

**ANNA BOBER, ŻANETA KUBIC,
ALEKSANDRA PODSTAWSKA-SOBIŁO**

Opracowanie przedmiotowe zbiorów bibliotecznych na świecie – przeгляд tendencji

W ostatniej dekadzie można odnotować wzmożone zainteresowanie bibliotek wykorzystaniem języków postkoordynowanych w opisie gromadzonych dokumentów. Przyczyny tego zainteresowania są wielorakie, a wśród najważniejszych wymienić należy trudności związane ze stosowaniem języków prekoordynowanych oraz zmiany potrzeb informacyjnych użytkowników bibliotek¹. Źródeł tych zmian upatrywać należy w gwałtownym rozwoju technologicznym.

Zespół Oddziału Opracowania Zbiorów Biblioteki Jagiellońskiej (OOZ), w odpowiedzi na zauważone tendencje, od kilku lat intensywnie zajmuje się zagadnieniem przyszłości prekoordynowanego języka haseł przedmiotowych KABA, stosowanego w bibliotekach naukowych w Polsce.

Zidentyfikowane przez OOZ problemy, dotyczące wykorzystywania tego języka, wynikają przede wszystkim z jego skomplikowanej struktury. Indeksowanie za jego pomocą wymaga wysoko wyspecjalizowanej kadry bibliotekarzy, tworzenie

¹ M. Cichoń, *Library data visibility and re-use: possibilities emerging from the National Library Descriptors Project*, „Polish Libraries” 2016, vol. 4, s. 6–39.

dużej liczby rozwiniętych haseł przedmiotowych nie jest automatyczne, zaś permutacja, czyli łączenie określników z tematami w różnych kombinacjach jest trudna, czasochłonna i w rezultacie wywołuje nadmiar informacji².

Problematyczne jest również kontrolowanie autorytatywności przez kartotekę wzorcową. Jeżeli katalog nie ma kartoteki wzorcowej mechanizm ten działa częściowo lub wcale. Poprawki i modyfikacje są trudne do wprowadzenia gdy trzeba ingerować w złożoną strukturę hasła i zmieniać jego poszczególne elementy. Ponadto jedno hasło przedmiotowe powinno odpowiadać jednemu pojęciu. Gdy w dokumencie występuje wiele tematów lub gdy chce się wyeksponować temat poboczny, czy kilka aspektów tematu, pełną charakterystykę treściową dokumentu wyraża się w formie kilku haseł przedmiotowych. Jeśli w indeksie haseł przedmiotowych brakuje odpowiedniego hasła rozwiniętego, bibliotekarz musi je utworzyć korzystając z listy określników swobodnych stosowanych w odpowiednich kategoriach tematów oraz z listy określników chronologicznych i formalnych. Kiedy brakuje odpowiedniego tematu należy utworzyć nowe hasło wzorcowe i za pomocą relacji szerszych, węższych i skojarzeniowych włączyć je do sieci już istniejących terminów. Tworzenie zarówno hasła przedmiotowego wzorcowego, jak i hasła przedmiotowego rozwiniętego wymaga od katalogującego dobrej znajomości reguł gramatycznych JHP i umiejętności ich poprawnego stosowania. Od tego zależy jakość opracowania przedmiotowego. Ale nie tylko od tego.

Proces opracowania przedmiotowego można podzielić na kilka etapów: analiza tekstu dokumentu, selekcja zebranych informacji, tworzenie charakterystyki słownej, jej przekład na JHP. Może się okazać, że na każdym z nich pojawiają się zagrożenia dotyczące jakości procesu, poczynając od fałszywej interpretacji treści, poprzez błędną selekcję informacji prowadzącą do jej nadmiaru lub niedostatku, skończywszy na złym przekładzie na JHP. Jakość opracowania przedmiotowego zależy w dużym stopniu od jakości przeprowadzonej analizy i syntezy treści dokumentu, od znajomości języka haseł przedmiotowych, od wiedzy ogólnej i specjalistycznej katalogującego, czyli jego predyspozycji intelektualnych. Gdy katalogujący nie posiada odpowiedniej wiedzy i kwalifikacji, gdy brakuje podręczników omawiających praktykę tworzenia i stosowania JHP,

2 Pionierem permutacji jest niewątpliwie Molier: „[...] «Piękna markizo, twoje piękne oczy sprawiły, iż umieram dla ciebie z miłości». Albo: «Iż z miłości umieram dla ciebie, twoje piękne oczy sprawiły, piękna markizo». Albo: «Sprawiły piękne oczy twoje, iż z miłości, piękna markizo, umieram dla ciebie». Albo: «Oczy twoje, iż umieram, piękna markizo, dla ciebie z miłości, sprawiły». Albo: «Dla ciebie, piękna markizo, iż umieram, sprawiły twoje piękne oczy, z miłości»”, Molière, *Mieszczanin szlachcicem, komedia w pięciu aktach z baletem*, tłum. T. Boy-Żeleński, akt II, scena VI.

gdy brakuje możliwości do podnoszenia własnych kwalifikacji, pojawiają się problemy.

