

Zdigitalizowane katalogi kartkowe w Bibliotece Uniwersyteckiej we Wrocławiu

Digitized card catalogs in Wrocław University Library

Abstrakt:

Zdigitalizowane katalogi kartkowe w Bibliotece Uniwersyteckiej we Wrocławiu rozwijane są obok głównego źródła informacji o zbiorach Biblioteki – katalogu komputerowego w systemie Virtua. Są one narzędziem informującym o zbiorach, które nie znalazły się jeszcze w katalogu komputerowym. Na przestrzeni 10 lat, począwszy od początków digitalizacji katalogów – oprogramowanie do ich obsługi ulegało wielu zmianom. Aby dostosować się do rosnących potrzeb Biblioteki i wymagań użytkowników rozbudowywane było o dodatkowe moduły zwiększające funkcjonalność całego systemu. Oprogramowanie jest wciąż udoskonalane i wzbogacane o nowe funkcje, które mają na celu umożliwienie użytkownikom uzyskania jak najszybszego dostępu do poszukiwanych pozycji z ogromnego księgozbioru Biblioteki Uniwersyteckiej. W najbliższej przyszłości planuje się wykorzystanie techniki rozpoznawania tekstu, w celu wdrożenia bardziej zaawansowanych metod wyszukiwania (indeksy, słowa kluczowe). Podjęte zostaną również prace nad integracją zdigitalizowanych katalogów kartkowych z innymi systemami i serwisami użytkowymi w Bibliotece.

Abstract:

Digitized card catalogs in Wrocław University Library are developed alongside of the main source of the Library's resources, namely VIRTUA electronic catalog. The digitized card catalogs are tools informing about the collections which are still absent from the electronic catalog. Throughout 10 years of digitizing them the software employed to use them has been subject to many changes. To adjust to the increasing needs of the Library and user's requirements the software has been upgraded with additional storage modules to increase functionality of the whole system. The software is still being improved and enriched with new functions, the aim of which is to enable users possibly the quickest access to the searched material out of the immense Library's stock. In the nearest future advantage is to be taken of text recognition technique in order to introduce more advanced searching methods (indexes, keywords). Attempts are also to be made to integrate digitized catalogs into other systems and services available in the Library.

Słowa kluczowe:

digitalizacja, katalogi kartkowe, karty katalogowe, skanowanie, OCR

Keywords:

digitization, card catalogs, catalog cards, scanning, OCR

Wstęp

Katalogi odgrywają w bibliotekach kluczową rolę jako przewodniki po zbiorach. Użytkownik nie będzie w stanie w pełni wykorzystać, ani docenić nawet najbogatszego księgozbioru, jeśli będzie miał problemy z dotarciem do interesujących go dzieł lub wręcz z poruszaniem się po nim. Stąd tak ważne jest, aby biblioteki posiadały uporządkowane wykazy swoich zasobów tak, aby można było odnaleźć poszukiwane pozycje oraz ustalić ich lokalizację i uzyskać informacje o ich dostępności.

Wyjątkowo bogata historia Biblioteki Uniwersyteckiej we Wrocławiu pozostawiła w spadku spisy, wykazy, katalogi tomowe i kartkowe, pozwalające na orientację w tym okazałym księgozbiorze. Łączenie bibliotek, z których wyrosła późniejsza Biblioteka Uniwersytecka, zniszczenia spowodowane działaniami wojennymi, przekatalogowywanie zbiorów spowodowały, że u progu automatyzacji Biblioteki obok katalogów głównych – alfabetycznego, czasopism i systematycznego – istniało kilkanaście innych katalogów, w tym katalogi zbiorów specjalnych oraz katalogi historyczne, odziedziczone po przedwojennej Bibliotece Miejskiej. Dokładny opis historii i rozwoju katalogów Biblioteki przedstawiła Ewa Pitak w swoim artykule „Informacja o zbiorach Biblioteki Uniwersyteckiej we Wrocławiu – od katalogu kartkowego do OPAC” [1, s. 26–44].

Po kilku latach funkcjonowania katalogu komputerowego, mimo podejmowanych prób retrokonwersji okazało się, że skatalogowanie wszystkich zbiorów Biblioteki w katalogu komputerowym byłoby zadaniem bardzo kosztownym i czasochłonnym. Dlatego należy szukać alternatywnych rozwiązań, które pozwolą na udostępnienie pełnej informacji o całości zbiorów. Digitalizacja wszystkich katalogów kartkowych wydała się najefektywniejszym sposobem umożliwienia dostępu poprzez Internet do tych zasobów Biblioteki, które nie zostały skatalogowane w systemie Virtua.

1. Geneza digitalizacji katalogów

Początek komputeryzacji Biblioteki Uniwersyteckiej we Wrocławiu w roku 1994 wniósł nową jakość w funkcjonowanie systemu bibliotecznego-informacyjnego, a mianowicie – katalog komputerowy (OPAC). Rewolucja związana z automatyzacją oznaczała przede wszystkim możliwość zdalnego dostępu do informacji o zbiorach. Jednak mimo ogromnego wysiłku, jaki wówczas podjęto, niezwykle trudno było w stosunkowo krótkim czasie zapewnić dostęp do informacji o wszystkich zasobach Biblioteki. Pod koniec roku 1999, w piątym roku komputeryzacji, katalog komputerowy zawierał 156 tysięcy rekordów egzemplarzy [2, s. 20] co było znakomitą wynikiem, ale stanowiło ciągle reprezentację poniżej 10% całości zbiorów. Podejmowane były prace w zakresie retro-

konwersji, ale wymagały one dużych środków finansowych i były bardzo czasochłonne. Przy klasyfikowaniu pozycji do retrokonwersji stosowano kryterium chronologii wydawniczej oraz kryterium przynależności do określonych kolekcji np. skryptów i podręczników oraz publikacji pracowników UWr. Zakres retrokonwersji podany jest na stronie internetowej Biblioteki <http://www.bu.uni.wroc.pl/katalogi/katalog-komputerowy-informacje>.

