

Agnieszka Dwornik
Biblioteka Uniwersytecka w Toruniu
agnieszka.dwornik@umk.pl

Dystrybucja metadanych bibliograficznych na świecie – źródła, sposoby pozyskiwania, problemy (wybrane aspekty praktyczne)

Streszczenie: Celem artykułu jest przegląd wybranych aspektów współczesnej dystrybucji metadanych bibliograficznych. Dane mogą powstawać lokalnie jako *original cataloguing* lub być wytwarzane w podmiotach zewnętrznych oraz pozyskiwane na drodze *copy cataloguing* – darmowej wymiany w ramach współkatalogowania lub też jako płatna subskrypcja i zakup. Przedstawiono zagadnienia dotyczące opisu bibliograficznego oraz jego jakości. Aspekty haseł wzorcowych i charakterystyki przedmiotowej wykraczają poza ramy tego artykułu.

Słowa kluczowe: metadane, katalogowanie, katalogi biblioteczne, wymiana danych

Wprowadzenie

Rozwój narzędzi informatycznych we współczesnym świecie dyktuje ciągłą metamorfozę oferowanych usług w coraz bardziej dynamiczny sposób. Potencjał dzisiejszych katalogów bibliotecznych jako źródła informacji znacznie różni się nie tylko od tradycyjnych katalogów kartkowych sprzed 50 lat, ale również od tych z początku komputeryzacji. Ich funkcja nie ogranicza się wyłącznie do prezentacji zasobów danej biblioteki. Katalogi są też źródłem informacji o tym, czy w bibliotece znajdują się dokumenty o podobnej tematyce, czy są one dostępne w innych formach, czy dane źródło jest dostępne fizycznie, czy instytucja proponuje dostęp elektroniczny. W przeciwieństwie do baz innego typu (np. służących administracji), które mogą być zbiorami zamkniętymi z lokalnymi schematami i regułami ich tworzenia, biblioteczne bazy danych powinny być otwarte, gotowe na ponowne wykorzystanie i tworzone przede wszystkim z myślą o użytkownikach.

Powyższe aspekty zmieniającej się rzeczywistości skłoniły międzynarodową społeczność bibliotekarską do rewizji zasad katalogowania i opracowania dokumentu *International Cataloguing Principles (ICP)*¹. Różni się on zasadniczo podejściem do danych od poprzednio wypracowanego dokumentu, tzw. Zasad paryskich. Nowe zasady skupiają się na użytkowniku i nowoczesnych narzędziach wyszukiwawczych. *Pośród dziewięciu zasad tworzenia przepisów katalogowania (...) takich jak: wygoda użytkownika, powszechność użycia, reprezentatywność, trafność, wystarczalność i konieczność, znaczenie, ekonomia, spójność i standaryzacja oraz integracja, które przyświecały autorom tego dokumentu, najważniejszą rolę przypisali komfortowi użytkownika*². Dynamicznie zmieniające się otoczenie sieciowe zmusza do aktualizacji

¹ *Deklaracja międzynarodowych zasad katalogowania* [online]. IFLA, 2009. [Dostęp 4.03.2018]. Dostępny w: https://www.ifla.org/files/assets/cataloguing/icp/icp_2009-pl.pdf.

² KRYNICKA, M. *Międzynarodowe zasady katalogowania: nowe zasady katalogowania oraz zmiany w terminologii, czyli nowoczesne podejście do katalogów* (preprint) [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.bn.org.pl/download/document/1329232478.pdf>.

dokumentu. W 2016 r. wydano jego kolejną wersję, poszerzoną o zagadnienia nowych kategorii użytkowników, środowiska open access, interoperacyjności i dostępności danych, przyszłości narzędzi do obsługi danych bibliograficznych. Ponownie podkreślono znaczące zmiany potrzeb i zachowań użytkowników³.

Innym dokumentem, który ma niebagatelny wpływ na kierunki rozwoju katalogowania i kooperacji bibliotek w zakresie tworzenia i wymiany danych, jest raport grupy roboczej powołanej przez Bibliotekę Kongresu na temat przyszłości kontroli bibliograficznej *On the record: report of the library of Congress Working Group on the future of bibliographic control*⁴. W raporcie zalecono zwiększenie wydajności poprzez ściślejszą współpracę i wymianę danych między instytucjami oraz skierowanie wysiłków na udostępnienie danych o rzadkich i unikalnych zbiorach bibliotek. Zwrócono uwagę na rozwój technologii, podkreślając, iż należy brać pod uwagę nie tylko nowe otoczenie sieciowe i zmiany technologiczne, ale także zmiany związane z użytkownikami metadanych, którymi są nie tylko ludzie. Dane powinny być czytelne również dla aplikacji komputerowych, które będą je analizowały i łączyły z danymi z innych źródeł. Wskazano też na potencjał schematu FRBR oraz potrzebę poszerzania umiejętności i kompetencji bibliotekarzy.

Jakie są więc współczesne metody zapewniania takiej jakości danych bibliograficznych, aby były interoperacyjne? Skąd dane te mogą pochodzić? Jak z nimi pracować, aby spełniały wymagania użytkowników i nowych narzędzi informatycznych?

Katalogowanie oryginalne

W literaturze angielskiego obszaru językowego można zauważyć silną opozycję dwóch sposobów pozyskiwania danych do katalogów, tzw. *original cataloguing* i *copy cataloguing*. W ostatnich latach obszar „oryginalnego katalogowania” podlega szybkim zmianom. Po latach dyskusji nad modelem konceptualnym FRBR do katalogów światowych wchodzi schemat RDA, który jest implementowany w stopniu pełnym lub częściowo. Pełną realizację schematu ogłosiły w 2013 r. m.in. Biblioteka Kongresu (Library of Congress, LC), Biblioteka Brytyjska (British Library, BL), Biblioteka Narodowa Australii (National Library of Australia, NLA). Częściowo lub w całości wdrażają je biblioteki fińskie, holenderskie, niemieckie, łotewskie czy czeskie. Podobnie dzieje się w Polsce – Centrum NUKAT planuje pełną realizację RDA⁵, Biblioteka Narodowa deklaruje wprowadzenie części elementów schematu⁶. Zmianom podlegają również języki informacyjno-wyszukiwawcze w kierunku większej „atomizacji” informacji, użyteczności dla wyszukiwania fasetowego⁷.

