

Marta Cichoń, Jakub Kalinowski, Grażyna Federowicz

Katalogowanie oparte na encjach

Wstęp

Artykuł ma na celu przedstawienie wprowadzanych przez Bibliotekę Narodową zmian w opracowaniu formalnym i rzeczowym – zarówno w kontekście globalnych trendów, które do decyzji o potrzebie zmian doprowadziły, jak i dalszych związanych z nimi analiz, które zmierzają do następnych etapów, będących logiczną kontynuacją zmian bieżących. Realizowany obecnie etap zakłada implementację modelu FRBR w zakresie, w jakim pozwalają możliwości formatu MARC 21 w zakresie struktury i kompletności danych bibliograficznych. Format MARC 21 nadal ma być narzędziem opracowania bibliograficznego, jednak w strukturze hybrydowej wprowadzającej do stosowanego standardu ISBD pewne elementy standardu RDA. Zgodnie z postulowanym przez IFLA Linked Data Study Group założeniem granulacja elementów RDA całkowicie pokrywa się z granulacją w standardzie ISBD (z pewnymi wyjątkami, z których znaczna część dotyczy strefy uwag)¹. Zagadnienia związane z kontekstem zmian w opisie i krótka analiza globalnych zmian w opracowaniu bibliograficznym, której są konsekwencją, jest głównym przedmiotem pierwszej części artykułu (M. Cichoń).

Większej atomizacji danych, polegającej na przeniesieniu znacznej części informacji zawartej do tej pory w polach uwag do nowych pól formatu z grupy 3XX, towarzyszy transformacja Języka Haseł Przedmiotowych Biblioteki Narodowej (JHP BN), mająca na celu uproszczenie jego słownictwa, likwidacja składni i przekształcenie go w Deskryptory Biblioteki Narodowej (DBN), których zasób miałby w przyszłości stanowić jedną bazę zarówno dla opisu formalnego, jak i przedmiotowego. Jest to przede wszystkim odpowiedź na stały spadek wykorzystania indeksu przedmiotowego w dotychczasowych katalogach online. Szczegóły zagadnień związanych z trudnościami, jakie wystąpiły w pracy przy Języku Haseł Przedmiotowych BN oraz z opracowaniem nowej bazy Deskryptorów BN są opisane w części drugiej artykułu (J. Kalinowski).

Kolejna część porusza zagadnienia ujednoczenia haseł kartoteki haseł formalnych i przedmiotowych (G. Federowicz) oraz uczynienia z zasobu Deskryptorów BN jednej bazy dla opisu formalnego i rzeczowego. Temat samej

1 G. Dunsire, *ISBD unconstrained elements and other extensions* – www.ifla.org/files/assets/cataloguing/isbdrg/isbd-extensions.pdf [dostęp: 08.02.2015].

integracji wychodzi już poza zakres zmian bieżących i nie jest w tej pracy szczegółowo poruszany. Analiza dotycząca ujednoczenia haseł jest istotnym elementem prac nad Deskryptorami BN w związku z rolą jednej bazy, jaką miałyby w przyszłości pełnić, dlatego uznaliśmy za zasadne poruszenie problematyki z nią związanej, opisaną z wykorzystaniem przykładów w części trzeciej.

1. Kontekst globalnych zmian w opracowaniu bibliograficznym

Bazy danych bibliograficznych a relacyjne bazy danych

Pierwsza generacja katalogów dostępnych online była zasadniczo produktem ubocznym procesu automatyzacji pracy w bibliotekach. Wczesne katalogi tego typu służyły przede wszystkim usprawnieniu procesów udostępniania zbiorów i wymiany metadanych między bibliotekami i nie były projektowane z myślą o potrzebach wyszukiwawczych użytkowników. Jest bardzo niewiele badań na temat potrzeb czytelników i sposobu wyszukiwania przez nich informacji z okresu sprzed rozwoju pierwszych katalogów dostępnych online (OPAC)². Już pierwsze spośród takich badań prowadzone w pierwszej połowie lat siedemdziesiątych XX wieku pokazywały, że wyszukiwanie przedmiotowe stanowi około 20-40% kwerend prowadzonych w katalogu³. Praktyki dotyczące sposobu wyszukiwania informacji były natomiast wielokrotnie analizowane w okresie późniejszym, kiedy katalogi OPAC zaczęły wchodzić do użycia, począwszy od drugiej połowy lat siedemdziesiątych. Również te badania wskazywały na szczególne problemy, jakie dotyczyły wyszukiwania przedmiotowego⁴. Wobec braku kompleksowego – formalnego i systematycznego – projektu zarówno w początkowym, jak i późniejszym okresie rozwoju katalogów elektronicznych zachodziło zjawisko – dość powszechne w rozwoju usług, jaką jest przecięż katalog w rozumieniu narzędzia dostępnego online – odpowiadania na nowe możliwości technologiczne ustalaniem celów krótkoterminowych kosztem długoterminowych. Jak wiadomo, cyfryzacja środków dostarczania i organizowania usług zmieniła oblicze całego sektora usługowego (prywatnego i publicznego),

2 A. Christensen, *Next generation catalogues: what users think?*, w: S. Chambers (ed.), *Catalogue 2.0: The Future of Library Catalogue*, Chicago 2013, s. 2.

3 R. R. Larson, *The Decline of Subject Searching: Long-Term Trends and Patterns of Index Use in an Online Catalog*, „Journal of the American Society for Information Science” 1991, t. 42, nr 3, s. 198 – www.researchgate.net/profile/Ray_Larson/publication/220435755_The_decline_of_subject_searching_Long-term_trends_and_patterns_of_index_use_in_an_online_catalog/links/02e7e5227424fb54e2000000.pdf [08.02.2015].

4 M. J. Bates, *Factors affecting subject catalog search success*, „Journal of the American Society for Information Science” 1977, t. 28, nr 3, s. 161.

stwarzając ogromny potencjał rozwoju. Wynikające z tego skoku technologicznego możliwości wymagają dogłębnej analizy semantyki usług i ich struktury dla opracowania projektu, który mógłby w pełni je wykorzystywać, są więc one zazwyczaj niedostatecznie rozumiane, w wyniku czego realizowane są przede wszystkim założenia krótkoterminowe⁵.

W rezultacie takiego podejścia elektroniczne katalogi biblioteczne zostały na początku swojego istnienia zaprojektowane jako elektroniczna wersja katalogów kartkowych, dająca możliwość przeszukiwania alfabetycznego określonej liczby prekoordynowanych indeksów (autorski, tytułowy, przedmiotowy) tak jak w katalogu kartkowym. W kolejnych generacjach elektronicznych katalogów bibliotecznych dodawano funkcjonalności takie jak przeszukiwanie indeksów postkoordynowanych (np. słowa kluczowe), zastosowanie języka zapytań z użyciem operatorów logicznych (operatorów Boole'a) w wyszukiwaniu zaawansowanym przez różne indeksy oraz grupowanie zapytań, a w końcu interfejs użytkownika umożliwiający korzystanie z katalogu w sieci internetowej. Katalogi biblioteczne w swojej istocie elektroniczne nie zmieniły się zasadniczo od początku, kiedy były narzędziem służącym do przeniesienia danych z katalogów kartkowych⁶.

Sam interfejs umożliwiający zdalne przeszukiwanie katalogu OPAC nie wpływa bezpośrednio na możliwość wyszukiwania zasobów katalogu w Internecie za pomocą zwykłych wyszukiwarek. Zawartość bazy danych katalogu nie jest widoczna dla eksponentów sieciowych i jest dostępna tylko dla użytkowników interfejsu katalogu. Tradycyjny katalog jest częścią „głębskiego Internetu” (*deep Web*), którego treści są ukryte na dynamicznie generowanych stronach internetowych⁷. Z tego samego powodu przy przenoszeniu danych z jednego katalogu do innego konieczna jest ich duplikacja – nie jest możliwe płynne odwołanie się do treści niezindeksowanych stron⁸.

Opisany powyżej rozdzźwięk pomiędzy możliwościami technologicznymi a obecnym ich wykorzystaniem w katalogach bibliotecznych nie dotyczy jedynie strukturalnych zasobów biblioteki. Dużo trudniejszym zagadnieniem są bariery formalne wynikające z samych zasad opracowania bibliograficznego.

5 L. V. Orman, *Service Semantics, Structure and Design*, „e-Service Journal” 2008, t. 6, nr 2, s. 59 – www.ssrn.com/abstract=1019041 [08.02.2015].

6 T. Kinstler, *Making search work for the library user*, w: S. Chambers, *Catalogue 2.0...*, s. 18.

7 K. R. Diaz, *The Invisible Web: Navigating the Web outside Traditional Search Engines*, „Reference and User Services Quarterly” 2000, t. 40, nr 2, s. 132 – <http://hdl.handle.net/1811/44703> [08.02.2015].

8 E. Bermes, *Enabling your catalogue for the Semantic Web*, w: S. Chambers, *Catalogue 2.0...*, s. 118.

Zasady katalogowania wyrażone w standardach takich jak AACR2 (Anglo-American Cataloguing Rules) czy nieco późniejszy ISBD (International Standard Bibliographic Description) opierają się na założeniach przyjętych podczas konstruowania katalogu kartkowego i w swojej zasadniczej części nie zmieniły się od czasów sprzed komputeryzacji⁹.

Zgodnie z podstawowymi założeniami projektowania baz danych oraz tworzenia systemów do ich zarządzania, projekt bazy danych jest zwykle podzielony na fazy specyfikacji i projektu konceptualnego (określanego często jako widok zewnętrzny, ponieważ dotyczy konceptu i struktury danych niezależnych od zastosowanej technologii) oraz projektu logicznego i fizycznego (określanego z kolei często jako widok wewnętrzny, ponieważ dostępna i zastosowana technologia ma zasadniczy wpływ na fizyczną strukturę bazy). Fazy te są w różnych modelach traktowane mniej lub bardziej rozdzielnie, ale dla optymalizacji projektu kolejność ma tu zasadnicze znaczenie. Po dokonaniu specyfikacji w pierwszej kolejności powinien zostać zdefiniowany konceptualny – zewnętrzny widok bazy. Celem jest określenie formalnej reprezentacji informacji zawartych w bazie oraz procesów między nimi, niezależnej od zastosowanego w dalszych fazach systemu zarządzania bazą danych (Database Management System – DBMS). W fazie projektu logicznego ten konceptualny schemat zostaje przełożony na logiczny model danych opracowywany na potrzeby wybranego systemu DBMS, taki jak na przykład model związków encji (model ER – *entity-relationship model*)¹⁰. Następnie zostaje on zaimplementowany w fazie projektu fizycznego bazy na potrzeby określonej technologii, w którym definiowany jest wewnętrzny schemat rozmieszczenia danych w bazie¹¹.