Wobec tego, że zdecydowana większość księgozbioru znajdowała się poza katalogiem komputerowym i była niedostępna dla czytelników korzystających z Internetu, a wizja kompleksowej retrokonwersji była ciągle bardzo odległa, w roku 2002 zapadła decyzja o digitalizacji katalogów kartkowych. Metoda ta wydawała się najtańszym i najbardziej efektywnym sposobem udostępnienia w Internecie informacji o zasobach Biblioteki. Po przeanalizowaniu dotychczasowych, nielicznych przedsięwzięć tego typu w kraju (m.in. sposobu zeskanowania katalogu kartkowego w Ossolineum) – Dyrekcja BU podjęła decyzję o zastosowaniu zupełnie nowej metody realizacji tego zadania – nie za pomocą dużej ilości zwykłych skanerów (jak to robiono wcześniej) lecz przy wykorzystaniu jednego specjalistycznego urządzenia tzw. skanera przelotowego, który nadawał się do skanowania luźnych kartek.

W celu zdobycia funduszy na ten cel i realizacji planu – w roku 2003 złożono do Komitetu Badań Naukowych wnioski o przyznanie dotacji na inwestycję aparaturową. We wniosku znalazła się m.in. pozycja: „Digitalizacja katalogu kartkowego BUWr.” Jako uzasadnienie inwestycji podano udostępnienie w Internecie katalogu kartkowego Biblioteki Uniwersyteckiej we Wrocławiu liczącego ponad 2 miliony kart. Wnioskowano o zakup skanera do kart wraz z oprogramowaniem. Wniosek, opisujący zdecydowanie nowatorskie, jak na ówczesne czasy przedsięwzięcie – zyskał akceptację oceniających i w roku 2004 pozyskano fundusze na realizację projektu.

Zakup odpowiedniego skanera okazał się zadaniem skomplikowanym. Biblioteka planowała zdigitalizować ok. 2 miliony kart rozmaitej jakości i to w możliwie jak najkrótszym czasie, zatem skaner musiał być urządzeniem najwyższej jakości i charakteryzować się dużą niezawodnością. Przed uruchomieniem procedury zakupu dokonano szczegółowej analizy rynku tego typu urządzeń. W specyfikacji wymagań technicznych, jakie miało spełniać urządzenie, zwrócono szczególną uwagę na obecność podajnika kart, konstrukcję toru prowadzenia papieru, możliwość skanowania dokumentów o wyższej niż przeciętna gramaturze, prędkość skanowania, warunki gwarancji i serwisu. W wyniku zorganizowanego przetargu nieograniczonego wybrany i zakupiony został skaner przemysłowy Scamax 401 firmy InoTec oraz oprogramowanie DPUScan.

2. Wczesny okres digitalizacji

Stanowisko do digitalizacji kart zostało zorganizowane na początku listopada 2004 roku. Zakupiony skaner wraz z zestawem komputerowym do jego obsługi umieszczono w Pracowni Reprograficznej znajdującej się w budynku Biblioteki na Piasku. Wkrótce po uruchomieniu stanowiska rozpoczęto testy urządzenia. Okres testów pozwolił na ustalenie optymalnych parametrów skanowania oraz na organizację pracy związaną z procesem digitalizacji katalogów. Przyjęte ustawienia łączyły odpowiednią wydajność skanowania z niewielkim rozmiarem plików oraz – co najważniejsze – jakością zeskanowanych obrazów, pozwalającą również na zastosowanie w późniejszych etapach digitalizacji techniki OCR. Stanowisko zostało podłączone do serwera, na którym miały być składowane zeskanowane pliki. Ustalono sposób organizacji plików na dysku. Aby nie komplikować i nie wydłużać prac przy skanowaniu uznano, że najlepiej będzie odwzorować istniejący układ kart w szufladkach, tzn. stworzyć strukturę folderów o nazwach etykiet z szufladek. Kolejne pliki miały otrzymywać nazwy składające się z sześciu cyfr. Główną częścią nazwy miał być kolejny numer pliku uzupełniony zerami do 4 znaków oraz z dodanymi dwoma zerami na końcu, np.: 000100.jpg, 000200.jpg. Przyjęcie takiej konwencji umożliwiało w przyszłości dodawanie nowych plików zawierających nowo zeskanowane karty katalogowe.

Typowanie katalogów do skanowania nie było zadaniem łatwym, ponieważ większość katalogów kartkowych była swoistym „żywym organizmem” i wymagała ciągłych aktualizacji. Nawet jeśli po digitalizacji jakiś katalog został zamknięty, to często jego „życie” przenosiło się do świata wirtualnego i to właśnie tam były dokonywane w nim dalsze zmiany. Aby zminimalizować modyfikacje katalogu po jego zeskanowaniu, w pierwszej kolejności wybrano do skanowania katalogi przygotowane najstaranniej. Digitalizacja głównego katalogu alfabetycznego została odłożona na termin późniejszy, ponieważ katalog ten ze względu na swoją wielkość i stan zachowania – wymagał znacznych prac przygotowawczych.

Pierwszy skanowany katalog miał stanowić pole doświadczalne dla całego przedsięwzięcia. Wybrano do tego celu katalog czasopism, zawierający wydawnictwa wydane w latach 1801–1997. Został on wcześniej przygotowany do tej operacji poprzez rozdzielanie na dwa szeregi abecedłowe. Jeden zawierał karty katalogowe z opisami czasopism objętych katalogiem komputerowym, drugi – tytuły dotychczas nie notowane w katalogu komputerowym. Dzięki temu, po zakończeniu digitalizacji i udostępnieniu efektów, w Internecie miała być dostępna kompletna informacja na temat zasobu opracowanych czasopism w zbiorach Biblioteki Uniwersyteckiej [1, s. 41]. Katalog czasopism liczył 83 000 kart w 101 skrzynkach. Umiejscowiony był w budynku Biblioteki przy ul. Szajno-

chy 7/9, zatem proces skanowania wymagał transportu skrzynek z kartami do Biblioteki Na Piasku. Skanowanie rozpoczęło się pod koniec 2004 roku, a zakończyło w styczniu 2005 roku.

