

Nie tylko druki – dane o e-książkach w katalogu Biblioteki Uniwersyteckiej w Toruniu

TEKST: AGNIESZKA DWORNIK, GRZEGORZ SZTURO

Biblioteka Uniwersytecka w Toruniu (BU UMK) współtworzy bazę NUKAT od początku jej istnienia, czyli od 2002 roku. Kooperacja w zakresie wymiany danych ma swe początki jednak już w roku 1996, kiedy biblioteka zaczęła przysyłać hasła do Centralnej Kartoteki Hasel Wzorcowych. BU UMK dostarczyła do bazy centralnej około 190 tys. opisów bibliograficznych dla wydawnictw zwartych, ponad 3 tys. dla wydawnictw ciągłych, a skopiowano ponad milion rekordów dla książek i 26 tys. dla czasopism¹. Stopień zgodności katalogu BU z bazą NUKAT jest konsekwentnie zwiększany poprzez bieżące katalogowanie, wprowadzanie danych retrospektywnych oraz prace melioracyjne. Obecnie niemal 86% opisów bibliograficznych w bazie katalogowej Horizon jest zgodne z bazą NUKAT. W dużym stopniu jest to też efekt scalania danych bazy lokalnej w projekcie „NUKAT – Autostrada Informacji Cyfrowej”.

Tak duży zbiór metadanych jest wykorzystywany przez książnicę UMK nie tylko do opracowania bieżących wpływów bibliotecznych. Zastosowanie znajdują one również w monitorowaniu krajowego ruchu wydawniczego² czy masowej konwersji danych z katalogu kartkowego do bazy komputerowej³. Dane te są też wykorzystywane do informacji o e-zasobach i ich promocji.

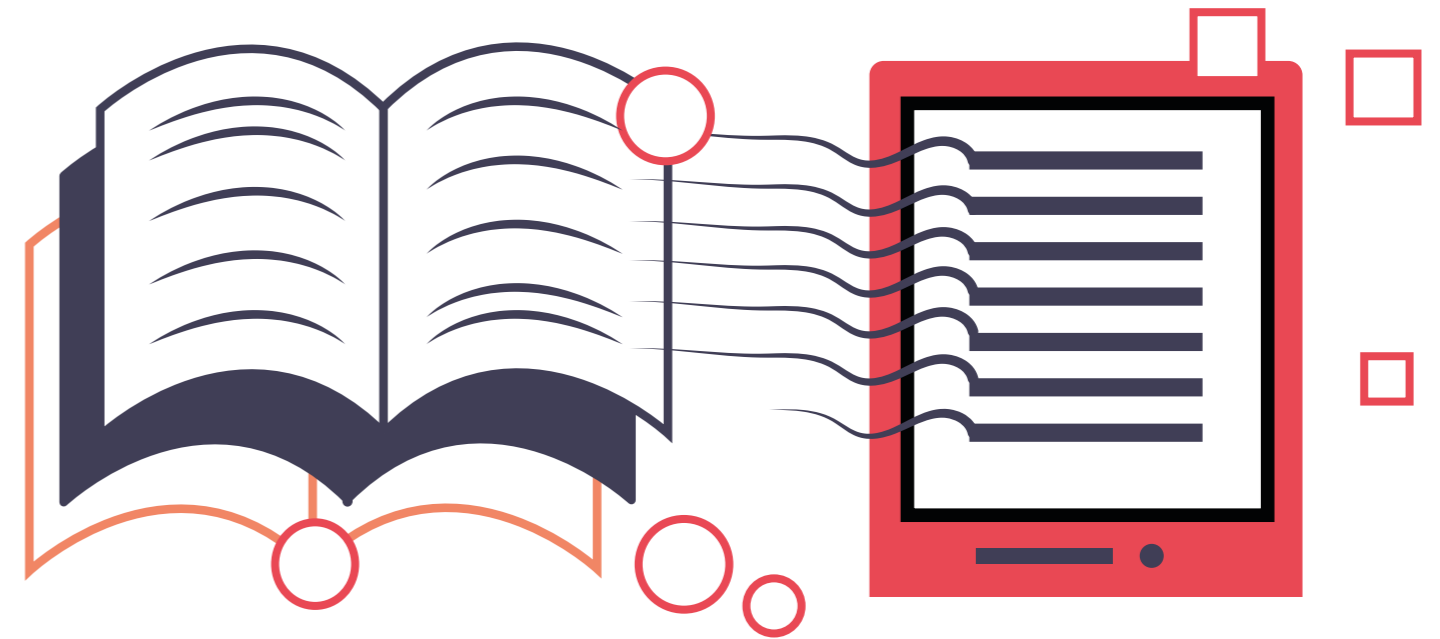
KPBC i nie tylko

Kujawsko-Pomorska Biblioteka Cyfrowa zainicjowała swoją działalność w 2005 roku. Biblioteka, tworzona obecnie przez