W przeszłości hasła przedmiotowe tworzono na użytek kartkowych katalogów rzeczowych czy przedmiotowych. Przeglądanie długich ciągów kart katalogowych nastręczało niemało trudności i było zajęciem bardzo czasochłonnym. Automatyzacja katalogów spowodowała, że komputerowy indeks haseł przedmiotowych służył użytkownikom jako baza terminologiczna do wyszukiwania przez słowa kluczowe. Obecnie katalogi komputerowe bibliotek nie wytrzymują konkurencji ze stronami internetowymi zapewniającymi szybki i bezpośredni dostęp do informacji. Tradycyjne katalogi OPAC (Online Public Access Catalog) oraz OPAC drugiej generacji są oceniane negatywnie jako narzędzia wyszukiwania, nie spełniając oczekiwań użytkowników. Główną przyczyną tej krytyki jest coraz szerszy dostęp do informacji, skutkujący koniecznością integracji wielorakich jej źródeł (Big Data, Linked Data).

Wśród najważniejszych wad katalogów OPAC wymieniane są: nieprzyjazne interfejsy, stosowanie indeksów, wymóg znajomości podstaw obsługi katalogu oraz terminologii bibliotekarskiej (np. „Tytuł serii” lub „Hasło przedmiotowe”), a także konieczność opanowania umiejętności wyszukiwania za pomocą operatorów logicznych opartych na algebrze Boole’a³.

Zmieniającym się potrzebom informacyjnym użytkowników powinny sprostać narzędzia wyszukiwawcze nowej generacji, które będą konkurować z prostotą wyszukiwarek internetowych Google’a⁴. W trzeciej generacji katalogów (3G) rezygnuje się zatem z wyszukiwania w oddzielnych indeksach na rzecz jednego indeksu scentralizowanego, który zapewnia dostęp do całości zbiorów bibliotecznych. Tego typu rozwiązanie pozwala na rozpoczęcie wyszukiwania na takim poziomie wiedzy, jaki się aktualnie posiada, aby później, korzystając z filtrów i dostępnych faset⁵, móc zawęzić lub poszerzać pole poszukiwań. Przyjmuje się założenie, że system sam podpowiada, w formie faset, możliwe

3 S. Skórka, *Next generation electronic catalogs [!] in Polish libraries. Usability and information architecture*, „Annales Universitatis Paedagogicae Cracoviensis. Studia ad Bibliothecarum Scientiam Pertinentia” 2016, t. 14, s. 3–17 – <http://sbsp.up.krakow.pl/article/view/20811861.14.1/3090> [dostęp: 01.12.2017].

4 A. Radwański, *System biblioteczny jako powszechna usługa sieciowa i baza danych*, w: *Systemy biblioteczne nowej generacji – platformy usług*, pod red. M. Odlanickiej-Poczobut, Gliwice 2015, s. 33 – <http://delibra.bg.polsl.pl/publication/40563> [01.12.2017].

5 Por. Faseta – podkategoria, klasy lub podklasy wynikające z zastosowania jednego kryterium podziału, Biblioteka Narodowa. Sternik: Słownik Terminologiczny z Zakresu Bibliografii i Katalogowania – <http://sternik.bn.org.pl/vocab/index.php> [19.12.2007]; *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*, oprac. B. Bojar [et al.], Warszawa 2002 – <http://bbc.uw.edu.pl/publication/639> [19.12.2007]. Zob. także PN-ISO 5127:2005. Informacja i dokumentacja. Terminologia.

sposoby dalszego wyszukiwania. Użytkownik, opierając się na rezultatach pierwszego zapytania, rozwija własną strategię wyszukiwawczą. Efektywność nawigacji zależy od trafności doboru faset oferowanych przez bibliotekę.

Wraz z pojawieniem się katalogów OPAC trzeciej generacji, opartych na wyszukiwarkach internetowych, znika także konieczność stosowania terminów w odpowiednim brzmieniu dla tematu czy określnika. Ich mechanizm wyszukiwania pozwala użytkownikowi na wpisanie dowolnych słów bez konieczności rozróżnienia pomiędzy tematem, a określnikiem rzeczowym, geograficznym czy formalnym. W tym kontekście fundamentalna zasada prekoordynacji hasła przedmiotowego traci na znaczeniu. Można przyjąć, iż prekoordynacja poszczególnych elementów hasła przedmiotowego lepiej oddaje treść dokumentu, pojawia się jednak pytanie, czy dla przeciętnego użytkownika katalogu komputerowego prekoordynacja jest na tyle zrozumiała, że będzie z korzyścią wykorzystywana do poszukiwań?

Doświadczenia w pracy z czytelnikami tego nie potwierdzają. Nakłady poniesione na kształcenie specjalistów opracowania przedmiotowego nie zwracają się w postaci wzrostu efektywności i satysfakcji użytkowników. Badania przeprowadzone w 2007 roku przez Bibliotekę Kongresu wskazują, iż ponad 3/4 użytkowników ma niewielką wiedzę nie tylko o tym jak korzystać z katalogów bibliotek, ale także niskie kompetencje informacyjne w zakresie przedmiotu swoich zainteresowań. Zaledwie 1% wszystkich użytkowników taką wiedzę i kompetencje posiada⁶. Dysproporcja pomiędzy potrzebami użytkownika, a jego kompetencjami stwarza poważny problem w odniesieniu do kwestii przygotowania opisów bibliograficznych.