Drugim skanowanym katalogiem był katalog alfabetyczny Gabinetu Śląsko-Lużyckiego. Liczył 135 000 kart w 145 skrzynkach. Jego skanowanie zakończyło się jeszcze w 2005 roku. W następnej kolejności skanowane były katalogi zbiorów specjalnych: Oddziału Starych Druków, Oddziału Zbiorów Graficznych, Oddziału Zbiorów Muzycznych oraz Oddziału Zbiorów Kartograficznych.

Główny katalog alfabetyczny liczący 1 780 skrzynek i prawie 2 miliony kart wymagał przeprowadzenia melioracji. Prace te polegały na: przekatalogowaniu książek z kartami ze skróconym opisem, zmeliorowaniu i przepisaniu kart pisanych ręcznie, ujednoliceniu haseł na kartach przepisywanych, wyłączeniu z katalogu kart głównych i pomocniczych z opisami dzieł wydanych po 1994 r. a dostępnych poprzez katalog komputerowy, uzupełnieniu brakujących elementów opisu, dopisaniu brakujących kart pomocniczych. Dzięki przeprowadzonej akcji melioracyjnej znacznie wzrosła wartość informacyjna katalogu [1, s. 42]. Do melioracji katalogu alfabetycznego przystąpiono w lipcu 2004 r., prowadząc ją w pełni w godzinach etatowych – toteż prace zostały zakończone dopiero na początku roku 2008. Wtedy też rozpoczęło się skanowanie katalogu. Ze względu na jego rozmiar stanowisko skanowania przeniesione zostało do budynku przy ul. Szajnochy 7/9. W Internecie katalog pojawiał się stopniowo – w miarę jak kolejne jego partie były skanowane. Kompletny katalog został udostępniony w Internecie na początku 2011 roku.

3. Rozwój oprogramowania

Równocześnie z rozpoczęciem skanowania kart, w roku 2005 rozpoczęły się prace nad oprogramowaniem do wizualizacji katalogów zdigitalizowanych. W rezultacie konkursu ogłoszonego przez Dyрекcję BU – opracowane zostały trzy różne projekty, których wspólnym założeniem było umożliwienie szerokiej rzeszy użytkowników dostępu do informacji o zbiorach, opisanych w każdym z zeskanowanych katalogów. Zgodnie z oczekiwaniami – najważniejszą funkcją oprogramowania miało być wygodne przeglądanie katalogu przez każdego użytkownika Biblioteki, najczęściej wyposażonego w średniej jakości sprzęt komputerowy i oprogramowanie.

Pierwszy projekt wykorzystywał ideę katalogu tradycyjnego. Układ kart katalogowych w zdigitalizowanym katalogu odpowiadał układowi kart w szufladkach katalogu kartkowego. Dotarcie do żadanego miejsca w katalogu możliwe było poprzez wybranie

liter, lub etykiety szufladki. W obrębie danej szufladki można było poruszać się w przód lub w tył o zadaną liczbę kart klikając odpowiednie przyciski (Ryc. 1).



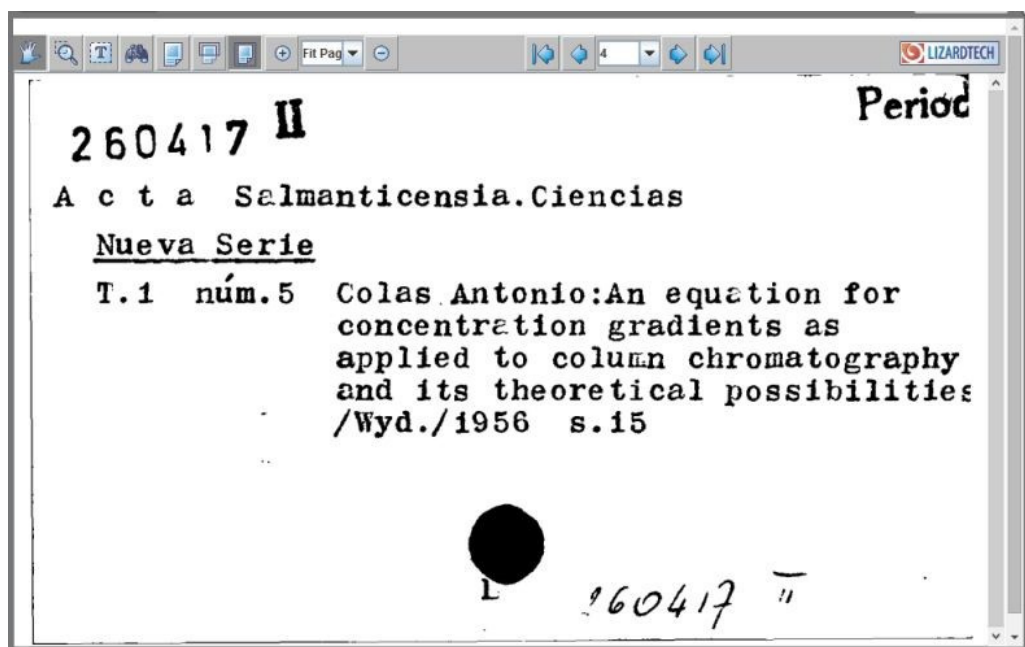
Ryc. 1.

Drugi projekt oparty był na podobnej idei, z tą różnicą, że posiadał też obsługę indeksów tytułów i sygnatur. Niestety indeksy należało wypełniać ręcznie, co wydłużało by czas potrzebny na udostępnienie całego katalogu (Ryc. 2).



Ryc. 2.

Trzeci projekt wiązał się z użyciem formatu plików DjVu. Poszczególne skrzynki katalogu zostały skonwertowane do pojedynczych plików DjVu (Ryc. 3). Konieczność przeprowadzenia konwersji skanów do tego formatu stanowiła pewne ograniczenie w elastycznym wykorzystaniu tego projektu. Ostatecznie, został on włączony do Biblioteki Cyfrowej i służy do przeglądania katalogu kartkowego czasopism w jego wersji historycznej z roku 2005, czyli momentu digitalizacji. Od tego czasu katalog kartkowy uległ znacznym zmianom, których nie odzwierciedla ta jego wersja.