³ Tamże.

⁴ *On the record: report of the library of Congress Working Group on the future of bibliographic control* [online]. 2008. [Dostęp 4.03.2018]. Dostępny w: <https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>.

⁵ Centrum NUKAT <http://centrum.nukat.edu.pl/pl/>. Wszystkie odesłania do stron internetowych przedstawiają wersję aktualną w dn. 4.03.2018 r.

⁶ Zob. nagranie z konferencji „Standard RDA – korzyści i problemy związane z jego wdrożeniem”: CICHONÍ, M. Elementy RDA stosowane w opisie bibliograficznym w ramach zmian wprowadzanych przez Bibliotekę Narodową. W: *Grupa Robocza ds. Standardów Bibliograficznych* [online]. Centrum NUKAT, 14.10.2016 r. [Dostęp 4.03.2018]. Dostępny w: <http://centrum.nukat.edu.pl/rda/>.

Jedną z ważniejszych kwestii podnoszonych w związku z rozszerzaniem użyteczności danych, a co za tym idzie, zwiększaniem zawartości rekordów z metadanymi bibliograficznymi, jest koszt ich wytworzenia. Coraz dynamiczniej zmieniają się wymagania co do kompetencji bibliotekarzy zajmujących się pracą z metadanymi. O ile do tej pory znajomość formatu, według którego w danej bibliotece opracowywano informację na temat gromadzonych zasobów była umiejętnością wystarczającą, to w dobie zintegrowanych systemów bibliotecznych, a za chwilę coraz bardziej powszechnych systemów nowej generacji tzw. systemów chmurowych, zmieniają się wymagania stawiane katalogerom⁸. Wśród najistotniejszych wymienia się: zrozumienie zmian technologicznych, znajomość różnych schematów metadanych, rozumienie zasad indeksowania, struktury baz danych, zasad wymiany danych.

Głównymi dostawcami wytwarzanych w ten sposób danych bibliograficznych i udostępnianych do ponownego powszechnego wykorzystania są tradycyjnie biblioteki narodowe. Jest to związane z ich statutowymi zadaniami, z których jednym z najważniejszych, obok gromadzenia i archiwizacji oraz ochrony dorobku piśmienniczego danego obszaru, jest jego dokumentacja. Udostępnianie tych danych różni się w zależności od wypracowanych rozwiązań narodowych. Część instytucji udostępnia je powszechnie i za darmo (np. Biblioteka Narodowa w Warszawie), inne dystrybuują dane odpłatnie (np. Biblioteka Kongresu).

Współkatalogowanie

Od wielu lat rozwija się kooperacja w zakresie wytwarzania metadanych bibliograficznych, często realizowana przez konsorcja biblioteczne tworzone właśnie z potrzeby wymiany informacji i optymalizacji kosztów działania bibliotek. Funkcjonują one na zasadzie utrzymania wspólnej bazy do wymiany danych dla bibliotek podobnego typu (np. francuski SUDOC <http://www.sudoc.abes.fr/>, prowadzony przez Agence bibliographique de l'enseignement supérieur, ABES), wykorzystywania jednego systemu bibliotecznego zakupionego wspólnie (np. LIBISnet <http://libis.be/libis/libisnet> w Belgii, wspólny katalog w systemie ALMA) bądź dobrowolnego zrzeszenia grupy bibliotek (np. katalog NUKAT <http://katalog.nukat.edu.pl/>), które działają na zasadzie katalogowania kooperatywnego (*cooperative cataloguing*), gdzie wszystkie biblioteki uczestniczące w projekcie mają takie same lub zbliżone uprawnienia do wprowadzania danych. W ostatnich latach podobne rozwiązania stosowane są również przez niektóre biblioteki narodowe. Korzystają one z możliwości współpracy z innymi instytucjami w opisywaniu dorobku piśmienniczego zgodnie z zakresem terytorialnym językowym oraz terytorialnym właściwym dla bibliotek narodowych. Jest to zasadne tym bardziej, że w dobie elektronicznych baz rozróżnienie na bazy typowo bibliograficzne i bazy katalogowe ze względu na wymagania użytkowników staje się mało użyteczne⁹, chociaż nie

⁷ SKÓRKA, S. Fasety na nowo odkryte: integrowanie systemów nawigacji i integrowania informacji. *Zagadnienia Informacji Naukowej* 2014, nr 2, s. 92–109. ISSN 0324-8194.

⁸ Przykładem opisu takich wymagań jest dokument *Core Competencies for Cataloguing and Metadata Professional librarians: approved by the ALCTS Board of Directors on January 23, 2017* [online]. January 23, 2017. [Dostęp 4.03.2018]. Dostępny w: http://www.lac.org.tw/files/core_competencies_cataloging_metadata_professional.pdf.

⁹ SIWECKA, D. *Światowy model informacji bibliograficznej: program i projekty (1952–2010)*. Wrocław [etc.]: Wydaw. SBP, 2015, s. 226. ISBN 978-83-64203-48-0.

niespotykane (np. Bibliographie nationale française <http://bibliographienationale.bnf.fr>). Model współkatalogowania przyjęty w kilku bibliotekach narodowych opiera się na tzw. katalogowaniu podzielonym (*shared cataloguing*).