Czytelnicy mają nowe wymagania odnoszące się do sposobu przeszukiwania oraz prezentacji metadanych. Rozwój technologii sieciowych (Web technologies) oraz obecność i użyteczność wyszukiwarek internetowych sprawiły, że użytkownicy opierają się na wielu różnorodnych źródłach (wyszukiwarki, księgarnie internetowe, specjalistyczne bazy danych). Serwisy komercyjne przyzwyczajają ich do stosowania jak najbardziej uproszczonych strategii wyszukiwawczych w myśl zasady: „szukaj mniej, znajdź więcej”⁷.

6 *On the Record. Report of the Library of Congress Working Group on the Future of Bibliographic Control, January 9, 2008, s. 30* – <https://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf> [01.12.2017].

7 Chan Ping Wah, J. Davasagayam, Ngian Lek Choh, *Search less find more. Redefining search environments to accommodate new user behaviors*, „Singapore Journal of Library & Information Management” 2009, vol. 38, s. 1–12 – <http://www.las.org.sg/wp/sjlim/files/SJLIM20091NgianSearch.pdf> [01.12.2017].

Zasadnicza zmiana w strategiach wyszukiwawczych stosowanych przez użytkowników nastąpiła dwie dekady temu. Zaobserwowano, że omijają oni w swoich poszukiwaniach katalogi bibliotek, a rozpoczynają poszukiwania od wyszukiwarek i innych źródeł internetowych⁸. Badania przeprowadzone w roku 2005 przez OCLC wykazały, że ponad 3/4 użytkowników katalogów przyjmuje taką strategię wyszukiwawczą (bibliotekarze również)⁹.

Z raportu Working Group on the Future of Bibliographic Control przy Library of Congress wynika, że nie można już mówić o „typowym użytkowniku” biblioteki:

Users of library materials are diverse, and a single individual will exhibit different needs, expectations, and behaviours as the purpose of his / her research changes. There is no „typical user”¹⁰.

Biblioteka powinna zatem służyć użytkownikom o różnych poziomach zaawansowania, wiedzy i kompetencji informacyjnych. Wyzwaniem staje się wobec tego wytwarzanie takich metadanych, które będą użyteczne dla jak najszerszej grupy osób. Wiele bibliotek naukowych w Polsce zdecydowało się tworzyć metadane w taki sposób, aby zaspokoić potrzeby najbardziej zaawansowanych użytkowników, pomimo faktu, że stanowią oni tylko niewielki procent ich ogólnej liczby. Stosowane jest niesprawdzone założenie, że wszyscy odniosą korzyści ze skomplikowanego sposobu opracowania. Warto jednak zadać sobie pytanie, czy biblioteki powinny operować językiem informacyjno-wyszukiwawczym o wysokiej złożoności oraz uczyć czytelników, jak z niego korzystać, czy też dostosować języki informacyjno-wyszukiwawcze do ich potrzeb? Wydaje się, iż pierwsze rozwiązanie nie jest wskazane, ponieważ prowadzi do sytuacji, w której biblioteka wkłada dużo energii, czasu i środków w budowanie narzędzia, które jest wykorzystywane w znikomym zakresie.

Praca na poziomie składniowym języka mobilizuje przeważającą część czasu i energii katalogujących, zbyt często kosztem pracy nad słownictwem. Słownictwo to, ponieważ jest kontrolowane (poziom terminologiczny) i zhierarchizowane (poziom semantyczny), stanowi prawdziwe bogactwo języka. Wydaje się, że można już dzisiaj zrewidować i odwrócić dotychczasowe priorytety, przyznać wyższość słownictwu, pozostawiając składnię na drugim planie

8 *On the Record. Report of the Library of Congress...*, s. 31.

9 *Perceptions of libraries and information resources. A Report to the OCLC Membership* – https://www.oclc.org/content/dam/oclc/reports/pdfs/Percept_all.pdf [01.12.2017].

10 *On the Record. Report of the Library of Congress...*, s. 30.

lub całkowicie z niej zrezygnować. Leksyka, jako fundamentalny rdzeń języka, może (w zasadzie) funkcjonować samodzielnie. Słownictwo nieobjęte regułami składni jest dla użytkowników zrozumiałe i umożliwia dostęp do szukanej informacji.

Słownictwo to powinno charakteryzować się następującymi cechami:

- trafnością, czyli zgodnością z zasadami języka polskiego, potwierdzoną przez podstawowe źródła języka polskiego, a także encyklopedie i słowniki specjalistyczne,
- jednoznacznością, polegającą na wykluczeniu dwuznaczności semantycznej (synonimii i homonimii),
- bogactwem w ekwiwalenty wyszukiwawcze, czyli wyrażenia dające wielokrotny dostęp do poszukiwanej informacji,
- spójnością, która wyklucza wszelkie powtórzenia i dublety semantyczne.

Od wielu lat w katalogach bibliotecznych na świecie testuje się języki oparte na postkoordynacji. W językach tych znormalizowana i ujednolicona jest jedynie forma oraz znaczenie jednostek leksykalnych, sposób ich łączenia jest zaś dowolny i zależny od inwencji użytkownika. Postkoordynacja oznacza również rezygnację z określników, każdy termin stosowany do oznaczenia treści dokumentu ma tę samą wagę, niezależnie od tego, czy oznacza przedmiot dokumentu, czy jego perspektywę chronologiczną lub geograficzną. Testowanie tego typu rozwiązań nie skutkuje na razie zdecydowanym odejściem od języków prekoordynowanych.