Ryc. 3.

Przez kilka miesięcy roku 2005 wszystkie trzy projekty były prezentowane w Internecie. Ostatecznie zdecydowano się pozostawić i rozwijać projekt pierwszy, jako najmniej skomplikowany i najpowszechniej dostępny (nie wymagał wysokiej jakości sprzętu komputerowego).

3.1. Moduł zamawiania

Wkrótce po udostępnieniu w Internecie dwóch pierwszych katalogów zdigitalizowanych pojawiła się potrzeba rozszerzenia funkcjonalności oprogramowania użytkowego o możliwość zdalnego składania zamówień na wybrane z katalogu pozycje. Składanie zamówień na pozycje ze zbiorów Biblioteki było dotychczas możliwe tylko z katalogu komputerowego oraz w tradycyjny sposób – na rewersach papierowych. Wdrożenie zamawiania online z katalogów zdigitalizowanych miało umożliwić zdalne wysyłanie formularzy rewersów w wersji elektronicznej, bez potrzeby wizyty w budynku Biblioteki. W podję-

tych pracach nie planowano jednak tworzyć kolejnej bazy czytelników (przedsięwzięcie zbyt skomplikowane programistycznie i zbyt kosztowne), wobec czego trzeba było pogodzić się z pozostawieniem oczywistej wady katalogów tradycyjnych – brakiem weryfikacji statusu czytelnika oraz zamawianej pozycji – podczas wysyłania formularza zamówienia. W praktyce, po stronie obsługi zamówienia niewiele się zmieniło, w stosunku do tradycyjnego sposobu zamawiania. Wysyłany przez Internet formularz drukowany był w Magazynie Biblioteki i dopiero Wypożyczalnia sprawdzała, czy dana pozycja może zostać wypożyczona czytelnikowi. W późniejszych latach udało się jednak wprowadzić do katalogu zdigitalizowanego mechanizm weryfikacji statusu czytelnika. Polega on na tym, iż system ma możliwość sprawdzenia, czy zamawiający nie ma zablokowanego konta i czy może złożyć zamówienie. Jeśli istnieją przeciwskazania – użytkownik otrzymuje komunikat o przyczynie, z powodu której operacja zamówienia nie może zostać wykonana.

Formularz zamówienia musiał zostać dostosowany do charakteru księgozbioru. Inny zestaw pól wymagany jest do złożenia zamówienia na wydawnictwa ciągle z Katalogu Czasopism, a inny na wydawnictwa zwarte z Katalogu Gabinetu Śląsko-Lużyckiego (Ryc. 4). W późniejszym okresie, po zeskanowaniu katalogów fotograficznych Oddziału Starych Druków konieczne było wykonanie formularza zamówienia do digitalizacji z jeszcze innym zestawem pól. Przy udostępnieniu katalogu alfabetycznego wykorzystano lekko zmodyfikowany formularz użyty wcześniej dla Katalogu Gabinetu Śląsko-Lużyckiego. W tym przypadku należało dodać listę lokalizacji skonstruowaną tak, iż użytkownik wybierając lokalizację wskazaną pieczęcią na karcie – mógł w efekcie uzyskać informację, gdzie konkretnie (nazwa agencji i lokalizacja budynku bibliotecznego) może odebrać zamawianą pozycję. Moduł zamawiania wdrożono na początku 2007 roku.

The image shows two different library order forms side-by-side. The left form is for 'British Union Cat.' and the right form is for 'Katalog komputerowy'.

Left Form (British Union Cat.):

- Top section: 'Pola oznaczone gwiazdką są wymagane'. Fields include 'Lokalizacja*' (with a note to provide location according to card info), 'Sygnatura*', and 'Tytuł*' (with a note to include series and volume).
- Middle section: Fields for 'Tom/Rocznik', 'Rok', 'Nr/Zeszyt', and 'Nr ogólnego zbioru'. A note says: '(Wpisać numer, lub w zamian miesiąc ewentualnie kwartał jeśli jest znany, zamawianego rocznika czasopisma.)'
- Bottom section: Fields for 'Nr karty bibliotecznej*', 'Nazwisko*', and 'Imię*'. Buttons for 'Zamawiam!' and 'Anuluj'.

Right Form (Katalog komputerowy):

- Top section: 'Pola oznaczone gwiazdką są wymagane'. Fields include 'Sygnatura*', 'Autor', 'Tytuł*', 'Rok wydania', 'Nr karty bibliotecznej*', 'Nazwisko*', and 'Imię*'. Buttons for 'Zamawiam!' and 'Anuluj'.
- Handwritten notes: '32276 IV Per.' and '32276E 9.5L.' are visible on the form.

Ryc. 4.

Pewnym problemem technicznym, przy realizacji modułu zamawiania było kierowanie formularzy do określonych drukarek. Kryteriami skierowania są: rodzaj katalogu oraz typ zamówienia (do wypożyczalni albo do czytelní), a przy organizacji obsługi w nowym gmachu Biblioteki Czytelní Czasopism Bieżących i Czytelní Śląskiej należało jeszcze uwzględnić kryterium bloków sygnatur. W momencie wdrażania modułu zamawiania istniały 3 stanowiska drukowania rewersów w dwóch budynkach: stanowisko druku zamówień do wypożyczenia w magazynie gmachu głównego Biblioteki przy ul. Szajnochy 7/9, stanowisko druku zamówień do Czytelní Głównej w tym samym miejscu oraz stanowisko druku zamówień do Czytelní Zbiorów Specjalnych w magazynie budynku Biblioteki Na Piasku przy ul. Św. Jadwigi 3/4. Po częściowej przeprowadzce Biblioteki do nowego budynku przy ul. Joliot-Curie 12 i uruchomieniu tam dwóch czytelní w roku 2014 powstały w tamtejszym magazynie 2 kolejne stanowiska druku.

3.2. Moduł edycji kart

Katalog kartkowy czasopism po zdigitalizowaniu został zamknięty. Oznaczało to, że nowe karty nie były już do niego dokładane, a istniejące karty nie były modyfikowane. Opisy czasopism są jednak o tyle specyficzne, że zawierają informacje o zasobach, czyli kolejnych pojawiających się wydaniach danego tytułu czasopisma i w związku z tym – zmieniają się. Podjęto zatem decyzję, że aby katalog zdigitalizowany zachował swą aktualność – zmiany te muszą być w nim rejestrowane.