Jeżeli spojrzeć na schemat zapisu danych w bazie bibliograficznej, należy zauważyć, że większość z nich określa strukturę danych na poziomie konceptualnym z dwoma wyjątkami – zasad dotyczących formatu zapisu danych oraz tworzenia punktów dostępu (te zasady dotyczą poziomu wewnętrznego). Z przyczyn opisanych wcześniej rozróżnienie poziomu zewnętrznego i wewnętrznego w przypadku standardów katalogowania jest zadaniem wtórnym, ponieważ w przypadku katalogu kartkowego, na którego potrzeby w pierwszej kolejności standardy te były tworzone, nie istniało rozróżnienie na te dwa poziomy¹².

9 S. Batley, A. Welsh, *Practical Cataloging: AACR, RDA and MARC2*, London 2012, s. 17.

10 P. Chen, *Entity-Relationship Modeling. Historical Events, Future Trends, and Lessons Learned* – http://bit.csc.lsu.edu/~chen/pdf/Chen_Pioneers.pdf [08.02.2015].

11 R.C. Goldstein, V. C. Storey, *Knowledge-Based Approaches to Database Design*, „MIS Quarterly” 1993, t. 17, nr 1, s. 25 – www.jstor.org/stable/249508 [08.02.2015].

12 M. Crandall, R. Fidel, *The AACR2 as a Design Schema for Bibliographic Databases*,

Zgodnie z jedną z elementarnych zasad tworzenia bazy danych, dziesiątym postulatem Codda, więzy spójności w bazie są niezależne – definiowane w bazie i niezależne od aplikacji. Ogólnie cztery z dwunastu postulatów Codda odnoszą się do niezależności danych i sprowadzają się do ścisłego rozdziału decyzji na poziomie konceptualnym, zewnętrznym i wewnętrznym. W sformułowaniach zapisu standardu ISBD wyraźne jest podejście, w którym nad odrębne i wyraźne zdefiniowanie poszczególnych elementów samej bazy danych przedkładany jest pragmatyzm w wyrażaniu elementów kolejnych stref charakterystycznych dla opisu bibliograficznego¹³. Zarówno struktura AACR2, jak i ISBD nie jest więc zgodna ze wspomnianymi postulatami niezależności danych. W AACR2, będącym poprzednikiem RDA, zasady odnoszące się do formatu zapisu danych nie były wyszczególnione, zostały ujęte w tych samych rozdziałach, co zasady odnoszące się do innych aspektów katalogu. Ponadto zasady katalogowania, które odnoszą się do formatu zapisu, nie są rozdzielone i najczęściej są przemieszane z innymi zasadami, na przykład tworzenia samych encji, relacji i atrybutów w bazie. Raya Fidel i Michael Crandall pokazują to zjawisko na przykładzie zasady 1.8C1 dotyczącej nadawania tytułu kluczowego, która brzmi: „dodaj tytuł kluczowy wydawnictwa ciągłego, jeżeli znajduje się na dokumencie lub jest łatwo dostępny w inny sposób po numerze ISSN. Podaj tytuł kluczowy nawet jeżeli jest identyczny z tytułem właściwym. Jeżeli brak numeru ISSN, nie podawaj tytułu kluczowego”¹⁴. Z jednej strony jest to zasada dotycząca zawartości, informująca, że tytuł kluczowy może być zawarty w rekordzie po numerze ISSN, z drugiej strony jest zasadą nadawania atrybutu z określeniem jego źródła. Zasada najpierw odnosi się do formatu zapisu danych – określa źródło i wskazuje, że tytuł nadawany jest po numerze ISSN, a dopiero na końcu określa reguły nadawania atrybutu tytułu kluczowego („dodaj tytuł kluczowy, nawet gdy jest identyczny z tytułem właściwym; nie dodawaj, jeżeli nie ma numeru ISSN”). W projektowaniu systemu zarządzania bazą danych zalecane byłoby odwrócenie tej kolejności – najpierw należałoby określić, czy i kiedy atrybut jest nadawany, potem zdefiniować regułę i źródła dla jego ustalenia („kiedy jest dostępny, nawet jeżeli jest identyczny z tytułem właściwym, ale tylko gdy podany jest numer ISSN”). Zasada dotycząca formatu zapisu (tytuł kluczowy powinien być zapisany po numerze ISSN) byłaby przy zastosowaniu takiego podejścia ujęta w oddzielnej sekcji. Przykład ten został zacytowany dla

„The Library Quarterly” 1988, t. 58, nr 2, s. 134-135 – www.faculty.washington.edu/fidelr/RayaPubs/AACR2asaDesignSchemaForBibliographicDatabases.pdf [08.02.2015].

13 *ISBD International Standard Bibliographic Description* – www.ifla.org/files/assets/cataloguing/isbd/isbd-cons_20110321.pdf [08.02.2015].

14 M. Crandall, R. Fidel, *op. cit.*, s. 134-135.

zilustrowania faktu, że praktykowane do tej pory zasady katalogowania nie zostały opracowane dla spełnienia założeń konstrukcyjnych bazy danych, jaką jest katalog, ale dla wyrażenia sekwencji czynności najbardziej użytecznej z punktu widzenia dotychczasowej praktyki katalogowania. W przypadku ISBD zasada ta jest właściwie taka sama, choć ujęta w nieco innym sformułowaniu w podpunkcie 8.2: „Tytuł kluczowy przyznawany jest przez sieć ISSN i powinien być podany nawet jeżeli jest identyczny z tytułem właściwym wydawnictwa ciągłego. Tytuł kluczowy powinien być podany, tylko jeżeli numer ISSN, do którego się odnosi, jest podany”. Format zapisu i inne reguły odnoszące się do tytułu kluczowego są również opisane w innych strefach dokumentu i wcale nie chodzi tu o oddzielną sekcję przeznaczoną na taką specyfikację, ponieważ w ISBD uwagi o formacie zapisu są również przemieszane z innymi, ale głównie o rozdział szósty – strefę serii i rozdział siódmy – strefę uwag, tak jak na przykład zasada 7.2.4.7 – uwagi o relacjach pomiędzy tytułami wydawnictw ciągłych¹⁵.

Model danych bibliograficznych opierający się na takich standardach znajduje odzwierciedlenie w słabym powiązaniu między konstrukcją – obiektami modelu koncepcyjnego i korespondującą strukturą w bazie danych. Jak już zaznaczono, źródłem tego stanu jest brak modelu pierwotnie zdefiniowanego na poziomie konceptualnym. Zarówno w procesie programowania, jak i obsługi bazy danych (czy innej aplikacji) zalecane jest rozpoznanie znaczenia przechowywanych w niej informacji w celu właściwej interpretacji danych, tak aby możliwe było stworzenie struktury danych łatwej do opisanego i zrozumiałej także dla użytkownika¹⁶. Z badania, które pod koniec ubiegłego stulecia zostało przeprowadzone na użytkownikach elektronicznych katalogów typu OPAC dziewięćdziesięciu ośmiu bibliotek w USA i Kanadzie, wynikało, że wśród czynników jakościowych mających wpływ na wybór bazy zasobów bibliotecznych, z jakiej korzystają czytelnicy, ogromne znaczenie poza oczywistym – przydatnością i unikalnością jej zasobów – mają łatwość korzystania w rozumieniu przeszukiwania dostępnych treści – intuicyjność i łatwa nawigacja interfejsu oraz zrozumiałe hasła wyszukiwawcze¹⁷.

O ile możliwe (i powszechne) jest oddzielenie logiki aplikacji od jej warstwy prezentacyjnej, o tyle semantyka danych w bazie determinuje możliwości

15 *ISBD International Standard Bibliographic Description...*

16 D. Maier, *Capturing More Meaning in Databases*, „Journal of Management Information Systems” 1984, t. 1, nr 1, s. 33-37 – www.jstor.org/stable/40384524 [08.02.2015].

17 C. Tenopir, E. J. Read, *Database Use Patterns in Public Libraries*, „Reference and User Service Quarterly” 2000, t. 40, nr 1, s. 44-47 – www.jstor.org/stable/20863899 [08.02.2015].

semantyki informacji prezentowanych w interfejsie, mówiąc trywialnie – potrzebujemy danych w odpowiedniej strukturze, aby móc za pomocą odpowiednich komend przetłumaczyć przypisane im wartości na potrzeby warstwy prezentacyjnej. Semantyczne modele danych służą precyzyjnej specyfikacji założeń konceptualnych za pomocą pewnej formalnej struktury oraz terminów i pojęć pochodzących z języka naturalnego. Przykładem może być model związków-encji (model ER) określający cztery typy obiektów: encje, atrybuty encji, relacje między encjami i atrybuty relacji. W modelu tym struktury informacji są łatwo opisywane za pomocą grafów, w których w klasycznej konwencji prostokąty wyrażają encje, romby – relacje, okręgi – atrybuty przypisane encjom i relacjom¹⁸. ER może służyć do modelowania zarówno na poziomie konceptualnym, jak i fizycznym¹⁹. Analiza związków encji została wykorzystana jako podstawa metodologii w opublikowanym w 1997 roku przez International Federation of Library Associations and Institutions (IFLA) raporcie określanym zwykle jako model FRBR. Wybór modelu ER uzasadniony został strukturalnym podejściem do analizy danych²⁰.