Biblioteka Brytyjska prowadzi strategię współkatalogowania egzemplarza obowiązkowego pod nazwą *The Legal Deposit Shared Cataloguing Programme*¹⁰. Model ten opiera się na porozumieniu sześciu bibliotek, które są beneficjentami egzemplarza obowiązkowego w Wielkiej Brytanii i Irlandii. Poza Biblioteką Brytyjską w projekcie uczestniczą: The Bodleian Library, Cambridge University Library, The National Library of Scotland, The National Library of Wales oraz The Library of Trinity College w Dublinie. Celem takiej kooperacji jest zarówno dzielenie wysiłku opracowania, jak również lepsza kontrola bibliograficzna dorobku piśmienniczego obszaru Zjednoczonego Królestwa. Stosunek wkładu danych do British National Bibliography biblioteki narodowej i bibliotek współpracujących wynosi 70:30. Dane są zgodne z RDA i oparte o AACR2. Wszystkie te biblioteki są też zaangażowane w *Program for Cooperative Cataloging* prowadzony przez Bibliotekę Kongresu (szerzej o tej inicjatywie w dalszej części pracy). Rekordy bibliograficzne mogą być tworzone od podstaw przez biblioteki, mogą też być oparte na opisach stworzonych w ramach *Cataloguing-in-Publication Programme*¹¹.

Podobny rodzaj kooperacji w zakresie tworzenia metadanych bibliograficznych stosowany jest w Czeskiej Bibliotece Narodowej (Národní knihovna České republiky, NK ČR), która również wypracowała własny program współkatalogowania dokumentów przekazywanych w ramach egzemplarza obowiązkowego. W realizacji tych zadań czynnie uczestniczą Moravská zemská knihovna oraz Vědecká knihovna v Olomouci. Do budowania bibliografii narodowej wykorzystywane są też dane z innych bibliotek mających prawo do regionalnego egzemplarza obowiązkowego, których opisy bibliograficzne są przekazywane do czeskiego katalogu centralnego CASLIN <http://www.caslin.cz/>. Co miesiąc zawartość bazy jest porównywana z danymi księżnicy narodowej, a Česká národní bibliografie jest wzbogacana o opisy brakujących publikacji do czasu zastąpienia ich danymi wykonanymi przez jedną z trzech bibliotek biorących udział w programie kooperacji. Korzyścią takiej organizacji pracy jest wysoki stopień kompletności rejestracji publikacji wchodzących w zakres bibliografii narodowej, jak również wysoka jakość metadanych¹².

W Niemczech obecnie brak centralnego katalogu narodowego, a za gromadzenie i dokumentację bibliograficzną dorobku piśmiennictwa odpowiedzialna jest Niemiecka Biblioteka Narodowa (Deutsche National Bibliothek, DNB). Wymiana danych i współkatalogowanie odbywa się w na bazie sześciu regionalnych systemów bibliotecznych (Gemeinsamer Verbundkatalog, GBV, Kooperativer Bibliotheksverbund Berlin Brandenburg, KOBV, HBZ-Verbund Datenbank, Hessische BibliotheksInformationSystem, HeBIS, Südwestdeutscher Bi-

¹⁰ The Legal Deposit Shared Cataloguing Programme. W: *British Library Collection Metadata: The British National Bibliography* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.bl.uk/bibliographic/clscp.html>.

¹¹ Cataloguing-in-Publication Programme w Wielkiej Brytanii jest prowadzony przez prywatną firmę BDS, która jest dostarczycielem wysokiej jakości danych bibliograficznych. W: *BDS* [online]. [Dostęp 4.03.2018]. Dostępny w: <https://www.bdslive.com/?pid=index>.

¹² LICHTENBERGOVA, E. *The Czech National Bibliography: new steps to completeness* [online]. Submitted on 17.06.2015. [Dostęp 4.03.2018]. Dostępny w: <http://library.ifla.org/1128/1/089-lichtenbergova-en.pdf>.

bliotheksverbund, SWB oraz Bibliotheks Verbund Bayern, BVB). Ich zbiory oraz zasoby innych bibliotek można przeszukiwać przez Karlsruher Virtueller Katalog (KVK <https://kvk.bibliothek.kit.edu/>). Współpraca między nimi odbywa się na poziomie standaryzacji koordynowanej przez Niemiecką Bibliotekę Narodową¹³, co zapewnia jednolitość danych i ułatwia ich wymianę między bazami. Właściwym katalogiem centralnym opartym na katalogowaniu kooperatywnym w Niemczech jest baza czasopism ZDB¹⁴, którą współtworzy 3700 bibliotek z Niemiec i Austrii. Zawiera ona opisy bibliograficzne czasopism drukowanych i oferowanych w dostępie elektronicznym bez ograniczeń językowych, chronologicznych bądź geograficznych. Prace nad tą bazą koordynuje Biblioteka Państwowa w Berlinie (Staatsbibliothek zu Berlin – Preussischer Kulturbesitz) oraz Niemiecka Biblioteka Narodowa, udostępniając dane bezpłatnie wszystkim instytucjom na licencji CC0 1.0 Universal¹⁵. Jest to wyraz dbałości o powszechny dostęp do danych wysokiej jakości.

Kooperacja ważna jest nie tylko przy bieżącym współkatalogowaniu, ale również w modyfikowaniu danych, które powstały według wcześniej obowiązujących standardów. Biblioteka Kongresu prowadzi w tym celu *Program for Cooperative Cataloguing (PCC)*¹⁶. Członkami tego przedsięwzięcia są różne biblioteki lub pojedynczy katalogerzy, którzy wypracowują standardy oraz metody działania, przygotowują materiały szkoleniowe i modyfikują dane. Współpraca jest prowadzona w czterech obszarach:

1. *BIBCO* – podnoszenie jakości danych w opisach bibliograficznych książek poprzez dodawanie formalnych i przedmiotowych punktów dostępu, akceptowanie różnych klasyfikacji i języków haseł przedmiotowych celem zwiększenia użyteczności danych na całym świecie; w projekcie uczestniczy 79 bibliotek.
2. *CONSER* – problematyka wydawnictw ciągłych, źródło autorytatywnych opisów bibliograficznych czasopism.
3. *NACO* – zagadnienia ujednoczonych nazw osobowych, korporatywnych, serii i tytułów ujednoczonych.
4. *SACO* – współpraca w zakresie opracowania przedmiotowego; członkowie mogą proponować nowe hasła do LCSH, Library of Congress Classification (LCC).