The image shows a digital interface for editing library cards. At the top, it says 'Zbiory Graficzne'. Below that, the card number '263999 III' is displayed in red. The title 'Fotografia' is written in a stylized font. A table lists issues for the years 1976, 1977, and 1978. A red button 'Zobacz katalog komputerowy' is located below the table. At the bottom, there is a checkbox 'W katalogu komputerowym', a dropdown menu 'Zmiana skierowania', a text input field 'Tekst do wstawienia:', and buttons 'Zapisz', 'Anuluj', and 'Przywróć oryginalną kartę'.

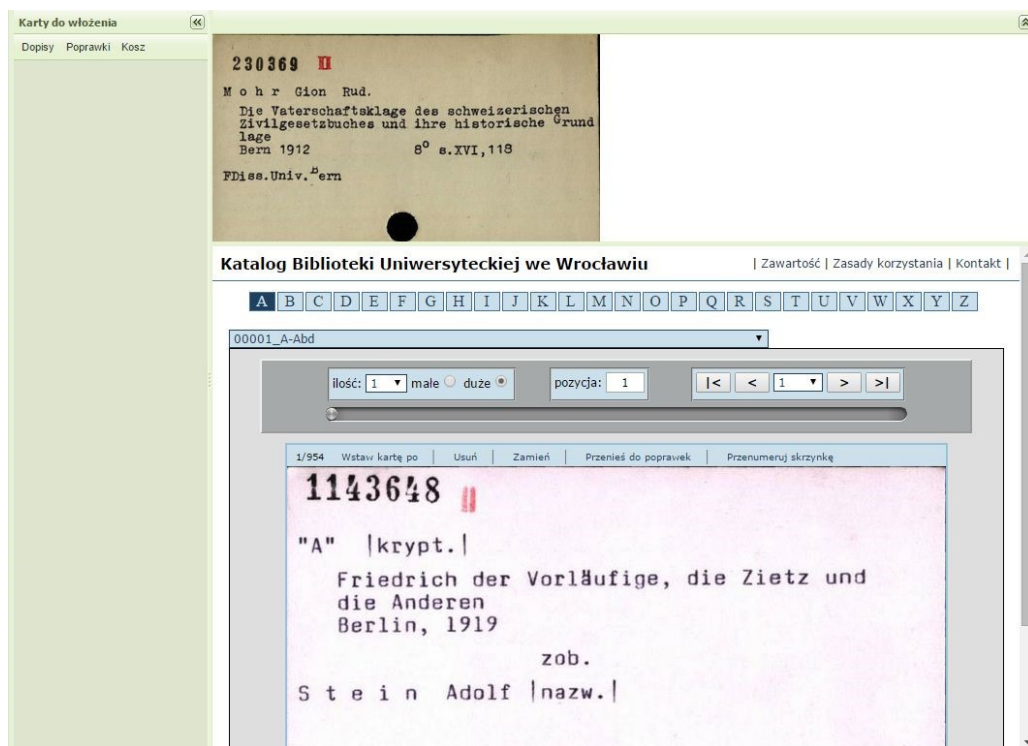
Ryc. 5.

Do tego celu opracowany został specjalny moduł do edycji kart. Ma on możliwość dopisywania na istniejących kartach nowych informacji, ale też umożliwia wymazywa-

nie wybranych informacji z kart, zmianę skierowania poprzez wymazanie istniejącej pieczęci i wstawienie nowego skierowania, dodanie komunikatu o kontynuowaniu danego tytułu w katalogu komputerowym (Ryc. 5). Moduł został wdrożony w roku 2007, natomiast w roku 2015 dostosowano go do obsługi katalogów d. Gabinetu Śląsko-Lużyckiego (alfabetycznego, czasopism i systematycznego), ze względu na podjętą meliorację tych katalogów.

3.3. Moduł modyfikacji katalogów

Trochę inaczej sytuacja wyglądała z katalogami, które nie zostały zamknięte. Dotyczyło to przede wszystkim głównego katalogu alfabetycznego. Katalog ten wymagał bardzo częstych operacji dodawania, usuwania i wymiany kart starych na nowe. Wraz ze zbliżającym się momentem udostępnienia tego katalogu w Internecie pojawiła się potrzeba stworzenia narzędzia do jego modyfikowania. Już w trakcie procesu skanowania, zeskanowane szufladki były niejako blokowane, aby nie dopuścić do powstania rozbieżności w katalogu tradycyjnym i zdigitalizowanym. Opracowany został program, który umożliwia wstawianie nowych, zeskanowanych kart, usuwanie oraz zastępowanie kart (Ryc. 6).



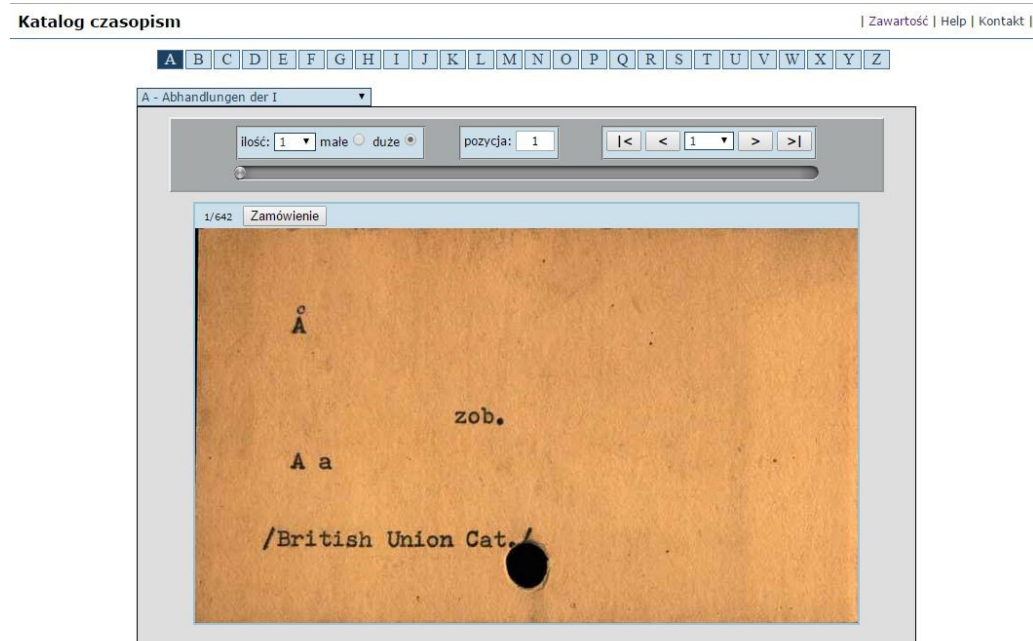
Ryc. 6.