W tym miejscu trzeba zaznaczyć, że celem pracy nie jest szczegółowe omawianie modelu FRBR, jednak zostanie on poniżej krótko opisany ze względu na jego znaczenie dla postulowanych zmian w opisie bibliograficznym, implementowanych w Bibliotece Narodowej.

Model FRBR jako model związków encji

FRBR jest modelem na poziomie konceptualnym, który wykorzystuje – jak wspomniano powyżej – analizę związków encji dla zdefiniowania struktury danych bibliograficznych. Został opracowany przez grupę badawczą IFLA w założeniu niezależnie od jakichkolwiek zasad katalogowania i ich zastosowania²¹. W tym sensie FRBR jest uzupełnieniem niewykonanej na wczesnych etapach automatyzacji analizy na poziomie zewnętrznym, niezależnym od rozwiązań systemowych. Jako model abstrakcyjny, definiujący encje i relacje między nimi, pozwala w wygodny sposób wyrażać dane za pomocą grafów – tak jak w modelu ER²². Encje zostały podzielone na trzy grupy. Pierwsza obejmuje wytwory

18 P. Chen, *Entity-Relationship Modeling. Historical Events, Future Trends, and Lessons Learned...*

19 D. Maier, *op. cit.*, s. 38-39.

20 *Functional Requirements for Bibliographic Records. Final Report*, München 1998, s. 9 – www.ifla.org/files/assets/cataloguing/frbr/frbr.pdf [08.02.2015].

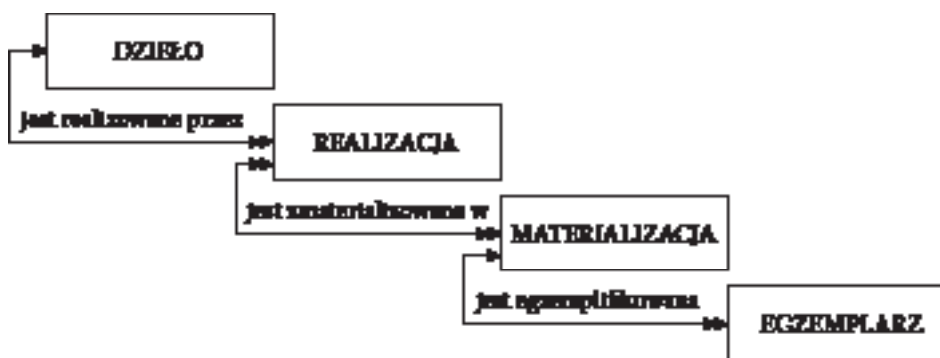
21 R. Callewaert, *FRBRizing your catalogue: the facets of FRBR*, w: S. Chambers, *Catalogue 2.0...*, s. 93.

22 E. Bermes, *Enabling your catalogue for the Semantic Web...*, s. 129.

aktywności intelektualnej i artystycznej, nazywane lub opisywane w rekordach bibliograficznych: dzieło, realizacja, materializacja oraz egzemplarz. Do drugiej grupy należą encje odpowiedzialne za zawartość intelektualną, artystyczną, produkcję i rozpowszechnianie lub będące dysponentami tych wytworów: osoba oraz ciało zbiorowe. Trzecia grupa encji dotyczy przedmiotu treści wytworów intelektualnych i artystycznych: pojęcie, obiekt, wydarzenie, miejsce. Relacje pomiędzy encjami grupy pierwszej zostały zobrazowane za pomocą następującego grafu, gdzie relacje zaznaczone są za pomocą strzałek.

Schemat 1

Encje pierwszej grupy modelu FRBR²³



Mają one charakter *jeden do wielu* (jedno dzieło może mieć wiele realizacji, jedna materializacja może mieć wiele egzemplarzy, ale jedna realizacja będzie realizować tylko jedno dzieło i jeden egzemplarz należy do jednej materializacji) oraz *wiele do wielu* (jedna realizacja może być ucieleśniona w jednej lub wielu materializacjach, a jedna materializacja może odnosić się do jednej lub wielu realizacji)²⁴. Jak wspomniano, praca nie ma na celu dokładnego opisu modelu FRBR, w którym podano szczegółowe definicje encji z przypisanymi im atrybutami oraz charakterystykę relacji. Zarówno dlatego, że cały raport jest łatwo dostępny, jak i dlatego, że jako model opracowany na wysokim poziomie abstrakcji z założenia niezależny od jego implementacji, pozostawia pewne pole do interpretacji, a jego szczegółowa analiza również nie jest celem tego tekstu. Najlepszym przykładem jest samo pojęcie dzieła, które wśród bibliotekarzy praktyków uważane jest jako pojęcie na zbyt wysokim poziomie abstrakcji (nieodwołujące się do określonego bytu fizycznego)

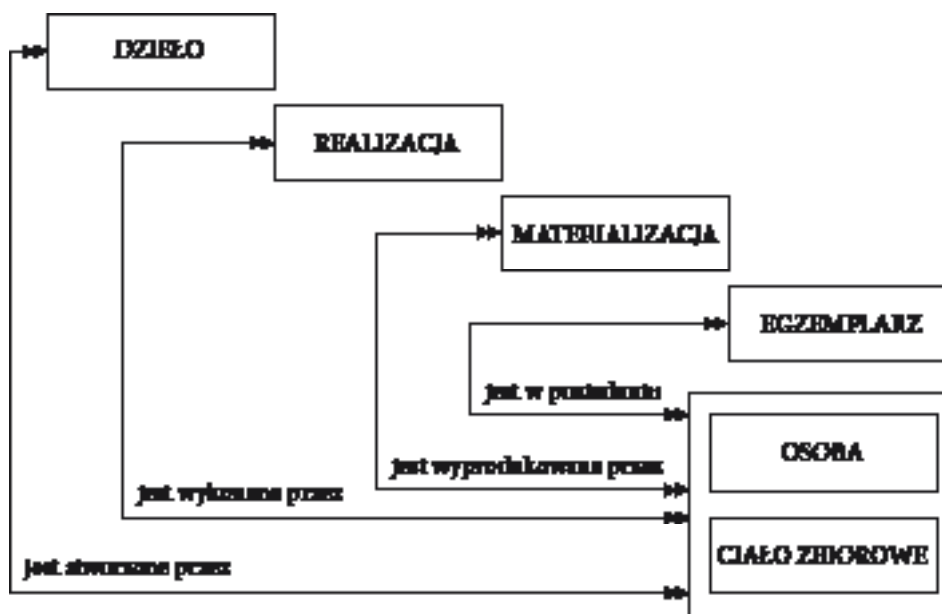
23 *Functional Requirements for Bibliographic Records. Final Report*, s. 13.

24 *Ibidem*, s. 13-14.

– pewien intelektualny twór z przypisanym autorstwem, który w systemie danych bibliograficznych reprezentowany może być jedynie jako pewien klaster opisów o atrybutach wspólnych dla różnych wydań, tłumaczeń i formatów. Zestaw wartości, który może przyjąć każdy z powyższych atrybutów, jest trudny do precyzyjnego określenia i w zależności od implementacji jest różnie interpretowany, co wpływa na rozumienie samego pojęcia *dzieło*²⁵. Dla uzupełnienia pierwotnego schematu FRBR przytaczamy jedynie poniżej przedstawione w raporcie FRBR grafy obrazujące relacje między encjami grupy drugiej oraz relacje między encjami grupy trzeciej (spośród których wszystkie mają charakter *wiele do wielu*)²⁶.

Schemat 2 i 3

Encje drugiej i trzeciej grupy modelu FRBR²⁷

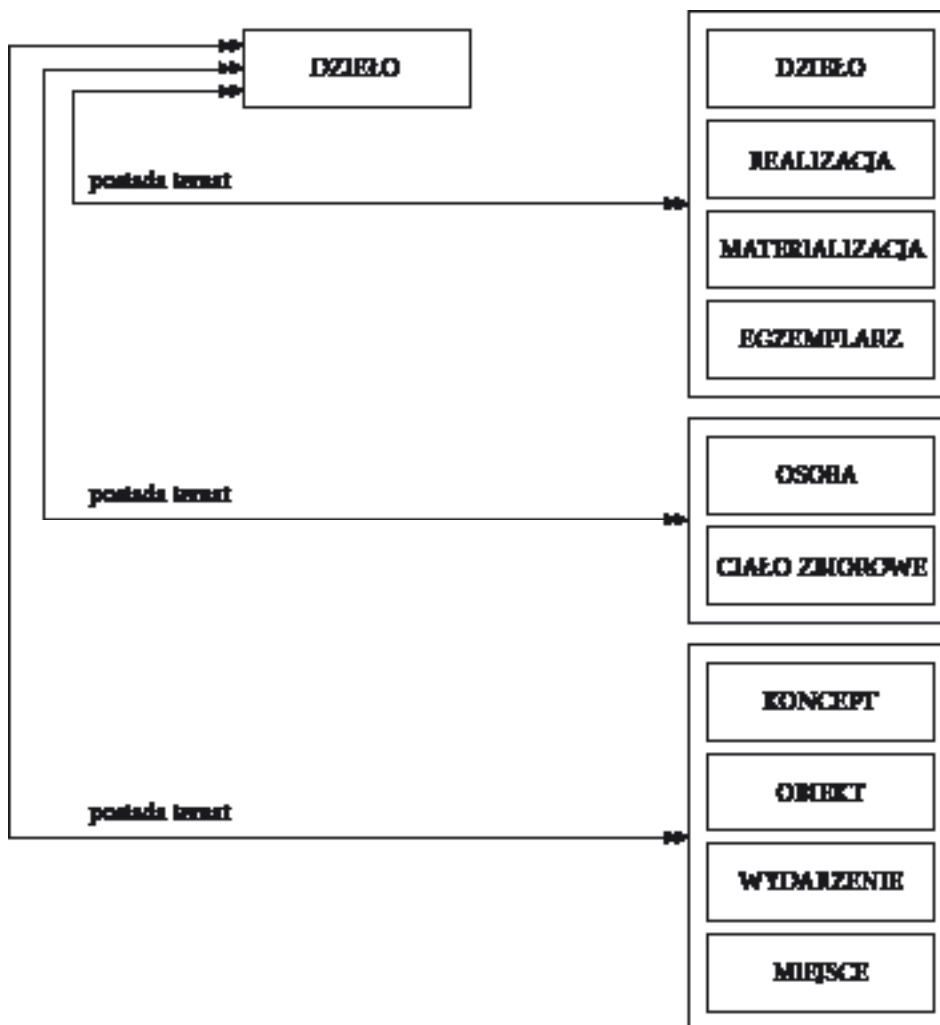


Model FRBR kładzie nacisk na zdefiniowanie zależności – relacji łączących katalogowany obiekt nie tylko z twórcą i innymi wydaniem, ale każdym bytem,

25 C. J. Godby, *The Relationship between BIBFRAME and OCLC's Linked-Data Model of Bibliographic Description: a Working Paper*, Dublin (OH) 2013, s. 11 – www.oclc.org/content/dam/research/publications/library/2013/2013-05.pdf [08.02.2015].