16 regionalnych instytucji, udostępniła ponad 140 tysięcy zdigitalizowanych obiektów. Na zasób biblioteki pod względem liczby dokumentów (gdzie np. każdy numer gazety to oddzielna publikacja) składają się głównie czasopisma. Książki stanowią zaledwie 3% ogółu dokumentów, generują jednak aż 38% wyświetleń dostępnych w bibliotece publikacji⁴. Schemat postępowania z danymi bibliograficznymi został opracowany przed przystąpieniem do procesu digitalizacji i udostępniania publikacji elektronicznych. Kluczową, a na początku prac wcale nie tak oczywistą, sprawą było przyjęcie za podstawę opisu katalogowego oryginalnego, drukowanego dokumentu piśmienniczego, a nie jego cyfrowej emanacji. Nie wszystkie biblioteki cyfrowe przyjęły taką zasadę. Ustalenia KPBC okazały się zbieżne z późniejszymi wytycznymi Centrum NUKAT. Chcąc zapewnić wysoki standard budowanego katalogu w programie dLibra, postanowiono w możliwie dużym stopniu wykorzystać dane z narodowego centralnego katalogu NUKAT⁵.

Metadane pilnie poszukiwane

Przed przystąpieniem do wprowadzania metadanych do programu dLibra zdecydowano, aby dane bibliograficzne konwertować z formatu MARC21. Przygotowywanie danych zgodnie z ujednoliconymi zasadami zapisu bibliograficznego jest najlepszym sposobem zapewnienia precyzyjnych hasel wyszukiwawczych, zwłaszcza w odniesieniu do elementów kontrolowanych w bazie NUKAT przez kartoteki hasel wzorcowych (twórcy, serie). Pomimo że biblioteki



cyfrowe posiadają samodzielne narzędzia przeszukujące ich zasoby, akcent położono na konieczność przestrzegania najwyższych standardów metadanych, aby ułatwić dotarcie do odpowiedniej publikacji, zwłaszcza korzystającym z opcji zaawansowanych. Po wytypowaniu książki z zasobów biblioteki UMK do digitalizacji, metadane publikacji są umieszczane w katalogu elektronicznym. Jeśli opis istnieje w bazie NUKAT, jest z niej pobierany, jeżeli brak dla niego gotowych danych, opis jest tworzony w bazie lokalnej i wysyłany do bazy centralnej. Następnie jest on konwertowany do formatu DublinCore i uzupełniany o dane charakterystyczne dla tego formatu (np. status praw autorskich). Wydaje się, że konwersja rekordów jest najbardziej ekonomicznym rozwiązaniem pozwalającym na zachowanie spójności danych w obu bazach. Przy pracach związanych z digitalizacją wzięto pod uwagę również pilną potrzebę przenoszenia danych z katalogu kartkowego do elektronicznej bazy bibliograficznej. Od września 2007 roku katalog kartkowy BG UMK udostępniany jest w formie zeskanowanych kart katalogowych. Po umieszczeniu danych w katalogu elektronicznym, skany kart głównych wraz z odsyłaczowymi zostają usunięte. Dzieje się tak w myśl przyjętej zasady, że książka pobrana z magazynu powinna przed zwrotem zostać skatalogowana w katalogu elektronicznym. W ten sposób cyfryzacja wspomaga retrokonwersję danych.