W bibliotekach krajów anglojęzycznych (Stany Zjednoczone, Kanada, Wielka Brytania, Australia) stosowany jest JHP Biblioteki Kongresu – Library of Congress Subject Headings – (LCSH) oraz Klasyfikacja Dziesiętna Deweya – Dewey Decimal Classification (DDC). LCSH został przyjęty także m.in. w Belgii, Czechach, Estonii, na Litwie, w Macedonii, Słowenii, Słowacji i w Szwecji, stał się podstawą do stworzenia wielu narodowych JHP (Kanada, Brazylia, Iran, Włochy)¹¹.

Klasyfikację Dziesiętną Deweya stosuje Deutsche Nationalbibliothek (DNB), na DDC opiera się klasyfikacja przedmiotowa Deutsche Nationalbibliographie. Kraje niemieckojęzyczne (Niemcy, Austria, Szwajcaria) stosują własny JHP – Regeln

11 Włoski JHP Soggettario został w 2007 roku zastąpiony przez tezaurs fasetowy Nuovosoggettario, por. D. Siwecka, *Światowy model informacji bibliograficznej: programy i projekty (1950–2010)*, Wrocław 2015, s. 185–192.

für den Schlagwortkatalog (RSWK)¹², ale z zapowiedzi wynika, że w związku z wprowadzeniem Resource Description and Access (RDA) słownictwo RSWK musi zostać przejrane, a same zasady muszą ulec zmianie. DNB współtworzy WebDewey Deutsch¹³ oraz WebDewey Search¹⁴.

W Portugalii stosowany jest JHP Siporbase (Sistema de Indexação em português), kraje związane z francuskim kręgiem kulturowym (Belgia, Szwajcaria i Kanada oraz kraje arabskojęzyczne) stosują język haseł przedmiotowych Bibliothèque nationale de France (BnF) – Répertoire d'autorité-matière encyclopédique et alphabétique unifié – RAMEAU¹⁵.

Instytucje odpowiedzialne za rozwój dwóch najpopularniejszych na świecie języków haseł przedmiotowych stosowanych w opracowaniu przedmiotowym – LCSH i RAMEAU – uznają iż języki te muszą ulec, jak ma to miejsce w przypadku RSWK i Soggettario, znacznej transformacji.

Library of Congress Subject Headings – LCSH

Już w 1998 roku w Online Computer Library Center (OCLC) we współpracy z Library of Congress rozpoczęto prace nad uproszczeniem LCSH. Skutkowało one utworzeniem opartego na LCSH schematu FAST (Faceted Application of Subject Terminology), który do dziś jest rozwijany (obecnie około 1 700 000 haseł)¹⁶. Jesienią 2016 roku w British Library, we współpracy z OCLC, rozpoczęto testy nad wdrożeniem FAST¹⁷.

Niezależnie od tych prac, powołana w 2006 roku Working Group on the Future of Bibliographic Control (Grupa Robocza ds. Przyszłości Kontroli Bibliograficznej), w opublikowanym na początku 2008 roku raporcie stwierdziła, iż LCSH:

12 Deutsche Nationalbibliothek. Regeln für den Schlagwortkatalog. RSWK, 3., überarbeitete und erweiterte Auflage auf dem Stand der 5. Ergänzungslieferung April 2009, Stand: April 2009 – http://files.dnb.de/pdf/rswk_gesamtausgabe.pdf [22.01.2018].

13 Deutsche Nationalbibliothek. WebDewey Deutsch – http://www.dnb.de/Subsites/ddcdeutsch/DE/DDCprodukte/WebDeweyDeutsch/webdeweydeutsch_node.html [01.12.2017].

14 Deutsche Nationalbibliothek. WebDewey Search – http://www.dnb.de/Subsites/ddcdeutsch/DE/DDCprodukte/WebDeweySearch/webdeweysearch_node.html [01.12.2017].

15 D. Siwecka, *Światowy model informacji bibliograficznej...*, s. 185–192.

16 FAST (Faceted Application of Subject Terminology) – <http://www.oclc.org/research/themes/data-science/fast.html> [01.12.2017]. Wykorzystuje go m.in. biblioteka narodowa Nowej Zelandii.

17 *Response to the consultation on Subject Standards* (11 July 2016) – <https://www.bl.uk/bibliographic/pdfs/british-library-response-survey-subject-standards.pdf> [01.12.2017].

- ma skomplikowaną strukturę, kłopotliwą zarówno ze względu na administrowanie danymi, jak i automatyzację procesu ich przetwarzania,
- stosowane w hasłach słownictwo jest przestarzałe, nieintuicyjne z punktu widzenia użytkownika,
- utworzone najwcześniej terminy nie są powiązane relacjami z innymi,
- tworzenie języka prekoordynowanego jest procesem czasochłonnym i skomplikowanym, trudnym do nauczenia,
- pomimo, iż prekoordynacja ukazuje ukryte relacje pomiędzy terminami (szersze, węższe), to tworzone przez katalogerów hasła złożone są niejednokrotnie niejasne i niezrozumiałe dla użytkowników,
- stosowane terminy są zbyt wąskie (główna zasada przy opracowaniu: najbardziej odpowiedni termin), podczas kiedy badania wykazały, że użytkownicy oczekują raczej terminów bardziej ogólnych. W hasłach stosowane są odesłania do terminów szerszych, a nie węższych,
- hasła LCSH są dostępne tylko w formacie MARC, co wyklucza ich użytkowanie poza środowiskiem bibliotecznym¹⁸.