26 *Functional Requirements for Bibliographic Records. Final Report*, s. 15-16.

27 *Ibidem*, s. 14-15.



jaki może być z nim związany²⁸. Jest to, pomimo trudności w interpretacji, duża zaleta modelu, ponieważ semantyka katalogu wyrażona za pomocą encji może przynosić korzyści na różnych płaszczyznach – z jednej strony stanowi model konceptualny niezależny od implementowanej technologii, z drugiej strony jest graficzną reprezentacją danych, która – jak potwierdziło badanie Leitheisera i Marcha – stosuje semantykę wygodniejszą do zrozumienia dla użytkowników potencjalnego systemu – zarówno katalogujących, jak i odbiorców katalogu²⁹. Ta sama metodologia modelowania związków encji została wyko-

28 S. Batley, A. Welsh, *Practical Cataloging...*, s. 8.

29 R. L. Leitheiser, S. T. March, *The Influence of Database Structure Representation*

rzyszta w opublikowanym przez IFLA raporcie – najpierw jako szkic (2007) i w ostatecznej wersji (2009): *Functional Requirements for Authority Data. A Conceptual Model (FRAD)*, prowadząc wraz z opublikowanym wcześniej modelem FRBR do koncepcji encji w nowym standardzie katalogowania RDA³⁰. W modelu FRAD rozbudowany został opis dotyczący relacji łączących encje z grupy pierwszej FRBR z osobą, rodziną i ciałem zbiorowym³¹. Terminy zastosowane w RDA zostały przejęte bezpośrednio z modelu FRBR, ale trzeba pamiętać, że często ich znaczenie jest inne niż w obowiązujących wcześniej standardach katalogowania, czego dobrym przykładem jest wspomniane pojęcie *działa*³².

Standard RDA będący sukcesorem AACR2, stosowany od 2013 roku przez Bibliotekę Kongresu, opisuje w sposób formalny różne rodzaje relacji potrzebnych dla połączenia encji FRBR oraz FRAD różnych poziomów w znormalizowany sposób³³. W przeciwieństwie do wcześniejszych standardów, zasady RDA nie są skonstruowane w podziale na typy dokumentów, ponieważ standard opiera się o ogólne zasady, dla których przewidziane są wyjątki lub opcje do zastosowania w różnych okolicznościach (takich jak specyficzny format dokumentu). O ile AACR2 był w pełni kompatybilny z ISBD, o tyle stosunkowo nowy standard RDA spotyka się z krytyką jako nie w pełni uwzględniający specyfikę zasad ISBD. Standard ISBD wprowadził wprawdzie niektóre elementy modelu FRBR w strefie 0 dotyczącej form zawartości i typu mediów, dające się łatwo mapować do wartości *rdacontent* i *rdamedia*, jednak w pozostałych kwestiach pomimo działań IFLA Linked Data Study Group nie dokonano pełnego uzgodnienia standardu ISBD z RDA i modelem FRBR. W związku z tym RDA Toolkit³⁴ pozwala na rozmieszczenie zasad RDA względem obszarów ISBD³⁵. W praktyce dla bibliotek stosujących zasady katalogowania oparte o standard ISBD – takich jak Biblioteka Narodowa – oznacza to możliwość zastosowania elementów opisu bibliograficznego wprowadzonych przez RDA w takim zakresie, w jakim jest to potrzebne dla optymalnego odzwierciedlenia encji i ich atrybutów zdefiniowanych na poziomie koncepcyjnego modelu danych

on *Database System Learning and Use*, „Journal of Management Information Systems” 1996, t. 12, nr 4, s. 206 – www.jstor.org/stable/40398183 [08.02.2015].

30 S. Batley, A. Welsh, *Practical Cataloging...*, s. 15.

31 *Functional Requirements for Authority Data. A Conceptual Model*, 2009 (rev. 2013), s. 31-46 – www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf [08.02.2015].

32 S. Batley, A. Welsh, *Practical Cataloging...*, s. 91.

33 R. Callewaert, *FRBRizing your catalogue...*, s. 105.

34 Interaktywna wersja przepisów RDA – www.rdatoolkit.org [08.02.2015].

35 S. Batley, A. Welsh, *Practical Cataloging...*, s. 91.

gromadzonych w bazie danych katalogu. Proces ten nosi nazwę *hybrydyzacji rekordów* i dotyczy dodawania elementów standardu katalogowania RDA do rekordów, które nie powstały w tym standardzie, bez względu na to, czy dokonywane jest to manualnie czy automatycznie. Jest to rozwiązanie znajdujące zastosowanie dla rekordów, które stworzono w starszych standardach (jak ISBD i AACR2). W efekcie powstaje rekord hybrydowy, który nie jest utworzony w standardzie RDA, ale zawiera elementy różnych standardów³⁶.

Pojawienie się standardu RDA spowodowało konieczność wprowadzenia do formatu MARC 21 szeregu pól, które wymagają zastosowania słownictwa kontrolowanego. Wzbogacenie opisu bibliograficznego o tak zdefiniowane atrybuty daje możliwość zastosowania Deskryptorów BN w rekordach w postaci klucza obcego (inaczej klucza zewnętrznego – *foreign key* – atrybutu, który w strukturze relacyjnych baz danych stanowi odnośnik do klucza głównego w tabeli nadrzędnej, w tym wypadku odpowiadającej właściwej encji wyrażonej za pomocą deskryptora z bazy deskryptorów). Jest to praktyka pożądana, ponieważ dzięki Deskryptorom BN możliwe jest uproszczenie haseł, a to sprzyja normalizacji bazy danych (database normalization) katalogu³⁷. Postać normalna bazy danych, czyli taka, w której nie występuje redundancja, może być osiągnięta tylko jeżeli wartości atrybutów są zatomizowane (niepodzielne). Chociaż obecnie trudno sobie wyobrazić całkowitą normalizację bibliograficznej bazy danych bez ryzyka utraty pewnych informacji lub stworzenia relacji pozabawionych odzwierciedlenia w świecie realnym, zastosowanie rozwiązań takich jak Deskryptory BN wydaje się być krokiem w dobrym kierunku. Wynika to z roli, jaką dla normalizacji bazy danych odgrywa nie tylko składnia zgromadzonych w niej danych i relacji między nimi, ale także jej semantyka, czyli to, co dane i relacje istotnie oznaczają, oraz sposób, w jaki mogą być względem siebie połączone. Konsekwencją takiego spojrzenia jest przyjęcie zasady, że aby encje i relacje w bazie były dobrze zdefiniowane, powinny móc one – a także ich atrybuty – być opisane za pomocą języka naturalnego³⁸.

Język naturalny umożliwia najwyższy stopień wyspecjalizowania i elastyczności w reprezentowaniu i wyszukiwaniu informacji. Słownictwo kontrolowane bazą deskryptorów umożliwia jednocześnie zmniejszenie swoistych problemów

36 *Report of the PCC Post-Implementation Hybrid Bibliographic Records Guidelines Task Group*, s. 3-4 – www.loc.gov/aba/pcc/rda/RDA%20Task%20groups%20and%20charges/PCC-Hybrid-Bib-Rec-Guidelines-TG-Report.docx [08.02.2015].

37 Normalizacja baz danych w rozumieniu informatycznym; zob. *Opis podstaw normalizacji baz danych* – <http://support.microsoft.com/pl-pl/kb/283878> [08.02.2015].

38 Q. F. Stout, P. A. Woodworth, *Relational Databases*, „The American Mathematical Monthly” 1983, t. 90, nr 2, s. 116.

(takich jak złożoność semantyczna i synonimia wyrażen języka naturalnego)³⁹. W użyciu języka naturalnego dla nazwania elementów bazy danych chodzi nie o ujęcie jak największej części języka naturalnego w słownictwie kontrolowanym bazą deskryptorów, ale o nadanie encjom, relacjom i atrybutom takich wartości, które mają swoje odzwierciedlenie w języku naturalnym i włączenie ich do bazy, jeżeli jest to konieczne, a nie o nadawanie im terminów o zmodyfikowanym zakresie znaczenia. Terminy, które ujęte są w kartotece, powinny być również rozpoznawane w codziennym języku. Zbytня złożoność semantyczna słownictwa może prowadzić do jego niezrozumienia⁴⁰. Między innymi dlatego Deskryptory BN zakładają uproszczenie dotychczasowego Języka Haseł Przedmiotowych BN. Polega to na przypisaniu poszczególnym encjom i ich atrybutom niepodzielnych wartości i może być przeprowadzone na tyle głęboko, na ile pozwoli to uniknąć terminów dwuznacznych. Najważniejszą przewagą katalogu biblioteki nad dużymi wyszukiwarkami internetowymi są bowiem i nadal pozostaną możliwości rozróżniania i ujednoznacznienia jednobrzmiących haseł⁴¹.