Idea prawidłowej modyfikacji zeskanowanych katalogów opiera się na tym, że operator programu musi jednocześnie wykonywać pracę na fizycznych skrzynkach kata-

logowych i na programie. Program został udostępniony pod koniec roku 2009. Od tego czasu, użyto go do dodania 24000 kart, wymiany 2 000, zmiany lokalizacji 7 300 oraz usunięcia 5 500 kart.

Nowa wersja programu do obsługi katalogów zdigitalizowanych

Pod koniec roku 2007 została udostępniona nowa wersja oprogramowania prezentacyjnego (Ryc. 7). Nowy program różnił się znacznie, pod względem wizualnym od poprzedniej wersji, ale pod względem funkcjonalności nie zmieniło się zbyt wiele, gdyż dodany został suwak, umożliwiający bardziej wygodne poruszanie się w obrębie wirtualnych szufladek katalogu. Istotniejsze zmiany tkwiły „wewnątrz” programu, m.in. pojawiła się możliwość bezpośredniego odwołania do konkretnej karty oraz połączenia programu z zewnętrznym źródłem danych, czego pozbawiona było pierwotna wersja. Rozszerza to potencjalne możliwości wyszukiwania o bezpośrednie przejście do poszukiwanej karty – po wpisaniu początku hasła. Połączenie z bazą danych zostało też wykorzystane do innych celów, o czym będzie mowa w dalszej części tekstu.

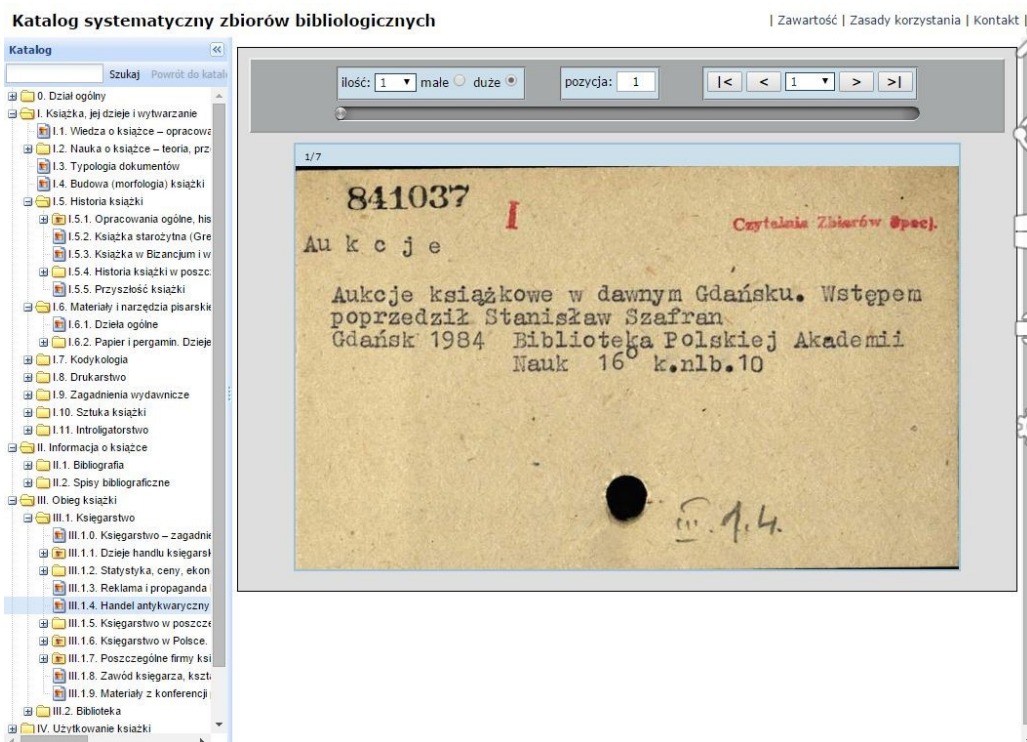


Ryc. 7.

3.4. Katalogi systematyczne

Oprogramowanie do prezentacji katalogów dobrze sprawdzało się w przypadku katalogów alfabetycznych. Problem pojawił się w przypadku katalogów systematycznych, które oparte są na wielostopniowym układzie działów wg przyjętego schematu klasyfi-

kacji. W roku 2007 zostały zeskanowane dwa takie katalogi – Katalog Systematyczny Zbiorów Bibliologicznych oraz Katalog Systematyczny Gabinetu Śląsko-Łużyckiego. Zwirtualizowanie drzewiastej struktury katalogów i połączenie jej z kartami wymagało włożenia sporego wysiłku. Przy skanowaniu tych katalogów struktura folderów i plików została odwzorowana zgodnie z układem hierarchicznym obecnym w szufladkach. W trakcie prac okazało się jednak, że w takim przypadku układ ten nie będzie zbyt wydajny. Należało więc tę strukturę „spłaszczyć”, a powiązania haseł i kart zapisać w bazie danych. Oryginalny schemat klasyfikacji należało wprowadzić praktycznie ręcznie – zapewniając powiązania hierarchiczne pomiędzy poszczególnymi działami. Do obsługi tych katalogów opracowano nowy interfejs, umożliwiający nawigację po drzewiastej strukturze haseł (Ryc. 8). Dodatkowo wyposażony on został w możliwość filtrowania nazw działów, co przy bardzo rozbudowanej strukturze obu katalogów (około 1 500 działów każdy) znacznie ułatwia wyszukiwanie. Prace wykonano w roku 2008.