Utrzymywanie istniejącego stanu rzeczy przyniesie między innymi dalsze odchodzenie od katalogów bibliotek na rzecz wyszukiwarek internetowych. Wyszukiwarki mogą być mniej skuteczne, ale korzystanie z nich nie wymaga ugruntowanej wiedzy w zakresie konkretnej dziedziny.

Rekomendowano m.in.:

- przekształcenie LCSH w elastyczne narzędzie do tworzenia i modyfikowania haseł przedmiotowych,
- otwarte udostępnianie LCSH,
- rezygnację z dotychczasowego układu alfabetycznego na rzecz tezaurusa tematycznego,
- zwiększenie relacji pomiędzy LCSH a Klasyfikacją Biblioteki Kongresu oraz DDC,
- zintensyfikowanie prac nad LCSH, szczególnie w odniesieniu do wykorzystywania tego języka przez narzędzia dostępne szerokiej publiczności,
- ocenę możliwości LCSH w zakresie obsługi wyszukiwania w fasetach,
- stworzenie powiązań pomiędzy LCSH a innymi językami informacyjno-wyszukiwawczymi (JIW) np. klasyfikacją MeSh (Medical Subject Headings) lub Searsa (Sears List of Subject Headings),

18 *On the Record. Report of the Library of Congress...*, s. 34–35.

- redukcję kosztów katalogowania poprzez zacieśnienie współpracy z OCLC,
- wykorzystanie potencjału automatycznego indeksowania w przypadku pełnotekstowych dokumentów elektronicznych¹⁹.

W 2009 roku w Library of Congress utworzono Linked Data Service, Authorities and Vocabularies, gdzie hasła prezentowane są przy użyciu Linked Data. Warto dodać, że w Resource Description and Access przewidywane jest stworzenie zasad opracowania przedmiotowego, trudno jednak stwierdzić, kiedy miałyby to nastąpić²⁰.

Répertoire d'autorité-matière encyclopédique et alphabétique unifié – RAMEAU

Z uwagi na fakt, że język haseł przedmiotowych BnF służył jako wzór w budowaniu kartoteki wzorcowej JHP KABA, to zmianom w nim zachodzącym należy poświęcić szczególną uwagę.

RAMEAU wywodzi się z Répertoire des vedettes matière (RVM), tworzonego od 1946 roku przez kanadyjską Bibliothèque de l'Université Laval w Québecu i jest francuskojęzyczną adaptacją języka LCSH (kartoteki wzorcowej Biblioteki Kongresu). W bibliotekach we Francji był rozwijany i stosowany od 1980 roku²¹.

Głównym powodem jego tworzenia była chęć wykorzystania go w katalogach bibliotek, ale także innych instytucji, zajmujących się dokumentowaniem dziedzictwa kulturowego²².

Ponieważ założono, że znajdzie tak szerokie zastosowanie, jego słownictwo charakteryzuje się różnym poziomem specjalizacji, zależnym od potrzeb

19 *On the Record. Report of the Library of Congress...*, s. 35–36.

20 *Resource Description and Access Steering Committee. RDA. A Prospectus*: „RDA is being developed as a new standard for resource description and access designed for the digital world” – http://rda-jsc.org/sites/all/files/5rda-prospectusrev7_0.pdf [01.12.2017]; Deutsche Nationalbibliothek. Standardisation – http://www.dnb.de/EN/Standardisierung/standardisierung_node.html [01.12.2017].

21 Zob. M. Mingam, *Rameau. Bilan, perspectives*, „Bulletin des bibliothèques de France” 2005, nr 5, s. 38–47 – <http://bbf.enssib.fr/consulter/bbf-2005-05-0038-001.pdf> [01.12.2017], por. idem, *RAMEAU, bilans i perspektywy*, w: *Opracowanie przedmiotowe dokumentów z zakresu nauk ścisłych: matematyczno-przyrodniczych i technicznych. Język haseł przedmiotowych KABA: teoria, praktyka, przyszłość*. Kazimierz Dolny, 20–22 września 2006 roku, Warszawa 2006 – <http://www.ebib.pl/publikacje/matkonf/kaba/mingam.php> [20.12.2017].

22 M. Mingam, *Rameau, les catalogues, le web*, „Bulletin des bibliothèques de France” 2015, nr 5, s. 120–131 – <http://bbf.enssib.fr/consulter/bbf-2015-05-0120-003.pdf> [01.12.2017].

użytkowników: od bardzo ogólnego do bardzo specjalistycznego. Cechuje je encyklopedyzm, kompletność i granulacja słownictwa.

RAMEAU jest encyklopedyczny, ponieważ obejmuje wszystkie dziedziny wiedzy i praktyki. Hasła nie powstają *a priori*. Utworzenie deskryptora zależy od potrzeb indeksacji dokumentu, od potrzeb katalogującego oraz instytucji dokumentującej. Nie rości sobie pretensji do kompletności, ale jest elastyczny i może być wzbogacany w nieskończoność. Granulacja słownictwa w ramach haseł jest doprowadzona do momentu, kiedy jest ono jeszcze precyzyjne.