Katalog biblioteki i jego dostępność w kontekście rozwoju sieci semantycznej

Trzeba w tym miejscu zauważyć, że pomimo tej niewątpliwej przewagi, jaką jest wieloletnia praktyka rozróżniania i ujednoznacznienia jednobrzmiących haseł, wraz z rozpowszechnieniem dostępu do Internetu, jego użytkownicy zaczęli przeprowadzać swoje kwerendy zwykle poza katalogami bibliotek, przede wszystkim dlatego, iż okazały się one zbyt trudne w korzystaniu w porównaniu z innymi platformami⁴². Badania wykazują, że studenci w pierwszej kolejności w poszukiwaniu wiedzy będą korzystać z Internetu, a znaczna liczba użytkowników bibliotek jest gotowa godzić się na rozwiązania, które wygodę przedkładają nad kompleksowość⁴³. Wynika to z pewnej znanej w aukach o informacji prawidłowości, według której ludzie wydają się być zadowoleni z tego, co można znaleźć łatwo, zamiast poświęcać więcej wysiłku na szukanie źródeł wysokiej jakości. Mamy tu do czynienia z obserwowaną również w wielu innych dziedzinach *zasadą najmniejszego wysiłku* (*principle of least effort*),

39 H. Chu, *Information Representation and Retrieval in the Digital Age*, Medford (NJ) 2010, s. 17.

40 D. B. Cleveland, A. D. Cleveland, *Introduction to Indexing and Abstracting*, Englewood (CO) 2013, s. 65.

41 S. Batley, A. Welsh, *Practical Cataloging...*, s. 12.

42 A. Christensen, *Next generation catalogues...*, s. 4.

43 A. Tariq, G. Puja Anand, S. Jaideep, *Developing Sustainable Digital Libraries: Socio-Technical Perspectives*, Hershey (PA) 2010, s. 134.

w nauce o informacji czasem określaną jako *zasada skąpstwa poszukujących informacji*, która jest też znana jako prawo Zipfa⁴⁴. W odniesieniu do nauk o informacji prawo to stwierdza po prostu, że „każda jednostka wybierze opcję uzyskania informacji, która wymaga najmniejszego wysiłku, i zaprzestanie dalszego poszukiwania po uzyskaniu wyników na poziomie akceptowalnym”. Ta prosta reguła, ma jak wiadomo, poważne konsekwencje dla bibliotek. Wyznacza trend, zgodnie z którym powinny one koncentrować się na zwiększaniu użyteczności i intuicyjności narzędzi, jakie dawałyby dostęp do ich zasobów⁴⁵.

Również ze względu na wyżej wspomnianą zasadę rośnie znaczenie tych spośród internetowych systemów wyszukiwania, które funkcjonują jako kanały prowadzące do innych źródeł informacji⁴⁶. W konsekwencji strategia wyszukiwawcza we współczesnym katalogu musi opierać się na tym, jak wyszukiwarki obchodzą się z danymi, aby zapewnić usługi dla odbiorców informacji w sieci internetowej. Interfejs biblioteczny nie stanowi w takim ujęciu już tylko interfejsu katalogu biblioteki, ale może być również wyszukiwarką, w której dostępne są powiązane informacje dotyczące konceptów takich jak dzieła, tematy, osoby (innymi słowy różne kategorie wspomnianych wcześniej encji modelu FRBR). Dzięki unikatowym adresom internetowym – URI (Uniform Resource Identifiers), reprezentującym poszczególne koncepty (naturalnie nie wszystkie reprezentowane w Internecie koncepty pochodzą z modelu FRBR, który dotyczy konkretnie danych bibliograficznych i danych z nim związanych) możliwe jest stworzenie stron internetowych agregujących informacje pochodzące z różnych źródeł. Inicjatywy takie jak Open Library⁴⁷, LibraryThing⁴⁸ czy Goodreads⁴⁹ zostały stworzone jako strony internetowe „oparte na poziomie encji” (*entity-level-based web pages*), w których zagregowane dane na temat autorów, dzieł, realizacji, materializacji, przedmiotu, *etc.*, układają się w pełną stronę. W zależności od zapytania, użytkownik trafia na jeden z poziomów takiej strony wskazującej linki do pozostałych poziomów prowadzących do końcowych usług takich jak sprzedaż lub wypożyczenie informacji w formie cyfrowej albo egzemplarza fizycznego, czy bezpośredni dostęp do publikacji cyfrowej⁵⁰. Ponieważ wyszukiwarki i komercyjne strony internetowe zawierające materiały bibliograficzne korzystają ze sposobu

44 Ze względu na prace lingwisty George’a Kingsleya Zipfa, który badał jej wpływ na dystrybucję słów.

45 J. R. Matthews, *The Digital Library Survival Guide*, Carlsbad 2011, s. 117-118.

46 H. Chu, *Information Representation and Retrieval in the Digital Age...*, s. 152.

47 Open Library – www.openlibrary.org [08.02.2015].

48 LibraryThing – www.librarything.com [08.02.2015].

49 Goodreads – www.goodreads.com [08.02.2015].

50 R. Callewaert, *FRBRizing your catalogue...*, s. 107.

wyszukiwania i prezentacji danych opartego na encjach, użytkownicy katalogów bibliotek oczekują tego samego od systemów narzędzi przeszukiwania zasobów bibliotecznych. Wyszukiwarki wprowadziły funkcjonalności takie jak karty informacji oparte na encjach do swoich wyników wyszukiwania, aby pokazywać użytkownikom dane odpowiadające ich zapytaniom w szerszych kontekstach powiązań encji odpowiadających tym zapytaniom z innymi zasobami. Systemy biblioteczne również mają potencjał, aby przekształcić katalog biblioteki z systemu opartego na wyszukiwaniu i wyświetlaniu rekordu na system wyszukiwania wyników oparty na encjach oraz prezentowania kart informacji o nich pod warunkiem, że systemy katalogowe, na których są zbudowane, również opierają się na encjach. Systemy takie mogą także zapewnić bardziej intuicyjny interfejs użytkownika i mogą mieć dodatkową wartość w podpowiadaniu scenariuszy wyszukiwawczych, jako że użytkownicy poprzez powiązania-linki są kierowani do kolejnych encji (encji powiązanych w korespondującym z danym systemem modelu ER)⁵¹. Wyraźne i jednoznaczne oznaczenie relacji jest kluczowym elementem FRBR, a tym samym jedną z cech, które czyni ten model, tak jak i standard RDA na nim oparty, gotowym narzędziem wyrażania informacji poza środowiskiem katalogu biblioteki i poziomem najbardziej znanym bibliotekom, jakim jest rekord bibliograficzny⁵². Produktem jest wówczas nie tradycyjnie pojmowany katalog, ale surowiec, z którego buduje się informację. Dane tworzone w ten sposób mogą być wykorzystywane w szerszym niż dotychczas kontekście. W systemie połączonych danych mogą one pełnić odmienną funkcję poza swoim oryginalnym kontekstem i biblioteki muszą sobie zdawać sprawę z wzajemnych powiązań między danymi. Potencjalnymi użytkownikami tych danych mogą być zarówno ludzie, jak i interfejsy maszynowe, co jest istotnym argumentem za stosowaniem otwartych standardów i znormalizowanych form zapisu informacji⁵³. Katalogowanie oparte na encjach daje katalogującym możliwość poprawienia jakości jednoznaczności danych, a w konsekwencji jakości wyników przeszukiwania przez użytkowników, między innymi dlatego, że poprawki dokonane na jednym poziomie grafu związków encji mogą być w takim systemie współdzielone z innymi poziomami (pozostałymi poziomami modelu FRBR)⁵⁴.

Sieć danych powiązanych relacjami za pomocą reguł logicznych oraz metod heurystycznych łącznie pozwala na syntetyzowanie znaczenia pojęć używanych w języku naturalnym. Ta wizja leży u podstaw rozwijanej przez

51 Th. Fons, *Authorities, Entities and Communities*, s. 2 – <http://library.ifla.org/1034/1/086-fons-en.pdf> [08.02.2015].

52 S. Batley, A. Welsh, *Practical Cataloging...*, s. 95.

53 R. Callewaert, *FRBRizing your catalogue...*, s. 112.

54 Th. Fons, *Authorities, Entities and Communities...*, s. 3.

konsorcjum W3C sieci semantycznej, która ma na celu połączenie niezależnie rozwijanych systemów danych bez potrzeby globalnego porozumienia co do warunków takiego połączenia⁵⁵. Celem tego artykułu nie jest szczegółowe zgłębienie siedmiowarstwowej struktury sieci semantycznej. Na jego potrzeby zaznaczyć trzeba, że pod pojęciem sieci semantycznej (*Semantic Web*) rozumiany jest powszechnie zestaw standardów, dzięki którym możliwe jest budowanie sieci połączonych danych, przede wszystkim model metadanych RDF (Resource Description Framework), języki tworzenia ontologii takie jak OWL czy RDF Schema i standardy przetwarzania danych jak SPARQL. Ideą tworzenia sieci danych połączonych jest publikowanie przez różne organizacje nieobjętych prawem autorskim metadanych zapisanych w standardzie wyrażen RDF. Składają się one z trzech elementów: podmiotu, predykatu i obiektu, z których z kolei każdy zidentyfikowany jest przez odpowiedni identyfikator URI, tworząc sieć relacyjną⁵⁶. Przedstawienie koncepcji sieci semantycznej jest o tyle istotne w kontekście zmian wprowadzanych w Bibliotece Narodowej, że obecnie wdrażane na poziomie katalogu rozwiązania – jego eferberyzacja z wykorzystaniem pól formatu MARC 21 wprowadzonych wraz z standardem RDA – są w pierwszej kolejności krokiem w kierunku stworzenia relacji między encjami opisu. W szerszym kontekście są one natomiast podstawą modelu relacyjnego i logicznego metadanych pochodzących z tego opisu zapisanego w standardzie RDF, który pozwoli włączyć katalog Biblioteki Narodowej w struktury globalnych danych połączonych w możliwie najbardziej efektywny sposób. Rozwijane obecnie na świecie modele danych połączonych (Linked Data Models) przeznaczone dla opisu zasobów bibliotecznych, takie jak model OCLC czy projekt Bibframe zainicjowany przez Library of Congress⁵⁷, zakładają, że hierarchia encji FRBR jest fundamentalna dla opisu zasobów bibliotecznych i operacji między nimi, chociaż ze względu na specyfikę opisu bibliograficznego encje te zdefiniowane są w różnych modelach na innych poziomach⁵⁸. Oczywiście model Biblioteki Narodowej nie może być tworzony w oderwaniu od innych i jego rozwój zakłada analizę podobnych projektów, z których większość – tak jak wspomniany projekt OCLC i Bibframe – jest w fazie postępujących prac. Taka analiza jest jednak tematem zbyt obszernym na potrzeby tego artykułu, ale wynikające również z niej znaczenie eferberyzacji katalogu i granulacji danych były istotnym czynnikiem dla wprowadzanych zmian opracowania, przede wszystkim dlatego, że w założeniu modelu danych połączonych powinny one

55 H. Chu, *Information Representation and Retrieval in the Digital Age...*, s. 274.

56 E. Bermes, *Enabling your catalogue for the Semantic Web...*, s. 119-120.

57 Bibliographic Framework Initiative – www.loc.gov/bibframe/ [08.02.2015].

58 C. J. Godby, *The Relationship between BIBFRAME...*, s. 14-23.

tworzyć spójny model związków encji, a określenie zgodności semantycznej pomiędzy heterogenicznymi zbiorami danych jest warunkiem koniecznym do ich zintegrowania⁵⁹.