Łączenie informacji

Po umieszczeniu dokumentu w KPBC, w katalogu biblioteki uniwersyteckiej UMK dodawane są linki lokalne w polu 957. Cyklicznie, co miesiąc, wg wygenerowanych list, uzupełnia się o linki do KPBC bazę NUKAT. Po zatwierdzeniu melioracji w katalogu centralnym linki lokalne zamieniane są automatycznie na pole 856. Bardzo duże nadzieje wiązane są z nowym narzędziem NUKATu, umożliwiającym dodawanie linków bez konieczności melioracji całego opisu bibliograficznego. Czynności te bardzo spowalniały pracę nad linkowaniem. Po udostępnieniu nowego narzędzia

1. Stan ma marzec 2016 wg statystyk na [www NUKAT-u](http://www.NUKAT-u)
2. G. Szturo, Wykorzystanie centralnych baz w pracy gromadzenia, [w:] Strategie gromadzenia zbiorów w bibliotekach polskich : materiały z IV Ogólnopolskiej Konferencji Naukowej, Pobierowo, 15–17 września 2011 / pod red. Urszuli Ganakowskiej, Mirosławy Różyckiej, Szczecin 2012, s. 206–215. Tekst dostępny również w Repozytorium UMK: <http://repozytorium.umk.pl/handle/item/278>
3. Projekt: Retrokonwersja masowa. 942/P-DUN/2015. (realizacja 2015–2016).
4. Ogólnodostępne dane na stronach Kujawsko-Pomorskiej Biblioteki Cyfrowej: <http://kpbc.umk.pl>
5. D. Czyżak, G. Szturo, M. Rakowska, Kujawsko-Pomorska Biblioteka Cyfrowa : problemy i zagadnienia związane z realizacją I etapu jej rozwoju, [w:] Biuletyn EBIB, nr 11 (81). Tryb dostępu: http://www.ebib.pl/2006/81/a.php?czyzak_szturo_rakowska

do tych prac będzie można zaangażować również osoby nie wprowadzające na co dzień danych przez klienta Virtui.

Oczywistym motywem umieszczenia linków w katalogu biblioteki jest

zapewnienie możliwie kompletnej informacji o zasobach, a także sposobach dotarcia do nich. W pierwszym etapie wyboru obiektów do udostępniania w KPBC dużą uwagę zwracano na obiekty w złym stanie zachowania. Ich digitalizacja umożliwiła wycofanie ich oryginalnych form z udostępniania czytelnikom, oferując jednocześnie dostęp do wersji cyfrowych dokumentów z poziomu katalogu OPAC. Informacje o zdigitalizowanej formie dokumentu ułatwiają kontrolowanie zbędnego cyfrowego powielania tej samej publikacji. Biblioteka Uniwersytecka w lokalnej bazie posiada niemal 6 tys. rekordów bibliograficznych z uzupełnionymi polami dostępu do zasobów cyfrowych KPBC. Liczba ta może wydawać się skromna, zwłaszcza jeśli porówna się ją z liczbą dokumentów w bibliotece cyfrowej. Należy jednak pamiętać, iż w bibliotekach zrzeszonych w Federacji Bibliotek Cyfrowych każdy numer wydawnictwa ciągłego jest liczony jako oddzielna publikacja. W rekordzie MARC21 natomiast odnotowujemy jedno łącze do tzw. opisu publikacji grupowej w formacie Dublin Core. Jeden link kieruje niejednokrotnie do np. kilkuset numerów gazety udostępnionej w zasobach elektronicznych.

Około 80% rekordów z linkami w bazie lokalnej posiada relację z bazą NUKAT. W początkowej fazie współpracy nie wszystkie utworzone opisy zbiorów specjalnych zostały wystane do bazy centralnej. W latach 2006–2007 na podstawie dokumentacji formatu MARC21 Library of Congress zostały opracowane lokalne formularze dla druków ulot-

Obecnie w katalogu BU UMK jest około 28 tys. rekordów z hipertąciami do polskich bibliotek cyfrowych. Największa liczba linków prowadzi do zasobów Wielkopolskiej Biblioteki Cyfrowej (ponad 2 tys.) oraz Polony (ok. 600)