Członkowie PCC w zależności od stopnia zaawansowania i posiadanych uprawnień biorą udział w projektowaniu, prezentowaniu i testowaniu nowych rozwiązań. Zajmują się również bieżącą modyfikacją i uzupełnianiem danych bazowych.

Pozyskiwanie danych bibliograficznych

Pobieranie danych z zewnętrznych źródeł nazywane jest w tradycji angloamerykańskiej *copy cataloguing* i stoi w opozycji do *original cataloguing*. W licznych bibliotekach pod względem ilościowym stanowi podstawowe źródło pozyskiwania danych bibliograficznych,

¹³ HUDZIK, K. *Biblioteki uniwersyteckie w Niemczech w epoce cyfrowej: ciągłość i zmiana*. Warszawa: Wydaw. SBP, 2017, s. 149. ISBN 978-83-65741-01-1.

¹⁴ *Zeitschriftendatenbank* [online]. Staatsbibliothek zu Berlin – Preussischer Kulturbesitz, 27.09.2017. [Dostęp 4.03.2018]. Dostępny w: <http://www.zeitschriftendatenbank.de/startseite/>.

¹⁵ Przekazanie do domeny publicznej [przyp. Red.].

¹⁶ Program for Cooperative Cataloguing. W: *Library of Congress* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.loc.gov/aba/pcc/>.

katalogowanie zaś oryginalne jest marginesem działalności. Instrukcje i procedury dotyczące tego zadania można znaleźć w serwisach internetowych bibliotek, co świadczy o dbałości o adekwatność danych. W celu zmniejszenia kosztów powszechnie jest delegowanie zadań na pracowników zwanych *paraprofessionals*¹⁷, którzy współpracują z katalogerami, a których rola zmienia się wraz z rozwojem nowych technologii. Czynności związane z pozyskiwaniem danych bibliograficznych w zintegrowanych systemach bibliotecznych można już rozpoczynać na etapie gromadzenia zbiorów. Pozyskane dane można następnie wykorzystywać w procesie katalogowania, co jest również zasadne z ekonomicznego punktu widzenia w przypadku retrokatalogowania.

Poza problematyką dostępności danych w środowisku sieciowym ważnym aspektem *copy cataloguing* jest techniczna możliwość ich pozyskiwania. Najpowszechniejszym sposobem kopiowania rekordów bibliograficznych do lokalnych katalogów jest protokół Z39.50, który został opracowany przez Bibliotekę Kongresu jeszcze w latach 70. XX w.¹⁸. Na podobnych zasadach działa protokół SRU (Search/Retrieval via URL) wykorzystywany m.in. w Niemieckiej Bibliotece Narodowej. W związku z szybkim rozwojem technologii informatycznych wiele baz oferuje kilka narzędzi do pobierania danych. Systemy chmurowe nowej generacji dysponują bardzo wygodnym narzędziem wbudowanym już w system biblioteczny. W ten sposób można pobierać opisy bibliograficzne z katalogów innych bibliotek, używających tego samego systemu lub z baz światowych podłączonych już przez dostawcę oprogramowania. Standardem jest przeszukiwanie i pobieranie danych w ten sposób z największego światowego katalogu WorldCat.

Crowdsourcing

Innym, bardzo interesującym z ekonomicznego i społecznego punktu widzenia, sposobem kompletowania i ujednolicania zawartości informacyjnej katalogów bibliotecznych jest crowdsourcing internetowy. Definiowany jest on jako *działalność heterogenicznych społeczności wirtualnych, angażująca ich wiedzę, umiejętności lub aktywa do wykonywania w Internecie (odpłatnie bądź bezpłatnie) zadań cechujących się pewną złożonością, zleconych w formie otwartego zaproszenia przez osobę prywatną, organizację, instytucję non profit bądź przedsiębiorstwo i przez nie nadzorowaną*¹⁹. W bibliotekach cechy takiej działalności ma angażowanie użytkowników w podnoszenie jakości katalogów poprzez zgłaszanie błędów w opisach bibliograficznych. Wykorzystywany jest w tym celu formularz typu „Zgłoś błąd” lub inne narzędzia do komunikacji umieszczone na stronie internetowej prezentującej katalog danej biblioteki. Potencjał crowdsourcingu może być wykorzystywany również do pozyskiwania metadanych wzorem projektu Biblioteki Brytyjskiej LibCrowds²⁰. Polega on na

¹⁷ RIDER, M.M. Developing new roles for paraprofessionals in cataloging. *The Journal of Academic Librarianship* [online]. 1996, Vol. 22, Iss.1, s. 26–32. [Dostęp 4.03.2018]. ISSN 0099-1333. Dostępny w: doi: [10.1016/S0099-1333\(96\)90031-2](https://doi.org/10.1016/S0099-1333(96)90031-2).

¹⁸ Z39.50 Resources. W: *Z39.50 International Standard Maintenance Agency* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.loc.gov/z3950/agency/resources/>.

¹⁹ KOWALSKA, M. *Crowdsourcing internetowy: pozytywny wymiar partycypacji społecznej*. Warszawa: Wydaw. SBP, 2015, s. 144–145. ISBN 978-83-64203-60-2.

²⁰ Crowdsourcing projects from the British Library. W: *LibCrowds* [online]. [Dostęp 4.03.2018]. Dostępny w: <https://www.libcrowds.com>.

zaangażowaniu społeczności internetowej do przenoszenia informacji z zeskanowanych kart katalogowych zawierających opisy zarówno książek, jak i czasopism kolekcji azjatyckich w języku chińskim i afrykańskich w języku urdu, do katalogu komputerowego. Zarejestrowani wolontariusze na podstawie skanów kart katalogowych wyszukują odpowiednie dane w bazie rekordów pobranych z WorldCat i dodają sygnatury do zasobu. Aby wykluczyć błędy, przyjęto zasadę, że każda karta sprawdzana jest przez trzy osoby. Jeżeli co najmniej dwóch wolontariuszy wskaże ten sam rekord bibliograficzny z bazy OCLC, dane importowane są do katalogu biblioteki narodowej. Karty z opisami bibliograficznymi, dla których nie dopasowano metadanych są przekazywane do ręcznego ich wprowadzenia. Sposób pozyskiwania danych w modelu crowdsourcingu jest bardzo obiecujący zwłaszcza w zakresie publikacji o szczególnym charakterze bądź wymagających rzadkich umiejętności (np. znajomość zbiorów specjalnych lub egzotycznych języków).

