

U PROGU MAŁOPOLSKIEJ BIBLIOTEKI CYFROWEJ

Pomysł stworzenia Małopolskiej Biblioteki Cyfrowej MBC narodził się w Wojewódzkiej Bibliotece Publicznej w Krakowie, dzięki współpracy i zaangażowaniu władz Województwa Małopolskiego w 2005 roku. Fundusze na realizację projektu pozyskano dzięki dotacji z Urzędu Marszałkowskiego Województwa Małopolskiego, której przeznaczeniem było finansowe wsparcie dygitalizacji archiwalnych dzienników i czasopism regionalnych. Głównym natomiast celem przyświecającym inicjatywie uruchomienia MBC było stworzenie dostępu (on-line) do pełnotekstowych materiałów bibliecznych ze szczególnym uwzględnieniem zbiorów regionalnych dotyczących Małopolski, a także

zapewnienie dostępu do zasobów informacyjnych instytucji publicznych regionu; świadczenie usług przez urzędy administracji publicznej regionu; umożliwienie kontaktu z tymi instytucjami; dostarczanie w przystępnej i atrakcyjnej formie informacji na temat regionu¹.

Wśród innych powodów należałoby wymienić ochronę zasobów bibliecznych, głównie czasopiśmienniczych wydanych po 1945 roku, które – jak powszechnie wiadomo – są zagrożone samodegradacją celulozy (zjawisko określane mianem problemu „kwaśnego papieru”; zagraża drukom i rękopisom wytworzonym pomiędzy latami 1850–1990). Pozostałe, lecz niemniej ważne cele, wskazują także ochronę zbiorów z szerokiej tematyki regionalnej (w tym *Cracoviana*).

Tworząc MBC wykorzystano polską platformę systemową – dLibra (Digital Library Framework), która jest dedykowanym środowiskiem zapewniającym prawidłową obsługę zabezpieczeń, umożliwiającą łatwe wyszukiwanie i współpracę z innymi bibliotekami cyfrowymi w kraju. Warto dodać, że do opisu zasobów MBC zastosowano zdobywający coraz większą popularność schemat metadanych DublinCore².

Przed wytypowaniem dokumentów do cyfryzacji powołano, spośród pracowników Wojewódzkiej Biblioteki Publicznej w Krakowie, Zespół ds. MBC, który tworzą wyspecjalizowani w danej gałęzi wydawniczej bibliotekarze. Jako pierwsze do skanowania wybrano prasę regionalną Małopolski z lat 1945–1990, a zatem „Dziennik Polski” (1945–

¹ A. P a s z k o, *Opis projektu* – <http://mbc/malopolska.pl> (2007.02.13).

² Schemat opisu danych DublinCore w platformie dLibra jest zawarty w ramach syntaktyki RDF oraz MARC. Wykorzystanie tego typu metadanych powoduje, że informacje o obiektach mogą być przenoszone na inne platformy oraz systemy.

1990), „Życie Literackie” (1951–1990), „Gazetę Krakowską” oraz „Tygodnik Powszechny”. Z czasem zakres zbiorów zostanie powiększony o następne wydawnictwa: „Życie” (1897, 1898, 1899, 1900/nr 1), „Tygodnik Ilustrowany” (1904, 1906, 1910, 1911), „Kalendarz Krakowski” Józefa Czecha (1881, 1885, 1888, 1894–96, 1899–1917) i „Rocznik Krakowski” (1898–1936). Do decyzji autorów i wydawców należeć będzie także kwestia dygitalizacji innych wydawnictw: „Znak”, „Student” (1967–89), „Zdanie” (1978–), „Magazyn Kulturalny” (1970–83) oraz „Kraków – magazyn kulturalny” (1984–1991), „Tem!” „Więści”, „Przekrój”, „Głos Nowej Huty”, i wybranych periodyków dzielnicowych Krakowa oraz pism parafialnych, np. „Tygodnik Salwatorski”. W perspektywie wybrane zostaną również materiały poświęcone problematyce regionalnej, w szczególności małopolskie periodyki, niskonakładowe kroniki, monografie, starodruki, rękopisy, biuletyny, broszury i foldery, bibliografie, dokumenty życia społecznego, mapy oraz ikonografie; w grupie tej przewidziany jest również rocznik „Małopolska”. Tematykę planuje się poszerzyć również o zasoby naukowe, edukacyjne, informacyjne a nawet urzędowe. Toteż główny zasób MBC został podzielony na odpowiednie kolekcje tematyczne: „Czasopisma współczesne”, „Dziedzictwo kulturowe”, „Nauka i dydaktyka”, „Regionalia”, „Wydawnictwa urzędowe”, „Zbiory specjalne (od 1945 r.)”.