nych, rękopisów i dokumentów ikonograficznych. Były one wykorzystywane przez kilka lat, do czasu ogłoszenia przez Centrum NUAKT oficjalnych przepisów katalogowania i formularzy. Po odpowiedniej melioracji również ta część metadanych będzie mogła być odnotowana w katalogu NUKAT. Biblioteka Uniwersytecka w Toruniu oprócz prowadzenia KPBC, uruchomiła również samodzielną bibliotekę cyfrową, tzw. Zasób Cyfrowy UMK, w której udostępnia skany książek, których prawo autorskie zabrania upowszechniać w domenie publicznej. Dotychczas udostępniono ponad 970 publikacji, głównie podręczników akademickich, ale również książki dostarczone jako egzemplarz obowiązkowy w formie plików pdf. Część wydawców i instytucji naukowych oraz społecznych coraz częściej przesyła uprawnionym bibliotekom swoje publikacje w formie pliku. W przypadku tego zasobu cyfrowego organizacja pracy wygląda niemal identycznie, jak w przypadku KPBC. Jediną różnicą jest pominięcie ostatniego etapu pracy, czyli dostarczenia hipertącza do zasobu w centralnym katalogu NUKAT. Rekord w bazie lokalnej zawiera za pomocą odnośnika w polu 957 dostęp do zeskanowanej książki.

Więcej linków, więcej informacji

W naszej bibliotece przyjęto zasadę, aby w katalogu elektronicznym gromadzić metadane również do zasobów innych bibliotek cyfrowych czy stron internetowych, które oferują nieograniczony dostęp do publikacji. Obecnie w katalogu BU UMK jest około 28 tys. rekordów z hipertąciami do polskich bibliotek cyfrowych. Największa liczba linków prowadzi do zasobów Wielkopolskiej Biblioteki Cyfrowej (ponad 2 tys.) oraz Polony (ok. 600). Pozostawia się w bazie, mimo że nie ma

W bibliotece druków, opisy książek z linkami, które importowane były z bazy NUKAT w celu kontroli egzemplarza obowiązkowego czy w trakcie prac związanych z retrokonwersją.

Subskrybowanie bazy danych w katalogu

Informacje o publikacjach dostępnych w bibliotekach cyfrowych z założenia są wprowadzane na stałe, posiadają niezmiennie hipertącza, status prawny publikacji i instytucji, które je udostępniają w przestrzeni publicznej, gwarantują nieograniczony dostęp wszystkim użytkownikom Internetu. Jak wiemy, zgoła odmiennie przedstawia się sytuacja z publikacjami elektronicznymi z subskrybowanych baz książek. Możliwość korzystania z nich jest bowiem ograniczona do uprawnionych użytkowników danej instytucji, limitowany umową jest również okres dostępu do platformy książek. Przy wprowadzaniu danych do lokalnej bazy ważna jest zatem możliwość szybkiego i łatwego ich wyselekcjonowania i usunięcia wraz z terminem wygaśnięcia umowy subskrypcyjnej. Równie ważna z organizacyjnego punktu widzenia jest dostępność pożądaných danych bibliograficznych z wewnętrznych baz danych, które łatwo można włączyć do lokalnego katalogu. Wprowadzenie przez zespół bibliotekarzy w krótkim czasie dziesiątek tysięcy rekordów jest nierealne ze względu na pracochłonność całego procesu⁶. Dodatkową kwestią jest badanie aktualności linków umieszczonych do elektronicznych książek.

IBUK Libra kompletny

Biblioteka Uniwersytecka w Toruniu wraz z Biblioteką Collegium Medicum UMK w Bydgoszczy kupuje od kilku lat dostęp do wybranych książek (niekiedy kolekcji lub podkolekcji) IBUK Libra. Obecnie uprawnieni użytkownicy UMK posiadają dostęp do niemal tysiąca publikacji. Liczba dostępnych dokumentów ulega ciągłej zmianie. Biblioteka Uniwersytecka w Toruniu dokupuje wskazane przez jej czytelników książki umieszczone na platformie, natomiast Wydawnictwo Naukowe PWN sukcesywnie zwiększa liczbę darmowych publikacji włączając je do kolekcji subskrybentów. Wszystkie książki, do których ma dostęp UMK, posiadają linki w rekordach bibliograficznych w lokalnej bazie. Do tekstu publikacji kierują one bezpośrednio czytelników łączących się z terenu uniwersytetu, a tych spoza uczelni poprzez autoryzowany uniwersytecki punkt logowania. Obecnie czytelnicy mogą odszukać w katalogu biblioteki ponad 1100 rekordów z linkami do książek na platformie IBUK Libra.