Koncepcja zmian w opisie bibliograficznym postulowana przez Bibliotekę Narodową zakłada z jednej strony eferberyzację katalogu poprzez wprowadzenie elementów opisu RDA. Ponieważ nie zakłada się odejścia od dotychczas przyjętego formatu katalogowania MARC 21, wprowadzenie elementów RDA i FRBR zamierzone jest na tym etapie tylko w takim zakresie, w jakim pozwala na to format (w szczególności poprzez dodanie nowych pól opisu formalnego i bibliograficznego z grupy 3XX). W związku z tym obecnie prace cały czas są prowadzone głównie na niższych poziomach modelu FRBR, przy założeniu, że część spośród rozbudowywanych przy pomocy nowych elementów opisu encji i atrybutów stanowić będzie podstawę do mapowania wyższych poziomów i relacji również poza samym środowiskiem katalogu w rozwiniętej strukturze danych połączonych. Z drugiej strony zachodzi proces uproszczenia dotychczasowego słownictwa i składni JHP BN, a następnie przekształcenia go w Deskryptory BN, które stanowiłyby słownictwo odpowiadające potrzebom relacyjnych struktur zgodnie z zasadą, że język jest narzędziem, w którym słowa i terminy są instrumentami i ich znaczenie jest określone przez sposób ich wykorzystania⁶⁰. To zagadnienie w sposób bardziej szczegółowy opisane jest w kolejnej części artykułu.

2. Deskryptory BN – uproszczony JHP BN

Baza Deskryptorów Biblioteki Narodowej
i jej funkcje w fasetowych systemach wyszukiwania informacji

Mechanizacja i automatyzacja katalogów bibliotecznych w systemach komputerowych pomimo ponaddwudziestoletniej praktyki w tym zakresie pozostaje nadal jednym z ważniejszych wyzwań stojących przed bibliotekami w Polsce. Zasadniczymi i ciągle aktualnymi problemami obu działań są z jednej strony świadoma decyzja, które elementy komunikacji użytkownika z katalogiem i w jakim zakresie poddawać jednemu lub drugiemu procesowi, z drugiej zaś ciągle dostosowywanie sposobu wyrażania elementów opisu bibliograficznego do rzeczywistości oraz wymagań środowiska sieciowego i relacyjnych baz danych. Przez automatyzację katalogu rozumie się tu całkowite zastąpienie

59 E. S. Soofi, H. Zhao, *Exploring Attribute Correspondences across Heterogenous Databases by Mutual Information*, „Journal of Management Information Systems” 2006, t. 22, nr 4, s. 306 – www.jstor.org/stable/40398821 [08.02.2015].

60 R. Raieli, *Multimedia Information Retrieval: Theory and Techniques*, Oxford, Sawston 2013, s. 25.

dowolnej czynności wykonywanej manualnie przez proces automatyczny⁶¹. Jednym z przykładów takiego działania może być proces zastąpienia wyszukiwania przez indeks wyszukiwaniem fasetowym, w którym podstawowymi jednostkami informacji są ustrukturyzowane instancje encji danych bibliograficznych, ich relacje w bazie danych, a także kontekst, w jakim zostały osadzone w rekordzie bibliograficznym. Mechanizacja natomiast nie eliminuje manualnego charakteru wyszukiwania w zbiorze danych, a jedynie go usprawnia, czego przykładem może być elektroniczny indeks odwzorowujący układ katalogu kartkowego⁶².

Obserwując proces informatyzacji polskich katalogów bibliotecznych, nie sposób oprzeć się wrażeniu, iż zarówno organizacja danych bibliograficznych, jak i możliwości ich przeszukiwania pozostają ciągle jeszcze pod dużym wpływem systemów manualnych (katalogów, bibliografii, indeksów). Abstrahując od częściowej strukturyzacji danych bibliograficznych, która była następstwem przyjęcia formatu MARC 21, katalogi te w większości przypadków pozostają nadal jedynie elektroniczną wersją katalogów manualnych o analogowej strukturze metadanych, co szerzej omówiono w pierwszej części artykułu. Przeniesienie indeksu haseł przedmiotowych, formalnych oraz analogowo zapisanej informacji bibliograficznej, czyli nagłówków kart katalogu i ich zawartości, do systemu komputerowego praktycznie nie wpływa na zmianę sposobu wyszukiwania informacji. Bez gruntownej przebudowy modelu metadanych pozostaje on w wielu aspektach taki sam jak w katalogu kartkowym i nie wykorzystuje wszystkich możliwości technicznych, jakie mogłyby zapewnić nawet najlepsze systemy organizacji, przetwarzania i wymiany informacji.

Współcześnie rozwijane koncepcje systemów informatycznych bibliotek i ich partnerów wymagają od bibliograficznych baz danych ustrukturyzowanej i ujednoliconej informacji zarówno na poziomie opisu bibliograficznego (rekordu) jak i elementu opisu bibliograficznego (encji i jej atrybutów) oraz wartości, jaką ten element przyjmuje (instancja encji wyrażonej deskryptorem)⁶³. Deskryptor w takiej strukturze danych pełni funkcję preferowanego

61 Zob. hasło „automatyzacja” – *Słownik języka polskiego PWN*, red. nauk. M. Szymczak, t. 1, A-K, Warszawa 1998, s. 95; por. też hasło „katalog zautomatyzowany” – *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*, oprac. B. Bojar, Warszawa 2002, s. 116.

62 Zob. hasło „mechanizacja” – *Słownik języka polskiego PWN*, t. 2, P-L, s. 125; por. też hasło „katalog manualny” – *Słownik encyklopedyczny informacji...*, s. 115.

63 Model ER został omówiony w rozdziale *Bazy danych bibliograficznych a relacyjne bazy danych*, s. 152-157.

znaku⁶⁴, którego desygnatami są pojęcia bytu⁶⁵ przedstawionego w publikacji lub stanowiący inną instancję encji opisu bibliograficznego. Z uwagi na to, że różne rozumienie bytu – sposób bytowania, zagadnienie istoty i szczegółowe koncepcje filozoficzne – nie wpływają w żaden sposób na praktyczny wymiar stosowania Deskryptorów Biblioteki Narodowej termin *byt* definiuje się w sposób uniwersalny jako element rzeczywistości, który istnieje, posiada istotę i może mieć dowolne cechy akcydentalne. W tym rozumieniu deskryptor ma znaczenie szersze od definicji stosowanej w teorii języków informacyjno-wyszukiwawczych, w której utożsamia się go jedynie z podstawową jednostką leksykalną języka deskryptorowego⁶⁶. Wskazane jest tu przyjęcie znaczenia zgodnego z terminologią międzynarodową stosowaną między innymi w dokumentach IFLA⁶⁷, normach NISO⁶⁸ i źródłach informacyjnych⁶⁹, w których deskryptor definiuje się jako preferowany termin służący do wyrażania pojęć i cech publikacji. Stosowanie właściwego rozumienia

64 Najbardziej rozpowszechnione słowo lub wyrażenie w języku naturalnym.

65 Termin „pojęcie” rozumiany jako obraz-znak treści bytu utworzony przez człowieka w intelektualnym poznaniu w celu oznaczenia poznanej rzeczy; *Powszechna encyklopedia filozofii*, kom. nauk. M. A. Krąpiec [et al.], t. 8, P-S, Lublin 2007, s. 318-319; zob. też poznanie pojęciowe w ujęciu platońsko-arystotelesowskim i relacja między słowem, pojęciem i rzeczą. M. A. Krąpiec, *Język i świat realny*, Lublin 1995, s. 50-65; por. triadyczna koncepcja znaku, Ch. S. Peirce, *Wybór pism semiotycznych*, wyb. dokonała H. Buczyńska-Garewicz, przekł. R. Mirek, A. J. Nowak, Warszawa 1997, s. 262.

66 Wiesław Babik wskazuje na odmienne znaczenie terminu deskryptor w polskiej terminologii języków informacyjno-wyszukiwawczych i terminologii anglojęzycznej (amerykańskiej i brytyjskiej). W. Babik, *Polska terminologia języków informacyjno-wyszukiwawczych w dobie globalizacji*, „Zagadnienia Informacji Naukowej” 2006, nr 1, s. 10; zob. też PN-N-09018: *Tezaurus jednojęzyczny: zasady tworzenia, forma i struktura*, Warszawa 1992, s. 2; *Słownik encyklopedyczny informacji...*, s. 47; J. Woźniak, *Dobry tezaurus, czyli jaki?*, „Notes Biblioteczny” 2006, nr 2, s. 47; W. Klenczon, *Indeksowanie zawartości treściowej w bibliograficznych bazach danych*, – www.ebib.pl/publikacje/matkonf/mat19/klenczon.php [08.02.2015].

67 Zob. definicja terminu „deskryptor” – *Guidelines for Subject Access in National Bibliographies*, Berlin, Munich 2012, s. 79.