Małopolska Biblioteka Cyfrowa staje się zatem narzędziem, które przenosząc postać tradycyjną materiału na elektroniczną (skanowanie), zamierza integrować rozproszoną prasę lokalną Krakowa i Małopolski oraz zasoby regionalne województwa. Warto podkreślić, że obiekt cyfrowy gwarantuje długowieczność materiału, a zatem nie tylko upowszechnienie, ale też przetrwanie. Istotną sprawą jest wybór odpowiedniej metody dygitalizacji. Wojewódzka Biblioteka Publiczna w Krakowie przyjęła jedną z wielu możliwych strategii i wynajęła w tym celu (w ramach ww. grantu) profesjonalną firmę zajmującą się wykonywaniem wielkoformatowych skanów. Przygotowując nowe publikacje do MBC należy liczyć się z poszanowaniem prawa autorskiego (status prawny). Niektóre materiały w niej zamieszczone mogą dodatkowo mieć – zgodnie z wolą autorów – limitowany dostęp do wybranych publikacji lub też zabezpieczenie hasłem.

Formatem zapisu danych wykorzystywanym przy tworzeniu kolekcji do MBC jest zdobywający coraz większe uznanie format DjVu³. Pliki tego typu mają bowiem większe możliwości, gdyż zastosowano w nich bardziej efektywne algorytmy kompresji⁴. Z tego powodu większość publikacji zgromadzonych w repozytorium cyfrowym została przygotowana lub przekonwertowana do tego formatu⁵. Obraz z plików DjVu cechuje znakomita jakość przetworzonego cyfrowo materiału w stosunku do wielkości pliku oraz zdolność OCR, która w praktyce przekłada się na możliwość przeszukiwania tekstu

³ Do najczęściej wykorzystywanych formatów zapisu treści tekstowo-graficznych w Internecie wciąż należą pliki PDF, TIFF oraz format HTML.

⁴ Mechanizm kompresji obiektu graficznego jest nawet do 1000 razy mniejszy niż w plikach TIFF. Kluczową sprawą jest jednak to, że pliki DjVu gwarantują skuteczny sposób archiwizowania danych oraz zachowują niezmiennosc wyświetlania przy użyciu rozmaitych platform systemowych.

⁵ Program umożliwiający przeglądanie dokumentów (LizardTech DjVu plugin) można pobrać z sieci nieodpłatnie – <http://lizardtech.com/download/> (2007.02.13).

w dokumentach⁶. Zastosowanie technologii OCR, sprawia także, że dokumenty te są indeksowane przez wyszukiwarki internetowe: <http://www.google.com>, <http://www.yahoo.com> i inne. Ważne jest jednak to, że pliki DjVu mają stosunkowo niskie rozmiary, dzięki czemu użytkownicy MBC, korzystający z niskich przepustowości sieciowych nie będą mieć problemów z ich ładowaniem się do pamięci komputera.

Można wskazywać mnóstwo zalet elektronicznej książki, jednak wśród najważniejszych wymienić należy wysokie bezpieczeństwo cyfrowego materiału (zabezpieczenia dLibra). Zastosowane w MBC międzynarodowe standardy zapisu elektronicznego uniemożliwiają dalsze przetwarzanie danych (format DjVu jest nieedytowalny), co chroni autorów przed nielegalnym powielaniem lub edycją tekstu⁷. Na uwagę zasługuje też łatwy, graficzny interfejs użytkowy serwisu, który dzięki uwzględnieniu najnowszych tendencji w budowaniu stron www, cechuje się przejrzystą nawigacją i intuicyjnością oraz możliwością zmiany języka portalu. Porównywalna platforma systemowa bibliotek cyfrowych w kraju spowodowała, że wszystkie te biblioteki są podobne pod kątem sposobu obsługi oraz funkcjonalności, charakteryzuje je m.in. analogiczny mechanizm przeglądania publikacji. Istotną zdaje się być także kwestia zastosowania uniwersalnej strony kodowej (UNICODE), która sprawia, że znikają bariery językowe. Wśród kolekcji naukowych znajdują się mogą frazy oraz fragmenty zawierające język znaków chińskich, arabskich, greckich, rosyjskich, a nawet pismo staro-cerkiewno-słowiańskie. Nie kończą się na tym zalety cyfrowych druków, wśród mniej widocznych, ale niemniej ważnych właściwości publikacji jest możliwość ich podziału na odpowiednie poziomy, a więc specjalne warstwy oddzielające grafikę od tekstu. Rozwiązanie to ma szerokie zastosowanie (zwłaszcza ekonomiczne) przy wydrukach lub przy publikacjach, w których brak wykorzystania procedury OCR powoduje trudności w rozróżnieniu tekstu od tła.