Ze względów organizacyjnych przyjęto zasadę, że linki dodawane są automatycznie do opisów dla publikacji w wersji drukowanej, na podstawie numeru ISBN i roku wydania. Natomiast dla książek, dla których nie ma odpowiednich

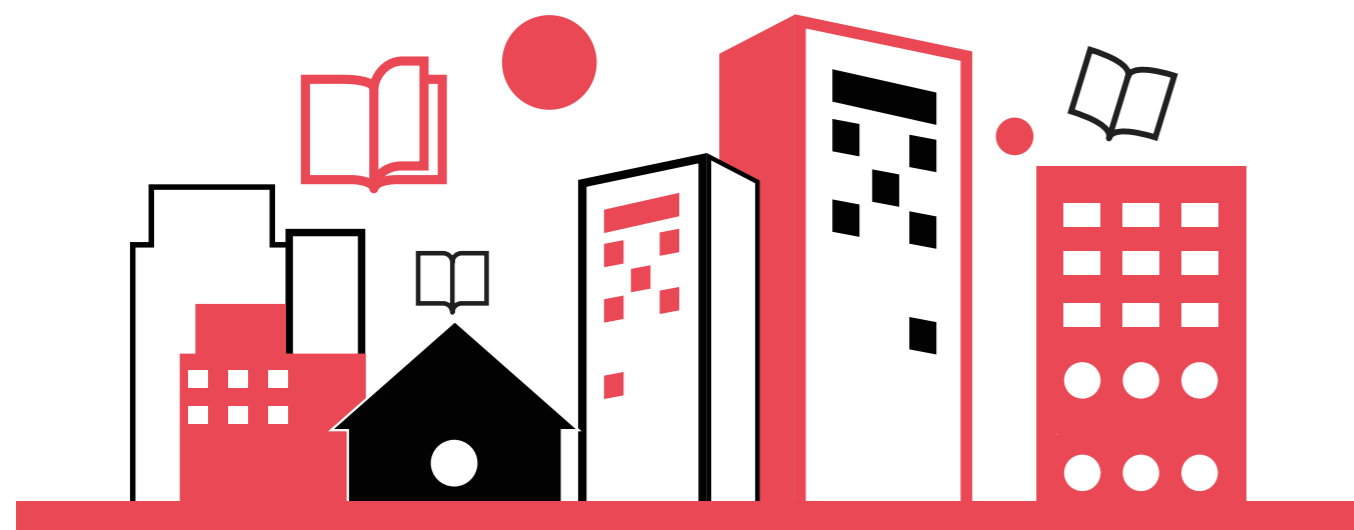
metadanych w bazie lokalnej ani w bazie NUKAT tworzy się opis na podstawie dokumentu elektronicznego i podaje się hipertącza w polu lokalnym. Opisy te traktowane są jako dane tymczasowe, widoczne w bazie tak długo, jak oferowany jest dostęp do publikacji. Z tego względu nie tworzy się do nich pełnych kartotek hasel wzorcowych, ale indeksuje się jedynie w polach kontrolowanych przez hasła formalne. Biblioteka Uniwersytecka w Toruniu przy wyborze książek dostępnych na platformie IBUK Libra kieruje się głównie zapewnieniem studentom dostępu do najbardziej popularnych podręczników akademickich, wykorzystywanych w procesie dydaktycznym. Planowane jest uruchomienie usługi wysyłania automatycznych e-maili do czytelników, którzy zarezerwowali w systemie wypożyczony egzemplarz książki, informujących o możliwości skorzystania z jego wersji online.

EBSCO i inne bazy

Znacznie większe pod względem liczby oferowanych tytułów, zarówno e-czasopism, jak i e-książek, są bazy EBSCO. Biblioteka Uniwersytecka w Toruniu subskrybuje kilka z nich. Najliczniej reprezentowana w katalogu jest eBook Academic Complete Worldwide. W bazie lokalnej mamy odnotowanych prawie 140 tys. rekordów z hipertąciami do książek EBSCO. Dzięki temu na poziomie katalogu biblioteki czytelnik może odszukać dane dostępnych książek wraz z trybem dostępu do nich. Opisy dostarczane są OCLC na podstawie danych EBSCO. Część z e-książek, która ma odpowiednik drukowany, jest linkowana do opisu dla takiej wersji (opisy pochodzą z bazy NUKAT). Metadane wprowadzane do katalogu z OCLC wyróżniają się m.in. tym, że pola kontrolowane przez kartoteki hasel wzorcowych poprzedzane są znakiem *. System ten przyjęto, aby zabezpieczyć dane przed automatycznym, niekontrolowanym przypisywaniem prac niewłaściwym autorom. Ponadto opisy te tworzone są w oparciu o zewnętrzne zasady bibliograficzne. Nie powodują one jednak konfliktów w systemie, są dobrym sposobem automatycznego implementowania danych.