68 *Guidelines for Abstracts* – ANSI/NISO Z39.14-1997, s. 1 – www.niso.org/apps/group_public/download.php/6609/Guidelines%20for%20Abstracts.pdf [08.02.2015]; *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* – ANSI/NISO Z39.19-2005 (R2010), s. 158, 164 – www.niso.org/apps/group_public/download.php/12591/z39-19-2005r2010.pdf [08.02.2015].

69 M. A. Drake, *Encyclopedia of library and information science*, New York 2003, s. 813-818; J. M. Reitz, *Online Dictionary for Library and Information Science* – www.abc-clio.com/ODLIS/odlis_d.aspx [08.02.2015].

deskryptora, jako *znaku pojęcia bytu* jest szczególnie istotne zarówno dla procesu globalnej wymiany danych, gdzie operuje się często zbiorami tych samych pojęć bytów (osób, instytucji, obiektów itp.), ale wyrażonych różnymi znakami (np. różne warianty językowe), jak też dla ustalania relacji między deskryptorami, które odczytuje się na poziomie pojęć, a nie znaków, ponieważ z natury nie posiadają one żadnego znaczenia, a tylko do czegoś odsyłają. Ponieważ desygnat deskryptora nie jest predestynowany do bycia instancją jednej, i tylko jednej określonej encji, więc może być traktowany jako instancja kilku różnych encji, np. *mapa* – typ publikacji; *mapa* – przedmiot publikacji; *Jan Kowalski* – autor publikacji; *Jan Kowalski* – przedmiot publikacji; *lekarze* – przedmiot publikacji; *lekarze* – atrybut osoby w jej rekordzie wzorcowym; *lekarze* – atrybut instytucji w jej rekordzie wzorcowym, itp. W takim modelu danych niezbędne jest narzędzie organizujące instancje poszczególnych encji opisu bibliograficznego oraz zarządzające relacjami zachodzącymi nie tylko między desygnatami, ale także między ich atrybutami. Optymalnym rozwiązaniem wydaje się jedna, zintegrowana z systemem wyszukiwawczym baza deskryptorów o funkcji tezauryusa / słownika wyszukiwawczego, która jest już nie tylko narzędziem pomocniczym przy wyborze formy ujednocionej, lecz stanowi główny element systemu wyszukiwawczego⁷⁰. Baza taka staje się też narzędziem kontroli wszystkich autoryzowanych punktów dostępu, odzwierciedla atrybuty encji dla poszczególnych instancji takie jak, zakres znaczenia, nota o zakresie stosowania i własności bytów denotowanych przez deskryptory na poziomie ich pojęć. Baza odzwierciedla też relacje logiczne, a właściwie stosunki teoriomnogościowe, pomiędzy desygnatami deskryptorów lub pomiędzy deskryptorami i ich wariantami w kontekście ich desygnatów⁷¹.

Relacja logiczna równoważności między deskryptorem i jego wariantami

Jeżeli D oznacza deskryptor, $W_{(D)}$ oznacza wariant deskryptora, x oznacza dowolny byt, $P_{(x)}$ oznacza pojęcie bytu, to relację równoważności można

70 Szerzej o funkcji tezauryusa wyszukiwawczego zob. E. Chmielewska-Gorczyca, *Funkcje tezauryusa w systemie informacyjno-wyszukiwawczym*, „Zagadnienia Informatyki Naukowej” 1995, nr 1/2, s. 3-17.

71 W literaturze stosowane są odmienne nazwy dla niektórych relacji / stosunków, np. u L. Borkowskiego zamienność, u K. Trzęsickiego równoważność, u Z. Ziemińskiego zamienność zakresów. L. Borkowski, *Elementy logiki formalnej*, Warszawa 1980, s. 77; K. Trzęsicki, *Logika z elementami semiotyki i retoryki*, Białystok 2009, s. 67; Z. Ziemiński, *Logika praktyczna*, Warszawa 1994, s. 39.

wyrazić następująco: $W_{(D)}$ pozostaje w relacji równoważności z D , wtedy i tylko wtedy, gdy $W_{(D)}$ denotuje $P_{(x)}$ i D denotuje $P_{(x)}$, np. *telewizor* i *odbiornik telewizyjny* pozostają w relacji równoważności, ponieważ denotują pojęcie tego samego bytu; pszczelarstwo i pszczelnictwo nie pozostają w relacji równoważności, ponieważ denotują pojęcia odrębnych bytów: *rzemiosło* i *naukę o tym rzemiośle*⁷².

Relacja logiczna podrzędności między pojęciami bytów denotowanych przez deskryptory

Jeżeli x i y oznaczają byty, $P_{(x)}$ i $P_{(y)}$ oznaczają pojęcia bytów x i y , $D_{P_{(x)}}$ i $D_{P_{(y)}}$ oznaczają deskryptory pojęć bytów x i y , to relację podrzędności można wyrazić następująco: $P_{(x)}$ denotowane przez $D_{P_{(x)}}$ pozostaje w relacji podrzędności do $P_{(y)}$ denotowanego przez $D_{P_{(y)}}$, wtedy i tylko wtedy, gdy każdy desygnat $D_{P_{(x)}}$ jest desygnatem $D_{P_{(y)}}$ i istnieje przynajmniej jeden desygnat $D_{P_{(y)}}$, który nie jest desygnatem $D_{P_{(x)}}$, np. pojęcie *podgrzybek zwyczajny* pozostaje w relacji podrzędności w stosunku do pojęcia *grzyby*, ponieważ każdy *podgrzybek zwyczajny* jest *grzybem*, a tylko niektóre *grzyby* są *podgrzybkami zwyczajnymi*.

Relacja logiczna nadrzędności między pojęciami bytów denotowanych przez deskryptory

Jeżeli x i y oznaczają byty, $P_{(x)}$ i $P_{(y)}$ oznaczają pojęcia bytów x i y , $D_{P_{(x)}}$ i $D_{P_{(y)}}$ oznaczają deskryptory pojęć bytów x i y , to relację nadrzędności można wyrazić następująco: $P_{(x)}$ denotowane przez $D_{P_{(x)}}$ pozostaje w relacji nadrzędności do $P_{(y)}$ denotowanego przez $D_{P_{(y)}}$, wtedy i tylko wtedy, gdy istnieje przynajmniej jeden desygnat $D_{P_{(x)}}$, który nie jest desygnatem $D_{P_{(y)}}$ i każdy desygnat $D_{P_{(y)}}$ jest desygnatem $D_{P_{(x)}}$, np. pojęcie *roślina* pozostaje w relacji nadrzędności w stosunku do pojęcia *wierzba*, ponieważ każda *wierzba* jest *rośliną*, a tylko niektóre *rośliny* są *wierzbami*.

72 W teorii języków informacyjno-wyszukiwawczych stosuje terminy ekwiwalencja wyszukiwawcza, synonim wyszukiwawczy. Terminy te odnoszą się do relacji między sztucznymi jednostkami jiw, nie zaś do relacji między pojęciami bytów, które te jednostki denotują. Por. J. Woźniak, *Dobry język informacyjno-wyszukiwawczy, czyli o pożytkach nieulegania stereotypom*, „Archiwa, Biblioteki i Muzea Kościelne” 1999, t. 71, s. 35; B. Sosińska-Kalata, *Języki informacyjno-wyszukiwawcze: współczesne tendencje w badaniach i zastosowaniu*, w: *Społeczeństwo informacyjne i jego technologie: miscellanea informatologica Varsoviensia*, praca zbiorowa pod red. B. Sosińskiej-Kalaty, K. Materskiej, W. Glińskiego, Warszawa 2004, s. 7-8; PN-N-09018: *Tezaurus jednojęzyczny...*, s. 7.

Asocjacja pojęć bytów jako relacja logiczna⁷³

Jeżeli x i y oznaczają byty, $P_{(x)}$ i $P_{(y)}$ oznaczają pojęcia bytów x i y , $D_{P_{(x)}}$ i $D_{P_{(y)}}$ oznaczają deskryptory pojęć bytów x i y , $A_{DP_{(x)}}$ oznacza atrybut deskryptora $D_{P_{(x)}}$ ⁷⁴, to relację asocjacji można wyrazić następująco: między $P_{(x)}$ i $P_{(y)}$ zachodzi relacja asocjacji, wtedy i tylko wtedy, gdy istnieje taka wartość $A_{DP_{(x)}}$, która jest równa wartości $D_{P_{(y)}}$, np. pojęcie *lekarz* pozostaje z pojęciem *medycyna* w relacji asocjacji, ponieważ zakresem działalności *lekarza* jest *medycyna*. Z uwagi na dużą liczbę relacji i ich różną przydatność wyszukiwawczą wskazane jest oznaczanie tylko takich, które odpowiadają na zapotrzebowania informacyjne użytkowników konkretnego systemu. W teorii języków informacyjno-wyszukiwawczych problematyka relacji asocjacyjnych (zwanych też kojarzeniowymi lub skojarzeniowymi) była poruszana wielokrotnie. W większości przypadków stosunki te określane są jako relacje paradygmatyczne między jednostkami języka informacyjno-wyszukiwawczego⁷⁵. Zdaniem autora błąd formalny (*non sequitur, fallacia consequentis*) w rozumieniu tej relacji polega na tym, że nie istnieje ona w ogóle na poziomie jednostek leksykalnych języka informacyjno-wyszukiwawczego, lecz na poziomie bytów w związku z ich cechami akcydentalnymi bądź istotowymi. Określenie istoty relacji asocjacyjnych jest klasycznym przykładem implikacji ontologicznej typu redukcyjnego, czyli okoliczności, w której przedmiot badań nauki szczegółowej, w tym wypadku opartej na założeniach językoznawstwa strukturalistycznego, wykracza poza jej metody badawcze i implikuje pytania natury metafizycznej⁷⁶.