Usługi komercyjne

Dane bibliograficzne najczęściej dostępne są w sieci bezpłatnie w sposób otwarty lub po nawiązaniu formalnej współpracy. Odmiennym modelem jest sposób oferowany przez Bibliotekę Kongresu, która w celu zwrotu ponoszonych kosztów świadczy subskrypcję płatną²¹. Nie jest to zatem działalność klasycznie komercyjna, nastawiona na wypracowanie zysków. Cena dostępu jest zróżnicowana w zależności od rodzaju nabywanych danych, np. od rodzaju nabywanych danych, np. opisów bibliograficznych książek we wszystkich językach lub tylko w języku angielskim, od rodzaju publikacji, zestawu metadanych itd. Tradycyjną i najpowszechniejszą praktyką jest wytwarzanie lub pobieranie danych bibliograficznych w oddziałach opracowania zbiorów w bibliotekach. Outsourcing tych zadań nie jest jeszcze powszechny. Można jednak przypuszczać, że wobec zmieniających się technologii i wymagań co do jakości danych oraz konieczności ekonomicznego wykorzystania posiadanych już przez biblioteki zasobów informacyjnych, będzie się on rozwijał. Dla potrzeb tego artykułu porównano ofertę usług firm Backstage Library Works²² oraz Casalini Libri²³. Bardzo różnią się one kompletnością dostarczanych serwisów. Backstage Library Works jest firmą obecną na rynku od 25 lat, oferującą szeroki pakiet usług związanych z wieloma aspektami działalności bibliotek: od digitalizacji, poprzez ochronę mikrofilmów, inwentaryzację po dostarczanie i zarządzanie metadanymi i optymalizację baz katalogowych. Firmie można dostarczyć pliki danych pochodzących np. z modułu gromadzenia w formacie MARC, ale również w postaci plików tekstowych czy arkuszy kalkulacyjnych, które zostaną opracowane zgodnie ze standardami, jakich wymaga biblioteka (opis skrócony, pełny, wg AACR2 lub RDA). Dane bibliograficzne pozyskiwane są podobnie jak dzieje się to w bibliotekach – za pomocą porównania i kopiowania z dostępnych baz bibliograficznych (katalog Biblioteki Kongresu, WorldCat, RLUK). W przypadku braku opisów w wyżej wymienionych bazach dokumenty są katalogowane prymarnie. Dodatkową usługą dostawcy jest opcja przesyłania danych w imieniu zamawiającej je biblioteki do wybranych baz lub katalogów konsorcyjnych. Firma oferuje obsługę danych również dla wydawców.

²¹ Cataloging Distribution Service. W: *Library of Congress* [online]. [Dostęp 4.03.2018]. Dostępny w: <https://www.loc.gov/cds/>.

²² *Backstage Library Works* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.bslw.com>.

²³ *Casalini Libri* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.casalini.it>.

Z kolei Casalini Libri od lat 60. specjalizuje się w dostarczaniu metadanych dla publikacji europejskich zwłaszcza na potrzeby bibliotek akademickich (od 2013 r. w standardzie RDA²⁴). W ofercie ma również wsparcie dla projektów digitalizacyjnych oraz dla wydawców. Obie firmy inwestują w rozwój najwyższej jakości metadanych. Casalini Libri podjęła współpracę z inicjatywą BIBFRAME, by być gotową na wprowadzenie Linked Data do swojej oferty. Backstage Library Work eksperymentuje z nowymi technologiami, m.in. wzbogacając autorytatywne punkty dostępu o adresy URL.

Innym źródłem danych, które mogą być widoczne w katalogach bibliotecznych są dane dostarczane przez dostawców kolekcji zasobów elektronicznych. Dane wraz z publikacjami wpływają masowo do baz bibliotecznych, dostępny zasób podlega częstej aktualizacji jednocześnie ze zmianą oferty dostawców oraz zmianą zapotrzebowania bibliotek i ich użytkowników. Stanowią coraz większą część oferowanych, zwłaszcza przez biblioteki naukowe, źródeł informacji. W zależności od przyjętych lokalnie rozwiązań mogą być one udostępniane użytkownikom jako osobne bazy lub włączane do katalogów. W dobie systemów bibliotecznych nowej generacji, których jedną z cech jest obsługa nie przez OPAC, a przez multiwyszukiwarki obejmujące wszystkie zbiory w danej instytucji, są bardzo ważnym źródłem danych bibliograficznych. Relatywnie niższa ich zawartość informacyjna, niedokładność i niespójność z innymi danymi katalogowymi może powodować niższe ich wykorzystanie, a co za tym idzie nieefektywność zainwestowanych w nie środków. Różnorodność rodzajów dokumentów w wersji elektronicznej (artykuły, książki, czasopisma, bazy danych i in.) i ich wzrastająca liczba powodują, że nie mogą podlegać bezpośredniej kontroli w zakresie metadanych przez specjalistów zatrudnionych w bibliotekach. W tym kontekście z zainteresowaniem należy się pochylić również nad polityką takich firm jak Springer, Elsevier czy Wiley.