Podobnie jak w przypadku książek EBSCO, uwzględniane są również publikacje z baz, do których biblioteka ma dostęp w ramach licencji krajowej. Tytułem przykładu, do książek Springer mamy ponad 9 tys. zlinkowanych opisów bibliograficznych pochodzących z katalogu NUKAT, do DOAB ponad

6. W latach 2012–2015 działał Zespół ds. opracowania e-książek, powołany przez Konferencję Dyrektorów Bibliotek Akademickich Szkół Polskich. Ze względu na nikłe zainteresowanie ze strony bibliotek proponowanym zagadnieniem (brano pod uwagę opracowanie kolekcji e-książek Springer o dostępie ogólnopolskim), prace zakończono tylko teoretycznym podsumowaniem tematu. Dokument dostępny na stronie http://kangur.uek.krakow.pl/kdbasp/wp-content/uploads/2013/10/e_ksiazki.pdf



4 tys. Odnotowywane są również publikacje z bazy Infona (ok. 10 tys. w znacznym stopniu pokrywają się z zasobami Springera) czy baz czasopism: Akademska Platforma Czasopism, DOAJ.

Najnowszą bazą, do której dostęp widoczny jest z poziomu katalogu bibliotecznego, jest Legalis. Dla części książek i czasopism udało się dodać hiperłącza do właściwych opisów bibliograficznych z bazy NUKAT na podstawie uwag o ograniczeniu dostępu dla subskrybentów bazy. Trwają konsultacje z dostawcą bazy nad wypracowaniem najlepszej identyfikacji metadanych.

Katalog nie tylko lokalny

Współczesne narzędzia informatyczne dają wiele możliwości łączenia danych katalogowych z zasobami zewnętrznymi. Katalog Biblioteki Uniwersyteckiej w Toruniu zawiera trzy dodatkowe relacje: z bazą NUKAT, kartoteką VIAF oraz Księgarnią Kopernikańską.

Za pomocą numerów kontrolnych NUKAT, widocznych w opisach katalogowych, łatwo można przejść do jego odpowiednika w bazie centralnej. Ten sposób lokalizacji publikacji poza macierzystą biblioteką może służyć jako początek kwerendy dla wypożyczalni międzybibliotecznej. Odnośniki do międzynarodowej kartoteki VIAF są drugim „oknem na świat”. Dodatkową opcją proponowaną czytelnikom jest zakup książki w księgarni uniwersyteckiej. Statystyki wykorzystania tych odnośników są obiecujące.

Kilka uwag...

Analizując możliwości automatycznego pozyskiwania danych do dokumentów elektronicznych, dostrzeżono pewne, niezamierzone przez udostępniające je instytucje, ograniczenia techniczne. Zdarza się, że brakuje linku zewnętrznego w oferowanej bazie. Hiperłącze kieruje tylko ogólnie do portalu, a nie

do konkretnej publikacji (np. NASBI). W Academia.pl brak jest możliwości podania linku zewnętrznego z katalogu do publikacji w bazie, gdzie w przypadku pozycji z ograniczonym przez prawa autorskie dostępem czytelnik miał możliwość zarezerwowania sesji i otwarcia dokumentu na terminalu we współpracującej bibliotece. Wydawcy często w danych tabelarycznych, obok istotnych identyfikatorów do publi-

kacji ISSN i ISBN (dla wersji papierowej i elektronicznej), podają też ISSN do serii czy ISBN dla całości wydawnictwa wielotomowego, co stwarza problemy przy automatycznym dopisywaniu linków do katalogu. Wydawnictwo Springer używa zaś do nowych wydawnictw elektronicznych te same ISBN, jakie w latach 70. ubiegłego wieku stosowało w wydawnictwach papierowych.

Bazy wielu polskich wydawców nie stosują standardów autoryzacji typu Athens czy Shibboleth. Budowa tych portali często utrudnia dostęp przez serwery proxy.