Ryc. 8.

3.5. Moduł do obsługi wypożyczeń

Moduł do obsługi wypożyczeń powstał w roku 2007 i był przeznaczony do obsługi zamówień składanych do czytelnii z katalogu komputerowego. Ponieważ zamówienia z katalogów zdigitalizowanych są realizowane głównie na miejscu, w czytelnich Bi-

blioteki, to również one zostały objęte tym programem. Idea programu oparta jest na dwóch ekranach wyświetlających wykazy pozycji znajdujące się w różnych statusach udostępniania. Każde złożone zamówienie zapisywane jest w bazie danych z odpowiednim statusem („zamówiony”). Na pierwszym ekranie znajduje się wykaz takich właśnie nowych zamówień. Gdy czytelnik przychodzi do czytelnicy po zamówioną pozycję, pracownik zmienia jej status na „wypożyczony”. Wykaz pozycji z tym statusem zawiera drugi ekran programu. Przy zwrocie, status zmieniany jest na „zwrócony”. Program jest bardzo pomocny w pracy czytelnicy, ale spełnia też rolę pseudo bazy czytelników i wypożyczeń. Przy składaniu zamówień na pozycję ze statusem „zamówiony” lub „wypożyczony”, czytelnik jest informowany o jej niedostępności. Baza ta pozwala też na ograniczenie liczby zamówień składanych przez jednego czytelnika, zgodnie z Regulaminem udostępniania zbiorów.

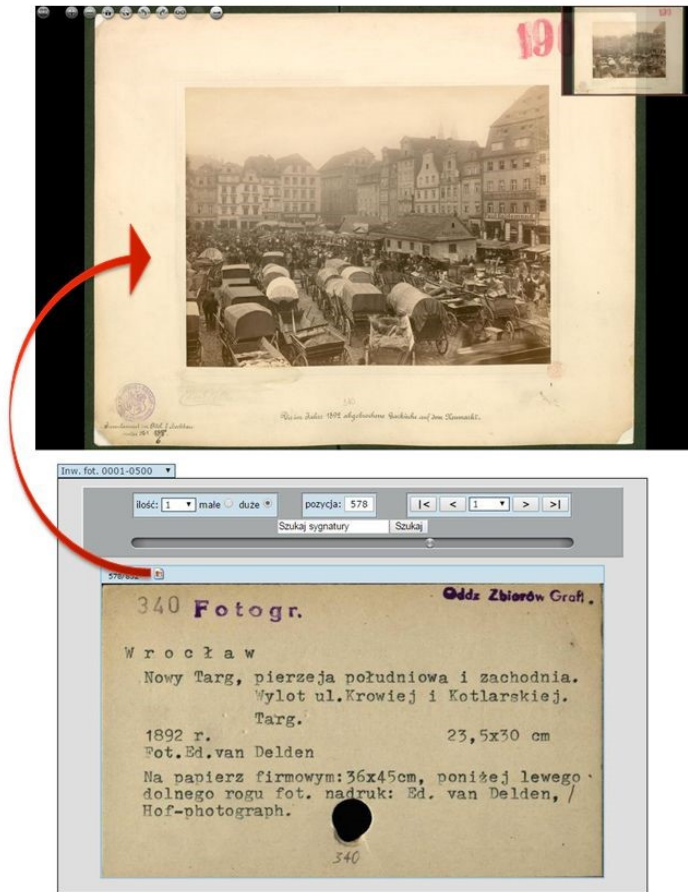
3.6. Funkcje eksperymentalne

Równoległe obok rozwijania zdigitalizowanych katalogów kartkowych Biblioteka prowadziła intensywne prace nad digitalizacją zbiorów. Stało się to zachętą do podjęcia prób uatrakcyjnienia funkcjonowania zdigitalizowanych katalogów poprzez np. połączenie zdigitalizowanych kart katalogowych ze zdigitalizowanymi materiałami bibliotecznymi, których te karty dotyczą. Przykładem takiego rozwiązania jest zdigitalizowany Katalog Topograficzny Fotografii Oddziału Zbiorów Graficznych. Fotografie, których opisy zawiera katalog, zostały zeskanowane i udostępnione poprzez Bibliotekę Cyfrową UW. Możliwość łączenia kart katalogowych z zewnętrznymi źródłami danych pozwoliła, w tym przypadku, na powiązanie kart z opisywanymi przez nie oryginałami (Ryc. 9).

Na takiej samej zasadzie istnieje możliwość łączenia dowolnych kart z odpowiadającymi im zasobami cyfrowymi dostępnymi w Bibliotece Cyfrowej lub z opisami znajdującymi się w katalogu komputerowym. Jest to jeden przykład z licznych możliwości integracji różnorodnych systemów informacyjnych i bazodanowych, funkcjonujących w Bibliotece, która stanowi kierunek rozwoju automatyzacji Biblioteki w następnych latach.

Możliwość podłączenia katalogów kartkowych do zewnętrznych źródeł danych pozwoliła też na wykonanie programu, umożliwiającego wpisywanie z kart określonych elementów opisu i wiązanie ich z odpowiednimi kartami. I tak na przykład wymieniony wyżej Katalog Topograficzny Fotografii Oddziału Graficznego, a także jego katalogi siostrzane – Autorów Fotografii oraz Rzeczowy Fotografii zostały w ten sposób wzbogacone o numery sygnatur, dzięki czemu w przyszłości będzie można połączyć te trzy, istniejące w rzeczywistości niezależnie, katalogi kartkowe – w jeden wirtualny katalog. Katalog autorów został ponadto uzupełniony nazwami autorów, a katalog rzeczowy hasłami rzeczowymi, co znacznie ułatwi korzystanie z tej kolekcji. Niestety, tego typu dane

należy mozolnie uzupełniać używając klawiatury, gdyż w katalogu tradycyjnym są one wpisane ręcznie na kartach przekładkowych, co uniemożliwia zastosowanie techniki OCR (*Optical Character Recognition*)¹.