Dyskusja nad tym zagadnieniem rozpoczęła się w środowisku bibliotekarskim w roku 2012²⁵. O objęcie formalnym pośrednictwem kontaktów między wydawcami, producentami agregatorów wiedzy oraz sprzedawcami baz wiedzy a bibliotekarzami zwrócono się do OCLC. Pod patronatem tej instytucji została powołana E-Data Quality Working Group, która w krótkim czasie opracowała dokument *Success Strategies for Electronic Content Discovery and Access: a Cross Industry White Papers*²⁶. Sprecyzowano w nim problemy, które zostały ujęte w trzech obszarach: niewystarczające i niekompletne dane bibliograficzne, nierównoczesna dystrybucja danych bibliograficznych i zasobów oraz dystrybucja danych w wielu różnych formatach. Przedstawiono równocześnie rekomendacje postępowania dla dostawców danych. W zakresie danych bibliograficznych były to: stosowanie formatu MARC, rozróżnienie ISBN dla wydań drukowanych i elektronicznych, opatrywanie dokumentów spójnymi i adekwatnymi metadanymi, korekta danych. Kolejna analiza i ocena serwisów wykazała znaczące podwyższenie jakości. Wydawcy oraz firmy zajmujące się dostarczaniem dokumentów elektronicznych zdały sobie sprawę, że wysokiej jakości me-

²⁴ ALDI, G. RDA implementation at Casalini. *JLIS.it.: Italian Journal of Library* [online]. 2018, Vol. 9, No. 1, s. 24-26. [Dostęp 4.03.2018] Dostępny w doi: [10.4403/jlis.it-12434](https://doi.org/10.4403/jlis.it-12434).

²⁵ RUSCHOFF, C. i in. E-data quality: how publishers and libraries are working together to improve data quality. *Collaborative Librarianship* [online]. 2016, Vol. 8, Iss. 4, Art. 7. [Dostęp 4.03.2018]. ISSN 1943-7528. Dostępny w: <https://digitalcommons.du.edu/collaborativelibrarianship/vol8/iss4/7>.

²⁶ *Success Strategies for Electronic Content Discovery and Access : a Cross Industry White Papers* [online]. Dublin, Ohio: OCLC Online Computer Library Center, 2014. [Dostęp 4.03.2018]. Dostępny w: <https://www.oclc.org/content/dam/oclc/reports/data-quality/215233-SuccessStrategies.pdf>.

tadane mają kolosalny wpływ na dostęp i wykorzystanie subskrybowanych przez biblioteki dokumentów. To zaś ma niebagatelny wpływ na dobór dostawców i kolekcji przez biblioteki w przyszłości. Porozumienie między wydawcami a bibliotekami można uznać za sukces, gdyż współpraca nawiązywana jest z następnymi dostawcami treści²⁷.

Kilka uwag na temat dystrybucji danych bibliograficznych

Na marginesie przeglądu możliwości pozyskiwania danych bibliograficznych rodzi się kilka pytań.

Dane bibliograficzne produkowane są w wielu agendach: w bibliotekach, agencjach bibliograficznych, w katalogach centralnych opartych na współkatalogowaniu. Dane są dostępne, można je importować, unikając dublowania pracy przy tworzeniu metadanych. Dostawcy systemów bibliotecznych oferują coraz prostsze sposoby pobierania danych z baz zewnętrznych. Rodzi się pytanie, czy zasadne i możliwe jest stosowanie jednego opisu bibliograficznego dla różnych rodzajów bibliotek, dla użytkowników z odmiennymi potrzebami informacyjnymi.

Pytania te postawiła sobie Kara Allison²⁸. Zbadła ona zmiany, jakie nanoszone są w rekordach bibliograficznych dostarczonych przez OCLC w katalogach lokalnych. Do badania wytypowano biblioteki akademickie i publiczne z sześciu krajów anglojęzycznych (Nowa Zelandia, Australia, Stany Zjednoczone, Australia, Wielka Brytania, Kanada i Południowa Afryka). Stwierdzono, że dane dostarczone przez OCLC były o wiele rzadziej modyfikowane przez biblioteki akademickie niż przez biblioteki publiczne. Uzupełnienia czynione przez pierwszą grupę najczęściej dotyczyły dodawania odnośników do elektronicznych form publikacji. Biblioteki publiczne o wiele częściej modyfikowały dane, co należy wiązać z innymi potrzebami użytkowników. Zmiany dotyczyły sfery *user-centre dediting* (związane z potrzebami lokalnego użytkownika) – tu szczególnie zmian haseł przedmiotowych oraz *enriched content* (wzbogacanie zasobu) najczęściej poprzez automatyczne dodawanie linków do płatnych baz elektronicznych oraz do innych e-źródeł oraz mediów społecznościowych. Powoduje to, że współczesne katalogi coraz bardziej oddalają się od powielania funkcji kart katalogowych, a stają się coraz bardziej powiązane z zasobami zewnętrznej sieci informacyjnej. Często również dodawano informacje na temat książek o podobnej tematyce czy o innych publikacjach danego autora. Tego rodzaju dodatki zastępują możliwości, które dawałoby zaimplementowanie FRBR. Modyfikacje dotyczące poprawności metadanych (*quality aspects*) były nieliczne. Badanie to wykazało, że jeden uniwersalny rekord bibliograficzny dla wszystkich może być fikcją, ale zauważono również, że rozwiązaniem są narzędzia nowej generacji i nowe schematy danych.

²⁷ OCLC signs agreements with leading publishers worldwide to make more content discoverable through libraries. W: OCLC [online]. Dublin, Ohio, 30 January 2018. [Dostęp 4.03.2018]. Dostępny w: <https://www.oclc.org/en/news/releases/2018/201803dublin.html>.

²⁸ KARA, A. Global records, local catalogues : investigating local editing practices in copy cataloguing. *New Zeland Library & Information Management Journal* [online]. 2014, Vol. 55, No. 1. [Dostęp 4.03.2018]. ISSN 1177-3316. Dostępny w: <https://lianza.org.nz/global-records-local-catalogues-investigating-local-editing-practices-copy-cataloguing>.