Małopolska Biblioteka Cyfrowa została w ostatnim czasie połączona z serwisem informacyjnym Urzędu Marszałkowskiego Województwa Małopolskiego – portalem Wrota Małopolski. Posunięcie to jest istotne nie tylko z punktu widzenia rozwoju i bezpieczeństwa technologicznego (zarządzanie mocą i pojemnością serwerów), ale przede wszystkim ze względu na charakter strategiczny i promocyjny, bowiem zadaniem Wrót jest m.in. upowszechnianie informacji na temat regionu⁸. Należy oczekiwać, że współistnienie MBC i Wrót przyniesie w dalszej perspektywie obustronne korzyści i zagwarantuje jego dalszą ewolucję, a może nawet sprawi, że MBC stanie się największą cyfrową skarbnicą w Polsce. Perspektywa ta jest wszelako możliwa, wystarczy wspomnieć pomysły stworzenia

⁶ Technologia ta wykorzystuje schemat podziału dokumentów DjVu na odpowiednie warstwy – warstwa tekstu i warstwa graficzna, przy czym wynik rozpoznania tekstu zapisywany jest wewnątrz pliku na jednej z warstw, którą można dalej kopiować, np. do pliku tekstowego (TXT). Wysoka jakość obrazu DjVu gwarantuje skuteczne rozpoznawanie tekstu; więcej – <http://djvu.com/pl/> (2007.02.07).

⁷ Powszechnie stosowane są zabezpieczenia i oznakowania kolekcji w postaci zastosowania tzw. identyfikatorów wodnych, np. siglum biblioteki, copyright itp.

⁸ Zarządzanie merytoryczne zawartością cyfrowych kolekcji pozostaje w gestii WBP w Krakowie, która w dalszym ciągu będzie budować zasób oraz zachęcać do projektu inne bibliotek publiczne w Małopolsce.

wspólnego katalogu rozproszonego o nazwie Fidkar Małopolski⁹, który objął swym zasięgiem niemal wszystkie biblioteki publiczne stopnia powiatowego i potężne grono bibliotek gminnych w województwie, przyczyniając się do stworzenia największego tego typu katalogu w kraju i głównej tężnicy baz regionalnych w Małopolsce. Obecnie Fidkar Małopolski ze względu na zróżnicowaną specyfikę katalogów elektronicznych został podzielony na trzy odrębne serwisy: BR@MA (Bibliografia Regionalna Małopolski), RK@M (Rozproszony Katalog Małopolski) oraz KR@K (Krakowski Rozproszony Katalog), z których pierwszy z wymienionych pełni funkcję konsolidującą bibliografie regionalne Małopolski¹⁰. Co ciekawe, multiwyszukiwarka Fidkar również została połączona z portalem Wrota Małopolski na tych samych prawach co MBC. Niewykluczone, że w przyszłości projekty te, które zostały zainicjowane przez WBP w Krakowie zostaną połączone w jedną całość. Obydwa bowiem, charakteryzując się profesjonalnymi mechanizmami wyszukiwawczymi, upowszechniają i propagują idee regionalne.

Rys. 1. Interfejs wyszukiwawczy platformy systemowej dLibra – Małopolska Biblioteka Cyfrowa

⁹ W. K o w a l e w s k i, *FIDKAR małopolski – szansa dla bibliografii regionalnej*, „Małopolska”. T. VIII: 2006, s.293–296.

¹⁰ Już dziś można opublikować bibliografię regionalną w postaci publikacji DjVu. Technologia umożliwia (po odpowiednim zaindeksowaniu pól) przeniesienie skomplikowanej struktury opisu rekordu bibliograficznego do standardowej formy dokumentu.

