучасному соціально-комунікаційному середовищі // Вісник ХДАК. – 2011. – Вип. 32. – С. 160–168.

<sup>5</sup> Климчук Л. В. Аналіз стану інформатизації архівів як основа вдосконалення системи документального забезпечення краєзнавчих досліджень Поділля // Бібліотекознавство. Документознавство. Інформологія. – 2011. – № 3. – С. 20–24.

<sup>6</sup> Юдіна Л. М. Використання інформаційних ресурсів Національного архівного фонду України: науково-методичні засади, напрями та форми (1993–2004 рр.): автореф. дис.... канд. іст. наук: 07.00.10 / Держкомархів України, Укр. НДІ архів. справи та документознавства. – К., 2005. – 16 с.

<sup>7</sup> Папакін Г. В. Сучасна українська Інтернет-археографія: основні форми та джерела електронних публікацій // Укр. іст. журн. – 2010. – № 5. – С. 153–166.

<sup>8</sup> Грум-Гзимальо Ю. В., Сабенникова И. В. Некоторые проблемы публикации архивных документов в электронных изданиях [Электронный ресурс]. – Режим доступа: [www.vestarchive.ru](http://www.vestarchive.ru). – Загл. с экрана.

<sup>9</sup> Залаев Г. З., Боброва Е. В. Архивы и Интернет // Отеч. архивы. – 2000. – № 2. – С. 42–50.

<sup>10</sup> Климчук Л. В. Аналіз стану інформатизації архівів як основа вдосконалення системи документального забезпечення краєзнавчих досліджень Поділля // Бібліотекознавство. Документознавство. Інформологія. – 2011. – № 3. – С. 24.

<sup>11</sup> Концепція Державної цільової програми «Електронний архів національної пам'яті України» [Електронний ресурс]: проект. – Режим доступу: [http://www.archives.gov.ua/Problems/Concept\\_EA\\_NPU-06.08.pdf](http://www.archives.gov.ua/Problems/Concept_EA_NPU-06.08.pdf). – Назва з екрана.

<sup>12</sup> Алексєєнко А. О. Інтернетизація документної сфери: впровадження електронних комунікацій у діяльність ЦДНТ України // Від XIX до XXI ст.: трансформація бібліотек у контексті розвитку суспільства: до 125-річчя ХДНБ ім. В. Г. Короленка. – Х.: РА «ІРІС», 2011. – С. 331–335.

<sup>13</sup> Боброва Е. В. Архивы via Интернет [Электронный ресурс] // НЛО. – 2005. – № 74. – Режим доступа: <http://magazines.russ.ru/nlo/2005/74/bob31-pr.html>. – Загл. с экрана.

<sup>14</sup> Основні правила роботи державних архівів України / Г. В. Боряк, О. П. Володіна, Л. О. Драгомірова [та ін.];



за заг. ред. Л. О. Драгомірової та К. Є. Новохатського. – К.: ДКАУ; УНДІАСД, 2004. – 228 с.

<sup>15</sup> Архіви України [Електронний ресурс]: офіц. веб-портал Державної архівної служби України. – Режим доступу: <http://www.archives.gov.ua>. – Назва з екрана.

<sup>16</sup> Алексєєнко А. О. Створення веб-сайтів державних архівних установ України: досвід ЦДНТА України // Збереження культурної спадщини в інформаційному суспільстві. Слово архівістів: матеріали Круглого столу, присвяченого 40-річчю ЦДНТА України. – Х.: СПДФО Михайлов Г. Г., 2010. – С. 6–8.

<sup>17</sup> Центральний державний науково-технічний архів України [Електронний ресурс]: офіц. веб-сайт ЦДНТА України. – Режим доступу: <http://www.archive.gov.ua>. – Назва з екрана.

<sup>18</sup> Дождьова О. Є. Документальні виставки on-line як форма використання документної інформації: досвід діяльності Центрального державного науково-технічного архіву України // Від XIX до XXI ст.: трансформація бібліотек у контексті розвитку суспільства: до 125-річчя ХДНБ ім. В. Г. Короленка. – Х.: РА «ІРІС», 2011. – С. 328–330.

В статье рассмотрены основные принципы публикации электронных образов архивных документов в сети Интернет, определены особенности подготовки онлайн-выставок.

*Ключевые слова:* интернетизация, электронная копия документа, он-лайн-выставка, ЦДНТА Украины.

The basic principles of publishing images of archival documents on the Internet, especially the preparation determined on-line documentary exhibits

*Key words:* connectedness, an electronic copy of the document on-line exhibition, CSSTA of Ukraine.

УДК [005.92:004.63]:004.932

Андрей Баранцев

## ПОИСК И ОТОБРАЖЕНИЕ ИНФОРМАЦИИ НА РАСТРОВЫХ ИЗОБРАЖЕНИЯХ ТЕКСТОВ, СОДЕРЖАЩИХСЯ В ЭЛЕКТРОННЫХ ДОКУМЕНТАХ

Предложен подход к созданию графического редактора растровых изображений, способного осуществлять поиск и выделение текстов на растровых изображениях документов.

*Ключевые слова:* растровое изображение, электронный архив, текстовая карта.

Непрерывное возрастание объемов архивных документов остро ставит вопрос предоставления оперативного доступа к ним. Наиболее эффективным решением этой проблемы является перевод документов с бумажными носителями в электронный вид и создание на их основе архивов электронных документов. Перевод документа с бумажным носителем в электронный вид подразумевает его сканирование и назначение ему обязательных атрибутов электронного документа. В результате такого перевода появится электронный документ, содержащий растровое изображение. Поскольку такое изображение электронного документа содержит текст, то при работе с ним было бы удобно иметь такой же набор возможностей, как и при работе с обычным текстовым документом.

Ведущие мировые проекты по созданию и накоплению электронных ресурсов (Google Book Search<sup>1</sup>, проект Национальной библиотеки Франции Gallica)<sup>2</sup> предоставляют возможности контекстного поиска фрагментов текста и наглядного отображения его результатов на растровом изображении. Это свидетельствует о востребованности подобного рода возможностей. Однако

характерной особенностью растровых изображений является необходимость использования сложных процедур для поиска и выделения необходимой информации, а проекты, успешно реализовавшие такие процедуры, не предоставляют описания их реализации.

Целью нашей статьи является ознакомление пользователей и разработчиков электронных архивов с разработанным нами механизмом отображения результатов поиска по содержимому документа на его растровом изображении.

Растровые изображения – это изображения, сформированные из сетки точек, имеющих разную яркость и цвет. Растровая графика позволяет создать (воспроизвести) практически любой рисунок, вне зависимости от сложности, в отличие, например, от векторной графики. Растровые изображения документов создаются, как правило, путем сканирования документа с бумажным носителем – создания его электронного образа с помощью специальных средств, обеспечивающих преобразование изображения на бумажном или ином материальном носителе в цифровую

© Андрей Баранцев, 2012