W ostatnich latach łatwo dostrzec gwałtowny wzrost wykorzystywania nowych narzędzi wyszukiwawczych, które zaczynają w polskich bibliotekach zastępować wyszukiwanie w tradycyjnych katalogach komputerowych. Wraz z rozszerzeniem oferty publikacji elektronicznych, jak również coraz większą popularnością tej formy prezentacji dorobku naukowego, biblioteki akademickie coraz częściej decydują się na zakup multiwyszukiwarek typu Summon, EDS, itp. Biblioteka Uniwersytecka w Toruniu nadal – podobnie jak część bibliotek, m.in. Biblioteka Jagiellońska – umożliwia czytelnikom korzystanie z tradycyjnego OPAC-a oraz multiwyszukiwarki (Summon od 2013). Efektywne wyszukiwanie źródeł w kilkudziesięciu bazach książek i czasopism, którymi dysponuje nowoczesna biblioteka, zapewnić może jedynie narzędzie przeszukujące wszystkie bazy dokumentów elektronicznych biblioteki. Czy zatem umieszczanie informacji o dostępie elektronicznym w tradycyjnym katalogu biblioteki ma nadal uzasadnienie? Częściowej odpowiedzi udzielają sami czytelnicy. W lutym 2016 roku odnotowano ponad trzy tysiące wyświetleń książek elektronicznych z poziomu katalogu biblioteki. Co ciekawe, na przestrzeni ostatnich kilku miesięcy można

zaobserwować powolny wzrost wykorzystania tej ścieżki dostępu. Prawie 30% z nich kierowało do zasobów EBSCO, około 10% do IBUK Libra, ponad 6% do KPBC. Jeśli porównamy statystyki wejść bezpośrednio z katalogu biblioteki do np. subskrybowanych zasobów IBUK Libra, możemy się pokusić o zmierzenie efektywności linkowania e-książek

w OPAC-u. I tak, dla ponad 6 tys. sesji do IBUK Libra w okresie od grudnia 2015 r. do końca lutego 2016 r. ponad 2 tys. wejść pochodziło z katalogu biblioteki. W tym samym przedziale czasowym do baz EBSCO również odnotowano podobną liczbę wejść. Przy ocenie statystyk nie można zapominać, iż znakomita większość informacji o zasobach elektronicznych jest importowana automatycznie do katalogu Biblioteki Uniwer-

syteckiej w Toruniu i nie generuje dodatkowych prac zespołu bibliotekarzy. Warto też zauważyć, że współpraca z Centrum NUKAT z zakresie udostępniania potrzebnych danych spotyka się zawsze z pełnym zrozumieniem i otwartością.

Zakładając, iż w najbliższym czasie nastąpi przejęcie przez multiwyszukiwarkę całości wyszukiwań w zasobach biblioteki, wydaje się, że warto jednak kontynuować dodawanie linków zdigitalizowanych książek do bazy lokalnej oraz NUKAT. Dzięki temu rekordy pobierane z centralnej bazy pojawiają się w katalogach innych bibliotek naukowych w kraju, zwiększając wykorzystanie KPBC i promując jej zasoby. W kontekście przyjęcia przez Centrum NUKAT zasad tworzenia rekordu analitycznego dla dokumentów niesamoistnych wydawniczo otwiera się szansa na promocję dokonań lokalnego środowiska naukowego w przestrzeni publicznej. Biblioteka Uniwersytecka w Toruniu otworzyła w 2012 roku repozytorium materiałów naukowych Rum@k. Obecnie zdeponowano tam

prawie 3 tys. tekstów. Utworzenie rekordów analitycznych dla dokumentów niesamoistnych wydawniczo, przestanie ich do NUKAT wraz z hiperłączem prowadzącym do repozytorium z pewnością wpłynie na większą widoczność prac naukowych Uniwersytetu Mikołaja Kopernika. Przejmowanie przez WorldCat rekordów bibliograficznych (również tych z polem dostępu elektronicznego) jest nie mniej istotne. NUKAT bowiem pełni również rolę agregatora danych o elektronicznych zasobach naukowych udostępnianych w Polsce. Rekordy z linkami do lokalnych zasobów cyfrowych są reprezentowane w światowym katalogu, niewątpliwie wzmacniając informacje o osiągnięciach krajowej nauki.

*Artykuł powstał w oparciu o dane dostarczone przez dra Wojciecha Sachwanowicza. Autorzy bardzo dziękują za pomoc i merytoryczną konsultację.