Ryc. 9.

Duże możliwości usprawnienia nawigowania w katalogu dostarcza użycie techniki rozpoznawania tekstu (OCR). W roku 2010 opracowana została procedura konwersji OCR kart katalogowych składająca się z szeregu programów i skryptów, w efekcie działania których uzyskuje się ustrukturyzowany materiał tekstowy, odpowiadający głównym elementom opisu. Od tego czasu prowadzona jest konwersja kart głównego katalogu alfabetycznego. Oczywiście należy mieć świadomość, że nie jest to technika doskonała. Uzyskane efekty konwersji mogą jedynie stanowić podstawę do dalszych prac korekcyjnych, ale jednak i tak dzięki temu znacznie może się skrócić czas potrzebny do udostępnienia treści katalogów. Analiza wyników konwersji przeprowadzona na próbce 500 kart pokazała, że

¹ OCR – Optical Character Recognition – optyczne rozpoznawanie znaków

całkowicie bezbłędnie zostało odczytanych 19% opisów. Wydaje się, że to niezbyt wiele, ale nie wszystkie elementy opisu są niezbędne do efektywnego wykorzystania katalogu. Trafność odczytywania hasła wyniosła 56%, a sygnatury i tytułu nawet ponad 70% (odpowiednio 74% i 77%).

Konwersja OCR głównego katalogu alfabetycznego jest niestety długotrwała, ze względu na jego wielkość. Problemów nastroczają też zmiany dokonywane ciągle w tym katalogu. Nawet po zakończeniu tych prac przygotowanie pozyskanego materiału do udostępnienia będzie stosunkowo długotrwałym przedsięwzięciem. Dlatego w roku 2015 postanowiono wykonać konwersję OCR na mniejszym katalogu, aby przetestować możliwości, które ta technika daje. Wybrano w tym celu dwa katalogi Oddziału Zbiorów Muzycznych – Alfabetyczny i Systematyczny. Materiał tekstowy uzyskany w wyniku konwersji został powiązany z odpowiadającymi mu kartami. Wyniki przetwarzania zostały udostępnione do korekty pracownikom oddziału. Elementy opisu poddane korekcie ograniczono do sygnatury, hasła oraz tytułu. Można się spodziewać, że po zakończeniu tych prac – oba zdigitalizowane katalogi kartkowe Oddziału Zbiorów Muzycznych będzie można przeszukiwać znacznie sprawniej, za pomocą słów kluczowych.

4. Plany na przyszłość

W okresie 10 lat funkcjonowania zdigitalizowanych katalogów kartkowych na witrynie internetowej BUWr – zwiększała się liczba udostępnianych katalogów oraz rosła funkcjonalność całego systemu poprzez obudowanie go dodatkowymi modułami. W nadchodzących latach najważniejszym zadaniem ma być usprawnienie możliwości wyszukiwania poprzez wykorzystanie techniki OCR.

Tekst uzyskany z kart dzięki konwersji OCR można wykorzystać na dwa sposoby. Połączenie obrazów kart z odczytanymi hasłami pozwoliłoby dotrzeć do poszukiwanego hasła – zaraz po wpisaniu go do odpowiedniego pola formularza i naciśnięciu przycisku Enter. Obecnie odnalezienie poszukiwanego hasła karty wymaga wykonania stosunkowo wielu kliknięć. Byłoby to więc znaczne przyspieszenie operacji wyszukiwania. Jednak jeszcze ciekawszym usprawnieniem byłoby umożliwienie przeszukiwania całej treści katalogu lub wybranych elementów opisu wg zadanych słów kluczowych i w odpowiedzi zwracanie zawęzonego zbioru kart zawierającego te poszukiwane słowa. Wdrożenie takiej funkcji wymaga jednak gruntownej przebudowy całego systemu, gdyż obecna wersja nie umożliwia stosowania filtrów, które byłyby tu niezbędne.

Niewątpliwym kierunkiem rozwoju katalogów zdigitalizowanych powinna też być ich integracja, udostępnienie możliwości przeszukiwania wszystkich zdigitalizowanych

katalogów kartkowych poprzez jeden wspólny interfejs. Dałoby to bardzo ciekawe możliwości wyszukiwawcze.

Zakończenie

Po latach użytkowania katalogów zdigitalizowanych wydaje się, że w najbliższym czasie nie będzie dla tej platformy innej alternatywy. Ewolucja oprogramowania do obsługi katalogów powinna zaspokoić potrzeby najbardziej wymagających użytkowników. Problemem pozostanie niestety konieczność używania kilku systemów katalogowych do odnalezienia poszukiwanych pozycji. Rozwiązaniem mogłoby być dokonanie OCR wszystkich katalogów zdigitalizowanych i podłączenie ich pod multiwyszukiwarke, wykorzystywaną w Bibliotece.

Bibliografia załącznikowa

5. PITAK, E. Informacja o zbiorach Biblioteki Uniwersyteckiej we Wrocławiu – od katalogu kartkowego do OPAC, W: *Komputeryzacja Biblioteki Uniwersyteckiej we Wrocławiu – dziesięć lat doświadczeń*, red. Grażyna Piotrowicz. Wrocław: Centrix.pl s.c. Krzysztof Cebula i Tomasz Kalota, 2006, s. 26–44. ISBN 83-921013-2-4, 978-83-921013-2-1. Dostępny w: <http://www.bibliotekacyfrowa.pl/publication/958>.
6. ŁUKASZEWICZ, J. Początki i przebieg komputeryzacji w Bibliotece Uniwersyteckiej we Wrocławiu w latach 1994–2004/2005 W: *Komputeryzacja Biblioteki Uniwersyteckiej we Wrocławiu – dziesięć lat doświadczeń*, red. Grażyna Piotrowicz. Wrocław: Centrix.pl s.c. Krzysztof Cebula i Tomasz Kalota, 2006, s. 13–26. ISBN 83-921013-2-4, 978-83-921013-2-1. Dostępny w: <http://www.bibliotekacyfrowa.pl/publication/958>.