Rys. 2. Przeglądanie publikacji w formacie DjVu

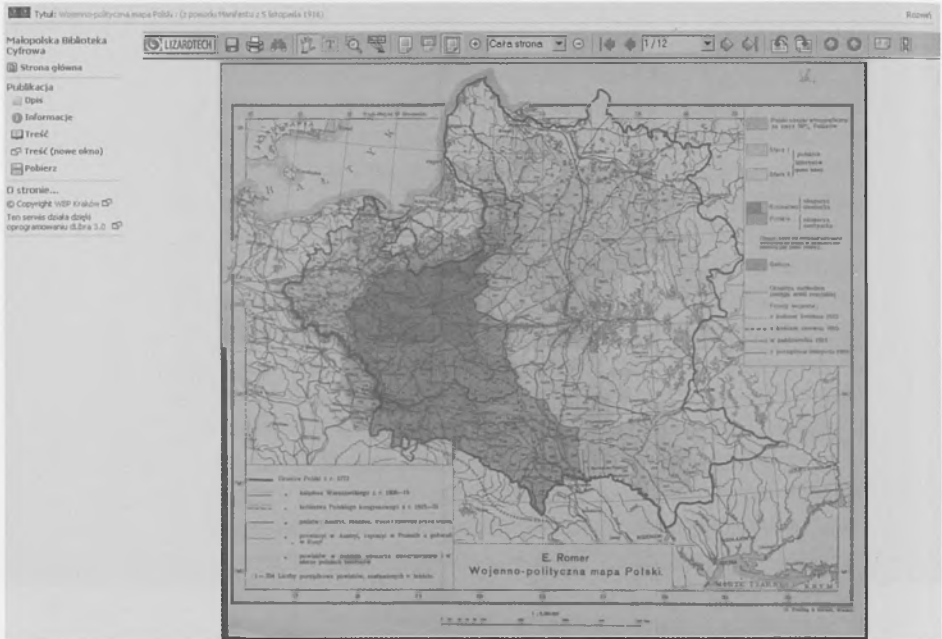


Na koniec chciałbym zauważyć, że każdy, kto nabył doświadczenie oraz umiejętność obsługi urządzeń skanujących, może stać się współtwórcą cyfrowej biblioteki¹¹. Czemu należałoby jednak poświęcić najwięcej uwagi? Otóż procesowi skanowania należy poddać dokumenty rzadkie i zagrożone samozniszczeniem, przy czym dużej ostrożności wymagają wartościowe druki i rękopisy (cymelia, inkunabuły, zabytkowe maszynopisy) oraz druki wielkoformatowe w złym stanie technicznym, gdyż łatwo można uszkodzić materiał archiwalny; w tym wypadku usługi skanowania należy zlecić odpowiednim firmom. Wojewódzka Biblioteka Publiczna w Krakowie zachęca zatem wszystkich, którym leży na sercu dziedzictwo kulturowe Małopolski, o zgłaszanie propozycji wydawniczych do MBC. Wśród naszych potencjalnych partnerów mogą znaleźć się biblioteki publiczne, instytucje kultury, wydawnictwa i stowarzyszenia regionalne, a także regionaliści i sami autorzy. Mile widziane będą publikacje w formie cyfrowej (PDF, TIFF) poświęcone zagadnieniom historycznym i regionalnym Małopolski, a w szczególności lokalne czasopisma, bibliografie, kroniki, rękopisy, opracowania naukowe dotyczące dialektologii i etnolingwistyki itd.

Dostęp do MBC: <http://mbc.malopolska.pl/>

¹¹ Wykonanie dobrej jakości obiektu wymaga skanowania w rozdzielczości 300 dpi (plik TIFF). Rozdzielczość ta umożliwi jego dalsze przetwarzanie (OCR i DjVu).

Rys. 3 i 4. Przeglądanie publikacji w formacie DjVu



GEOGRAFIA "TRYLOGII" H. SIENKIEWICZA MAPY W OPRAC. TADEUSZA PRAUSSA (ZE ZBIORÓW WBP W KRAKOWIE)



Polecamy powyższe mapy w pełnych rozmiarach [DjVu]: <http://mbc.malopolska.pl/publication/1492>



© Malopolska Biblioteka Cyfrowa, Kraków 2007

by: wrmk