Problemem, który dotyczy innych niż anglojęzyczne bazy bibliograficzne jest to, że dostęp do danych we współczesnych katalogach organizowany jest w jednym punkcie – tak jak wcześniej wspomniano, multiwyszukiwarki zastępują dotychczasowy OPAC biblioteczny. W jednym punkcie wyszukiwane więc są (lub będą) dane produkowane przez lokalnych bibliotekarzy oraz dane dostarczane przez dostawców kolekcji elektronicznych. Są one w znakomitej większości opatrywane aparatem wyszukiwawczym w języku angielskim (słowa kluczowe, hasła przedmiotowe). W bazach stosujących aparat informacyjno-wyszukiwawczy w językach narodowych fakt taki będzie implikował konieczność powtarzania przeszukiwania lokalnych źródeł informacji według metadanych „narodowych” oraz dostarczonych przez firmy tworzące bazy w języku angielskim. Dodatkową komplikacją może być również fasetowanie według kryteriów formalnych. U progu czekających nas zmian związanych z implementacją tzw. systemów chmurowych warto obserwować, jak te dane mogłyby się integrować, dając czytelnikowi jak najbardziej pełną odpowiedź na zapytanie wyszukiwawcze.

Dystrybucja danych bibliograficznych w Polsce

W Polsce wymiana danych bibliograficznych w zasadniczych kwestiach wygląda podobnie jak na świecie. Biblioteki bardzo często stosują kombinację *original cataloguing* i *copy cataloguing*, nazywając procesy te katalogowaniem i pobieraniem lub ściąganiem opisów. Mamy do dyspozycji dwie wielkie bazy katalogowe, z których można bezpłatnie pobierać dane: katalog Biblioteki Narodowej (tworzony samodzielnie) oraz katalog NUKAT (tworzony na zasadzie współkatalogowania). Innym katalogiem, o którym warto wspomnieć jest Centralny Katalog Federacji Bibliotek Kościelnych FIDES (<http://katalog.fides.org.pl>), który jest tworzony na zasadzie współkatalogowania w programie KOHA. Możliwe jest również pobieranie danych z lokalnych katalogów bibliotek dostępnych poprzez wyszukiwarkę KARO (<http://karo.umk.pl/Karo/>). W poprzednich latach użytkownicy systemu Horizon praktykowali wymianę opisów bibliograficznych w ramach jednego systemu bibliotecznego. Zakup danych bibliograficznych nie rozwinął się w Polsce.

Zakończenie

Trudno wyrokować, jak szybko będą następowały kolejne zmiany w obszarze danych bibliograficznych i jaki to będzie miało wpływ na ich dystrybucję. Idea FRBR ma już 20 lat, a nadal nie jest powszechnie stosowana w katalogach w formie klasycznej. Pozostaje jako zastosowania eksperymentalne²⁹. O wizję przyszłości technologii związanych z udostępnianiem informacji przez biblioteki w swoim artykule pokusił się Michael Keeble BUCKLAND³⁰. Trudno się nie zgodzić, że zasoby informacyjne bibliotek na całym świecie w coraz większym stopniu nie są budowane lokalnie przez samych bibliotekarzy, którzy kiedyś do-

²⁹ Na przykład wyszukiwarka książek kucharskich CookBook Finder zob. CookBook Finder [online]. W: OCLC. [Dostęp 4.03.2018]. Dostępny w: <https://www.oclc.org/research/themes/data-science/cookbook-finder.html>.

³⁰ BUCKLAND, M.K. Library technology in the next twenty years. *Library Hi Tech* [online]. 2017, Vol. 35, Iss. 1, s. 5–10. [Dostęp 4.03.2018]. ISSN 0099-1333. Dostępny w doi: [10.1108/LHT-11-2016-0131](https://doi.org/10.1108/LHT-11-2016-0131).

stosowywali kryteria dostępu do zbiorów lokalnych. Obecnie i w przyszłości udostępnianie informacji o zbiorach będzie polegać na pozyskiwaniu metadanych ze źródeł zewnętrznych. Katalogi lokalne coraz powszechniej będą zawierały dane z repozytoriów instytucji lub kolekcji wydawców, a tym samym będą dostarczać bardzo zróżnicowane metadane. Dane te opatrywane są w aparat wyszukiwania przedmiotowego, który może być znany tylko wąskiej grupie doświadczonych, wyspecjalizowanych czytelników i bibliotekarzy. Wiodący jest silny trend w kierunku dużych i zróżnicowanych zbiorów danych. Dostęp do nich zasadniczo nie różni się od innych źródeł informacji, ale opisy bibliograficzne są dla nich mniej rozwinięte (np. mikropublikacje typu blogi). Zadaniem bibliotekarzy w przyszłości będzie zapewnienie dostępu nie tylko do znanej, lokalnej kolekcji, ale również pośredniczenie między bardzo zróżnicowanymi jednostkami o niejednorodnej strukturze danych opisu bibliograficznego. Aby zapewnić niezawodność i wydajność, które do tej pory charakteryzowały katalogi biblioteczne, potrzebna będzie inwestycja w mechanizmy pośredniczące między czytelnikami a danymi bibliograficznymi.

Bibliografia:

1. ALDI, G. RDA implementation at Casalini. *JLIS.it.: Italian Journal of Library* [online]. 2018, Vol. 9, No. 1, s. 24-26. [Dostęp 4.03.2018] Dostępny w doi: [10.4403/jlis.it-12434](https://doi.org/10.4403/jlis.it-12434).
2. *Backstage Library Works* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.bslw.com>.
3. *BDS* [online]. [Dostęp 4.03.2018]. Dostępny w: <https://www.bdslive.com/?pid=index>.
4. BUCKLAND, M.K. Library technology in the next twenty years. *Library Hi Tech* [online]. 2017, Vol. 35, Iss. 1, s. 5–10. [Dostęp 4.03.2018]. ISSN 0099-1333. Dostępny w doi: [10.1108/LHT-11-2016-0131](https://doi.org/10.1108/LHT-11-2016-0131).
5. *Casalini Libri* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.casalini.it>.
6. Cataloging Distribution Service. W: *Library of Congress* [online]. [Dostęp 4.03.2018]. Dostępny w: <https://www.loc.gov/cds/>.
7. CICHÓN, M. Elementy RDA stosowane w opisie bibliograficznym w ramach zmian wprowadzanych przez Bibliotekę Narodową. W: *Grupa Robocza ds. Standardów Bibliograficznych* [online]. Centrum NUKAT, 14.10.2016 r. [Dostęp 4.03.2018]. Dostępny w: <http://centrum.nukat.edu.pl/rda/>.
8. CookBook Finder [online]. W: OCLC. [Dostęp 4.03.2018]. Dostępny w: <https://www.oclc.org/research/themes/data-science/cookbook-finder.html>.
9. *Core Competencies for Cataloguing and Metadata Professional librarians: approved by the ALCTS Board of Directors on January 23, 2017* [online]. January 23, 2017. [Dostęp 4.03.2018]. Dostępny w: http://www.lac.org.tw/files/core_competencies_cataloging_metadata_professional.pdf.
10. Crowdsourcing projects from the British Library. W: *LibCrowds* [online]. [Dostęp 4.03.2018]. Dostępny w: <https://www.libcrowds.com>.
11. *Deklaracja międzynarodowych zasad katalogowania* [online]. IFLA, 2009. [Dostęp 4.03.2018]. Dostępny w: https://www.ifla.org/files/assets/cataloguing/icp/icp_2009-pl.pdf.
12. HUDZIK, K. *Biblioteki uniwersyteckie w Niemczech w epoce cyfrowej: ciągłość i zmiana*. Warszawa: Wydaw. SBP, 2017. ISBN 978-83-65741-01-1.
13. KARA, A. Global records, local catalogues : investigating local editing practices in copy cataloguing. *New Zealand Library & Information Management Journal* [online]. 2014, Vol. 55, No. 1. [Dostęp 4.03.2018]. ISSN 1177-3316. Dostępny w: <https://lianza.org.nz/global-records-local-catalogues-investigating-local-editing-practices-copy-cataloguing>.
14. KOWALSKA, M. *Crowdsourcing internetowy: pozytywny wymiar partycypacji społecznej*. Warszawa: Wydaw. SBP, 2015. ISBN 978-83-64203-60-2.

15. KRYNICKA, M. *Międzynarodowe zasady katalogowania: nowe zasady katalogowania oraz zmiany w terminologii, czyli nowoczesne podejście do katalogów* (preprint) [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.bn.org.pl/download/document/1329232478.pdf>.
16. LICHTENBERGOVA, E. *The Czech National Bibliography: new steps to completeness* [online]. Submitted on 17.06.2015. [Dostęp 4.03.2018]. Dostępny w: <http://library.ifla.org/1128/1/089-lichtenbergova-en.pdf>.
17. OCLC signs agreements with leading publishers worldwide to make more content discoverable through libraries. W: OCLC [online]. Dublin, Ohio, 30 January 2018. [Dostęp 4.03.2018]. Dostępny w: <https://www.oclc.org/en/news/releases/2018/201803dublin.html>.
18. *On the record: report of the library of Congress Working Group on the future of bibliographic control* [online]. 2008. [Dostęp 4.03.2018]. Dostępny w: <https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>.
19. Program for Cooperative Cataloguing. W: *Library of Congress* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.loc.gov/aba/pcc/>.
20. RIDER, M.M. Developing new roles for paraprofessionals in cataloging. *The Journal of Academic Librarianship* [online]. 1996, Vol. 22, Iss. 1, s. 26–32. [Dostęp 4.03.2018]. ISSN 0099-1333. Dostępny w doi: [10.1016/S0099-1333\(96\)90031-2](https://doi.org/10.1016/S0099-1333(96)90031-2).
21. RUSCHOFF, C. i in. E-data quality: how publishers and libraries are working together to improve data quality. *Collaborative Librarianship* [online]. 2016, Vol. 8, Iss. 4, Art. 7. [Dostęp 4.03.2018]. ISSN 1943-7528. Dostępny w: <https://digitalcommons.du.edu/collaborativelibrarianship/vol8/iss4/7>.
22. SIWECKA, D. *Światowy model informacji bibliograficznej : program i projekty (1952–2010)*. Wrocław [etc.]: Wydaw. SBP, 2015. ISBN 978-83-64203-48-0.
23. SKÓRKA, S. Fasety na nowo odkryte: integrowanie systemów nawigacji i integrowania informacji. *Zagadnienia Informacji Naukowej* 2014, nr 2, s. 92–109. ISSN 0324-8194.
24. *Success Strategies for Electronic Content Discovery and Access : a Cross Industry White Papers* [online]. Dublin, Ohio: OCLC Online Computer Library Center, 2014. [Dostęp 4.03.2018]. Dostępny w <https://www.oclc.org/content/dam/oclc/reports/data-quality/215233-SuccessStrategies.pdf>.
25. The Legal Deposit Shared Cataloguing Programme. W: *British Library Collection Metadata: The British National Bibliography* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.bl.uk/bibliographic/clscp.html>.
26. Z39.50 Resources. W: *Z39.50 International Standard Maintenance Agency* [online]. [Dostęp 4.03.2018]. Dostępny w: <http://www.loc.gov/z3950/agency/resources/>.
27. *Zeitschriftendatenbank* [online]. Staatsbibliothek zu Berlin – Preussischer Kulturbesitz, 27.09.2017. [Dostęp 4.03.2018]. Dostępny w: <http://www.zeitschriftendatenbank.de/startseite/>.

DWORNIK, A. Dystrybucja metadanych bibliograficznych na świecie – źródła, sposoby pozyskiwania, problemy (wybrane aspekty praktyczne). *Biuletyn EBIB* [online] 2018, nr 2 (179), Dane bibliograficzne – jakość, trendy. [Dostęp 25.04.2018]. ISSN 1507-7187. Dostępny w: <http://open.ebib.pl/ojs/index.php/ebib/article/view/628>.