RAMEAU został zaprojektowany jako język prekoordynowany – zespół reguł, pozwalających w sposób systematyczny tworzyć hasła wzorcowe, ale jest również bogatym zasobem kontrolowanego i ustrukturyzowanego słownictwa, które trudno zredukować tylko do składni.

Język RAMEAU, dzięki swojej strukturze [...] umożliwia [...] trzy uzupełniające się sposoby wyszukiwania przedmiotowego w katalogach. Te sposoby mogą być wykorzystywane przez biblioteki niezależnie jeden od drugiego lub łącznie, dla lepszej obsługi użytkowników: jako język kontrolowany, łącząc terminy przyjęte i odrzucone, RAMEAU pozwala na wyszukiwanie przez słowa kluczowe; jako język zhierarchizowany pozwala na nawigację semantyczną w słownictwie za pomocą relacji szerszych, węższych i skojarzeniowych pomiędzy terminami; [...] jako język prekoordynowany rządony zasadami składni, pozwala na wyszukiwanie poprzez przeglądanie indeksów²³.

W związku z nowymi perspektywami, związanymi z możliwością prezentowania danych w sieci, przystąpiono do systematycznej rewizji słownictwa, która miałaby znacząco poprawić jego jakość i spójność. Podjęto decyzję, aby główny nacisk położyć na aspekty terminologiczne (sprawy niejasności języka naturalnego) i semantyczne (zarządzanie relacjami pojęciowymi między deskryptorami), koncentrując się na hasłach wzorcowych, tworzących repozytorium.

Dzięki kontrolowanemu słownictwu RAMEAU stara się ograniczyć w jak największym stopniu niejasności języka naturalnego (homonimy, synonimy). Celem jest osiągnięcie słownictwa dokładnego co do treści, jednorodnego pod względem formy i jednoznacznego pod względem zastosowania (gdzie deskryptor = pojęcie i pojęcie = deskryptor) przy jednoczesnym zwiększeniu

23 M. Mingam, *RAMEAU, bilans i perspektywy...*

punktów dostępu (poprzez zapewnienie odpowiedniej liczby odsyłaczy). Wybór formy preferowanej (ujednoliconej) został poddany surowym regułom. Hasło główne musi być równocześnie dokładne, trafne, niedwuznaczne i zgodne z poświadczonym w źródłach stosowaniem.

Przedmiotem szczególnej uwagi pozostają synonimy. Podstawową funkcją deskryptorów w języku indeksowania jest nie tyle wyrażać zawartość dokumentu, ile go wskazać, dać do niego dostęp, aby użytkownik mógł następnie go znaleźć poprzez terminy wyszukiwawcze, formy odrzucone, równoważne do hasła głównego, w ramach zintegrowanej całości, która ma postać „kiści”, „bukietu” (*grappe / bouquet*) terminologicznego. Chodzi głównie o zwielokrotnienie relacji równoważności, próbę przewidzenia typów zapytań, które mogą być formułowane przez użytkowników. Słownictwo powinno być zarówno popularne, jak i naukowe. Różne warianty mogą stać się formami preferowanymi, w zależności od potrzeb.

Celem ewolucji terminologicznej jest więc zaproponowanie słownictwa (na ile to możliwe) jak najściślej ustalonego, jednoznaczego, oferującego wiele możliwości uzyskania dostępu do dokumentu.

Kontrolowane słownictwo RAMEAU może przedstawiać się również jako sieć połączonych w sposób uporządkowany deskryptorów, których zadaniem jest umożliwienie użytkownikowi poruszania się pomiędzy hasłami wzorcowymi, pozwalającymi zawężać, uogólniać i przekierowywać wyszukiwanie (hasła związane). Sieć relacji semantycznych powinna być budowana w taki sposób, żeby w efekcie powstał zestaw tezaursów, zgrupowanych wokół 15 deskryptorów, odpowiadających wielkim dziedzinom wiedzy i działalności (filozofia, religia, prawo, nauki społeczne, nauki humanistyczne, nauki techniczne, nauki ścisłe, technika, medycyna, sztuka, literatura...). Począwszy od tych deskryptorów rozwijać się będzie całe, powiązane relacjami, słownictwo. Otwiera to nadzwyczaj interesujące perspektywy dla użytkownika końcowego w zakresie wyszukiwania informacji zarówno w odniesieniu do katalogów, jak sieci Web.

Skupienie się na terminologii (słownictwie) i semantyce nie oznaczało całkowitego zarzucenia składni. RAMEAU nie przestał jeszcze być językiem prekoordinowanym, który ma reguły syntaktyczne. Założono, że znacząca zmiana tych reguł lub rezygnacja z nich naruszy spójność katalogów i w efekcie spowoduje powstanie niekorzystnego dla użytkownika zjawiska „szumu informacyjnego”. Prace miały się więc ograniczać do aktualizacji reguł, wzbogacania podziałów, wprowadzania do haseł wzorcowych informacji o stosowaniu. Już wtedy jednak uważano, że uproszczenie składni jest bardzo pożądane.

W 2017 roku ukazał się raport Groupe de travail national sur la syntaxe de Rameau, w którym mówi się o rewizji i granulacji słownictwa, ale także o przejściu od logiki indeksu do zwielokrotnienia relacji między encjami oraz o uproszczeniu reguł rządzących składnią, a nawet uwolnieniu RAMAU od reguł składniowych²⁴. Wnioski płynące z prac wskazują na kierunek przyszłych zmian – przekształcanie RAMEAU w język o charakterze deskryptorowym i rezygnacja z prekoordynacji.

Pozostaje kwestia w jaki sposób potencjał metadanych RAMEAU może zostać wykorzystany zarówno w katalogach, jak i w sieci, mając na uwadze potrójny wymiar tego narzędzia indeksacji. Odpowiedź na pytanie nie zależy tylko od samego RAMEAU, ale także od już istniejących i przyszłych możliwości technicznych.

Aby można było wykorzystać potencjał RAMEAU na poziomie terminologicznym, niezbędnym minimum, zdaniem jego twórców, jest zarządzanie formami odrzuconymi. Konieczne jest również zapewnienie nawigacji semantycznej w obrębie haseł wzorcowych, ponieważ znacząco wspomaga to użytkownika w poszukiwaniu informacji i zwiększa możliwości dotarcia do niej. Prekoordynacja RAMEAU pozwala także na jego prezentację w formie indeksów alfabetycznych.

W 2008 roku słownictwo RAMEAU przetłumaczono w ramach TEL (The European Library) w formacie SKOS (Simple Knowledge Organization System) zgodnie ze standardem RDF (Resource Description Framework), pozwalającym na prezentację i obsługę wszystkich typów słowników kontrolowanych²⁵. W 2011 roku powstał projekt data.bnf.fr, który prezentuje zasoby BnF w standardach sieci semantycznej, posługując się bogatą i ustrukturyzowaną terminologią języka RAMEAU²⁶.

Jakie podstawowe pytanie zadali sobie twórcy data.bnf.fr? Nie to, w jakim formacie i przy zastosowaniu jakich narzędzi technicznych tworzyć. Pierwsze pytanie brzmiało: czego i w jaki sposób szukają internauci? Jakie są ich praktyki wyszukiwawcze? Szukają konkretnych rzeczy, pojęć, miejsc geograficznych,

24 F. Menard, *Rapport du Groupe de travail national sur la syntaxe de Rameau. Préconisations et pistes d'évolution* (Version 1.0; 05/05/2017) – http://rameau.bnf.fr/chantier_syntaxe/pdf/rapport_final_syntaxe_rameau.pdf [01.12.2017].

25 Por. A. Isaac, T. Bouchet, *RAMEAU et SKOS*, „Arabesques” 2009, nr 54, s. 13–14 – <http://www.abes.fr/Publications-Evenements/Arabesques/Arabesques-n-54> [22.01.2018].

26 Bibliothèque nationale de France. 7e journée nationale RAMEAU. Jeudi 6 novembre 2014, A. Simon, *RAMEAU dans data.bnf.fr* – http://www.bnf.fr/fr/professionnels/anx_journees_pro_2014/a.jp_141106_rameau.html#SHDC_Attribute_BlocArticle5BnF [01.12.2017]; eadem, *Rameau dans le web sémantique* – http://www.bnf.fr/documents/rameau2014_simon.pdf [01.12.2017]; A. Simon, R. Wenz, *Des outils automatiques pour le signalement en bibliothèque: expérimentations autour du projet data.bnf.fr.*, „Bulletin des bibliothèques de France” 2012, nr 5, s. 39–43 – <http://bbf.enssib.fr/consulter/bbf-2012-05-0039-008> [22.01.2018].

autorów, itd. Nie szukają przez katalogi komputerowe, których logiki nie rozumieją, szukają na stronach Web klikając w interesujące ich tematy, pojęcia, nazwy. Zdecydowało to o wyborze języka RAMEAU do publikacji zasobów BnF. Okazało się to możliwe po konwersji RAMEAU na format RDF według modelu SKOS. Bogata terminologia, sieć relacji semantycznych wiążących ze sobą terminy, symbole ARK (Archival Resource Key), identyfikujące poszczególne hasła wzorcowe spowodowały, że RAMEAU stał się użytecznym narzędziem, służącym porządkowaniu danych biblioteki. Jej zasoby, skupione wokół haseł – autorów, tytułów i pojęć, tworzą strony tematyczne, po których użytkownik ma możliwość swobodnej nawigacji. Słownictwo RAMEAU zostało udostępnione za pośrednictwem data.bnf.fr, co pozwala na jego wielokrotne wykorzystanie i powoduje, że sam język staje się pełnoprawnym i liczącym się w świecie uczestnikiem tworzenia sieci semantycznej i danych internetowych. Ustrukturyzowane metadane, jakimi są hasła wzorcowe RAMEAU, mogą być użytecznymi narzędziami do korzystania z sieci i pozwalają na jej uporządkowanie.

* * *

Planowany rozwój RAMEAU, leżącego u podstaw JHP KABA, powinien stać się wyznacznikiem kierunku ewolucji języka stosowanego przez polskie biblioteki naukowe. Należy przy tym wykorzystać te zasady opracowania zbiorów, które powstały w odpowiedzi na zmieniające się wraz z rozwojem technologicznym otoczenie sieciowe. Najważniejsze z nich zawarte zostały w Statement of International Cataloguing Principles (2009). Nowe wytyczne, zastępujące obowiązujące dotychczas Zasady Paryskie (1961), przeniosły punkt ciężkości z kwestii istotnych dla bibliotekarzy na kwestie istotne dla użytkowników. Za najważniejszą zasadę katalogowania uznano wygodę użytkownika – „decyzje dotyczące tworzenia opisów i kontrolowanych form nazw jako punktów dostępu powinny być podejmowane z myślą o użytkowniku”. Pozostałe zasady to:

- **Powszechność użycia.** Słownictwo stosowane w opisach i w punktach dostępu powinno być zgodne ze słownictwem używanym przez większość użytkowników.
- **Reprezentatywność.** Opisy i kontrolowane formy nazw powinny odpowiadać sposobowi, w który jednostka się samookreśla.
- **Trafność.** Opisywana jednostka powinna być wiernie przedstawiona.

- **Wystarczalność i konieczność.** W opisach i kontrolowanych formach nazw jako punktach dostępu, powinny się znaleźć tylko te elementy danych, które odpowiadają potrzebom użytkownika i są niezbędne do jednoznacznej identyfikacji jednostki.
- **Znaczenie.** Elementy danych powinny być znaczące bibliograficznie.
- **Ekonomia.** Gdy istnieją alternatywne drogi osiągnięcia celu, należy dać pierwszeństwo sposobowi najbardziej wydajnemu (tzn. najniższy koszt lub najprostszy sposób).
- **Spójność i standaryzacja.** W miarę możliwości należy znormalizować opisy i budowę punktów dostępu. Umożliwi to większą spójność, która z kolei zwiększy możliwość wymiany danych bibliograficznych i wzorcowych.
- **Integracja.** Opisy wszystkich typów dokumentów i kontrolowane formy nazw wszystkich typów jednostek powinny być zgodne, w miarę możliwości, ze wspólnymi zasadami.

Zasady stosowane w przepisach katalogowania powinny być uzasadnione, a nie narzucone arbitralnie. W pewnych sytuacjach te zasady mogą być wzajemnie sprzeczne, wówczas należy zastosować uzasadnione, sprawdzone rozwiązanie²⁷.

Wobec powyższych wytycznych warto się zastanowić, w jakim zakresie stosowane w Polsce języki informacyjno-wyszukiwawcze je spełniają. W tym kontekście ważną kwestią staje się rewizja dotychczas stosowanych języków informacyjno-wyszukiwawczych (JIW) i sformułowanie pytań odnoszących się do przyszłości opracowania przedmiotowego zbiorów bibliotek naukowych w Polsce. Analiza światowych tendencji, najnowszych wytycznych oraz identyfikacja najważniejszych problemów związanych z wykorzystywaniem języka prekoordynowanego w codziennej praktyce stały u źródła decyzji o podjęciu prac w Bibliotece Jagiellońskiej na rzecz zastąpienia JHP KABA przez język postkoordynowany, przy zachowaniu wypracowanego przez lata bogatego słownictwa. Ponieważ efektem powyższych zadań byłoby niewątpliwie przekształcenie języka prekoordynowanego w język deskryptorowy, postanowiono skorzystać z doświadczeń Biblioteki Narodowej w tym zakresie. W 2016 roku obie instytucje, a następnie również Biblioteka Uniwersytecka w Toruniu, podjęły prace przygotowawcze do przekształcenia JHP KABA na Deskryptory BN.

²⁷ *Deklaracja międzynarodowych zasad katalogowania. Wprowadzenie (2009)* – http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-pl.pdf [01.12.2017].

Prace realizowane są przez bibliotekarzy z BJ mających doświadczenie w pracy z JHP KABA, zarówno w zakresie stosowania, jak i tworzenia tego języka.

Pierwszym etapem jest sporządzenie zestawień jednostek leksykalnych JHP KABA i DBN. Zestawienia te stanowią podstawę do zbudowania schematów przekształceń haseł rozwiniętych JHP KABA na DBN, w ramach drugiego etapu prac. Jako etap ostatni planowane jest przekształcenie, w większości automatyczne, opisów przedmiotowych JHP KABA na DBN oraz weryfikacja wyników przekształcenia. Prowadzone prace mają na celu wzbogacenie zasobu słownictwa DBN, zarówno pod względem szczegółowości słownictwa, jak i odpowiedników z języków LCSH i RAMEAU oraz utworzenie jednego języka opracowania rzeczowego dokumentów dla wszystkich polskich bibliotek.

ANNA BOBER, ŻANETA KUBIC, ALEKSANDRA PODSTAWSKA-SOBIŁO

Subject cataloguing of library collections in the world: overview of tendencies

The article discusses inevitable changes that occur in the field of subject cataloguing in library catalogues in Poland and in the world. It points to the reasons for the creation of a new language – the Descriptors of the National Library of Poland (DBNs). The DBNs are an alternative to the hitherto used language of subject description, which is outdated and impractical to use both for librarians and library users. The authors present the context that brought about the changes in traditional tools used in librarians' work. They quote examples of changes that have already been implemented or are underway in libraries all over the world. They indicate that the pre-coordinated subject description languages still in use in Poland are obsolete and unsatisfying for modern-day users of library collections catalogues.