73 Relację można też rozpatrywać jako metafizyczną relację kategoryalną między bytami. Problem ten wykracza poza zakres artykułu i nie jest jeszcze szczegółowo rozpoznany. Szerzej o relacjach kategoryalnych zob. M. Gogacz, *Ważniejsze zagadnienia metafizyki*, Lublin 1973, s. 70 – http://katedra.uksw.edu.pl/gogacz/ksiazki/wazniejsze_zagadnienia_metafizyki.pdf [08.02.2015]; M. Gogacz, *Elementarz metafizyki*, Suwałki 1996, s. 45-49 – http://katedra.uksw.edu.pl/gogacz/ksiazki/elementarz_metafizyki.pdf [08.02.2015]; M. A. Krąpiec, *Język i świat realny...*, s. 285-298.

74 Desygnatem atrybutu deskryptora jest pojęcie cechy akcydentalnej lub istotowej bytu, którego pojęcie denotuje ten deskryptor. Zob. też *Atrybuty encji* na s. 184 tego artykułu.

75 Por. PN-N-09018: *Tezaurus jednojęzyczny...*, s. 2; *Słownik encyklopedyczny informacji...*, s. 227-228; B. Wereszczyńska-Cisło, *Relacje kojarzeniowe w językach informacyjnych*, „Zagadnienia Informacji Naukowej” 1985, t. 24 nr 1 (46), s. 69.

76 Termin „implikacje ontologiczne typu redukcyjnego” zostały wprowadzone do dyskursu naukowego przez K. Kłósaka, jako rozwinięcie „implikacji testowych hipotez

Deskryptory Biblioteki Narodowej – projekt uproszczenia JHP BN

Język Haseł Przedmiotowych Biblioteki Narodowej, który z założenia został zaprojektowany dla systemów manualnych, spełnia w tych systemach funkcję informacyjno-wyszukiawczą dzięki kilku cechom, które w systemach zautomatyzowanych opartych na encjach nie mają zastosowania. Pierwszą z tych cech jest narzędzie organizujące zbiór wyszukiawczy. W przypadku JHP BN jest nim gramatyka pozycyjna i abecedłowe szeregowanie w indeksie zdań języka informacyjno-wyszukiawczego z charakterystyk wyszukiawczych publikacji⁷⁷. Szeregowanie to ma charakter wielopoziomowy i odbywa się najpierw w obrębie jednostek autosyntaktycznych (tematów), a następnie w obrębie zwykle uogólniających jednostek synsyntaktycznych (określników), które modyfikują zakres pojęciowy jednostki nadrzędnej. Narzędzie to umożliwia wyszukiwanie manualne, zwane też popularnie *przełądaniem indeksów*, które jest podstawową metodą komunikacji użytkownika z katalogiem manualnym lub manualnym zmechanizowanym. Od wielu lat można obserwować stopniowe odchodzenie od tego sposobu wyszukiwania w komputerowych systemach informacyjnych. Świadczy o tym między innymi duża i stale rosnąca popularność systemów wyszukiwania informacji opartych na indeksowaniu współrzędnym, wykorzystujących w mechanizmie wyszukiwania związki encji kontrolowane słownikiem wyszukiawczym lub przez odpowiednio zaprojektowany algorytm. Zjawisko to występuje również w części interfejsów bibliograficznych baz danych o charakterze komercyjnym, gdzie albo wprowadzono wyłącznie wyszukiwanie pełnotekstowe z uwzględnieniem treści odpowiednio nasyconych słowami kluczowymi abstraktów, np. JSTOR⁷⁸, albo zastosowano wyszukiwanie pełnotekstowe i wyszukiwanie współrzędne za pomocą derywowanych z tekstu słów kluczowych, np. bazy ProQuest⁷⁹. Na uwagę zasługuje też rozwiązanie wdrożone w szwajcarskim katalogu RERO, agregującym opisy 220 bibliotek z obszaru frankofońskiej Romandii, gdzie zdecydowano się na deskryptorowe opracowanie zbiorów z wykorzystaniem słownictwa

naukowych” C. G. Hempła: K. Kłósak, *Z teorii i metodologii filozofii przyrody*, Poznań 1980, s. 150-151; por. C. G. Hempel, *Podstawy nauk przyrodniczych*, Warszawa 1968, s. 17-18.

77 Por. E. Chmielewska-Gorczyca, *Przed wyborem języka informacyjno-wyszukiawczego*, „Przegląd Informacyjno-Dokumentacyjny” 2000, t. 30, nr 1, s. 34-35.

78 *New to JSTOR? Learn more about us* – <http://about.jstor.org/10things> [08.02.2015].

79 ProQuest – www.proquest.com/ [08.02.2015].

Rameau⁸⁰. Spadek zainteresowania wyszukiwaniem manualnym jest widoczny również w statystykach wykorzystania indeksu przedmiotowego katalogu OPAC Biblioteki Narodowej, który do grudnia 2014 roku stanowił jedyne narzędzie wyszukiwania informacji o przedmiocie publikacji ze zbiorów BN⁸¹. O ile w roku 2009 średnie wykorzystanie tego indeksu stanowiło ok. 20-25% wszystkich zapytań⁸², to w latach 2013-2014 odsetek ten stanowił średnio 11,48% ogółu zapytań, a w ujęciach miesięcznych 6-13%⁸³.

Warto w tym momencie postawić dwa pytania. Pierwsze z nich dotyczy sposobu wyszukiwania informacji o zawartości publikacji i jej cechach formalnych w systemach komputerowych. Czy biblioteki powinny odstąpić od modelu silnie osadzonego w systemach manualnych na rzecz wyszukiwania zautomatyzowanego w systemie, którego podstawą jest współrzędne indeksowanie zawartości dokumentów i fasetowa prezentacja wyników wyszukiwania? Czy JHP BN i inne języki haseł przedmiotowych mogą być sprawnym wyszukiwawczo i efektywnym ekonomicznie narzędziem w takich właśnie systemach? Chcąc odpowiedzieć na te pytania, należy w pierwszej kolejności przeanalizować proces tworzenia instrukcji wyszukiwawczej JHP BN i innych języków informacyjno-wyszukiwawczych, a także sposób wykorzystania jej przez użytkowników katalogu. JHP BN cechuje ścisła prekoordynacja zarówno na poziomie autosyntaktycznych jednostek leksykalnych (tematów), jak i całych zdań (haseł przedmiotowych). Oznacza to, iż katalogujący ustala z góry instrukcję wyszukiwawczą, tzn. hasła denotujące bezpośrednio katalogowaną publikację⁸⁴, co skutkuje ograniczeniem możliwości wyszukiwaw-

80 P. Durussel, *RAMEAU dans le réseau suisse RERO. Le choix de la post-coordination*, 2014 – www.bnf.fr/documents/rameau2014_rero.pdf [08.02.2015]; *Directives pour l'indexation des matières dans RERO*, 2012 – www.rero.ch/pdfview.php?section=indexation&filename=directives_indexation_20120824.pdf [08.02.2015].

81 19 XII 2014 Biblioteka Narodowa uruchomiła nowe narzędzie przeszukiwania zbiorów Encore dostarczane przez firmę Innovative Interfaces Inc. – <http://www.bn.org.pl/aktualnosci/823-nowy-katalog-internetowy-bn.html> [08.02.2015].

82 W. Klenczon, *Indeksowanie zawartości treściowej...*, s. 5.

83 Statystyki wyszukiwań wg Indeksów w katalogu OPAC BN w okresie 15.01.2013-05.02.2014.

84 W teorii języków informacyjno-wyszukiwawczych stosuje się aksjomat w postaci pojęcia denotacji bezpośredniej i pośredniej wyrażań jiw, co jest następstwem apriorycznego podziału rzeczywistości na dokumentacyjną i pozadokumentacyjną. Zagadnienie to wykracza poza ramy tego artykułu i powinno być szczegółowo omówione w osobnej pracy uwzględniającej ontologiczne i logiczne implikacje tego problemu. Por. B. Sosińska, *Denotacja wyrażań języków informacyjno-wyszukiwawczych*, „Zagadnienia Informacji Naukowej” 1986, nr 1, s. 29-40; W. Babik, *Termin i jego status*

czych do haseł złożonych z tematów i modyfikujących ich znaczenie określników. W systemie manualnym użytkownik, który chce wyszukać dokument na określony temat, powinien znać tę instrukcję lub przynajmniej potrafić wyszukać w indeksie odpowiednie hasło denotujące grupę interesujących go publikacji (denotacja bezpośrednia jednostek języków informacyjno-wyszukiwawczych). Dysponuje przy tym narzędziem organizującym zbiór danych w postaci układu abecadłowego katalogu i gramatyki pozycyjnej. Efektem wyszukiwania jest wynik o maksymalnym współczynniku trafności, który dodatkowo może być zwiększany przez odpowiednie nasycenie katalogu odsyłaczami całkowitymi. Możliwe jest także uzyskanie wysokiego współczynnika kompletności przez zastosowanie rozbudowanej siatki odsyłaczy uzupełniających.

W systemie zautomatyzowanym jednostki syntaktyczne oraz synsyntaktyczne stanowią samodzielne punkty dostępu, co umożliwia znalezienie określonych przedmiotów i treści publikacji zarówno według instrukcji wyszukiwawczej zaproponowanej przez katalogującego, jak i według reguł indeksowania współrzędnego przez tworzenie dowolnych strategii, których nie przewidział autor instrukcji. Niewątpliwie rozszerza to możliwości wyszukiwania i wychodzi naprzeciw użytkownikom, którzy znają analogiczne mechanizmy z serwisów komercyjnych. Mogą w nich samodzielnie kształtować instrukcję wyszukiwawczą i dowolnie zawężać wyniki za pomocą faset. Rozważając możliwość wykorzystania w wyszukiwaniu fasetowym JHP BN, jak i innych narzędzi wyszukiwania informacji stosujących gramatykę pozycyjną, można stwierdzić, że mogłyby być one stosowane w systemach fasetowych jedynie w przypadku traktowania tematów i określników jako samodzielnych punktów dostępu. Wynika to z tego, że format MARC 21, w którym sporządzane są opisy, w sposób naturalny kategoryzuje poszczególne jednostki leksykalne w odpowiednich fasetach. Przy odpowiednio zaprojektowanej wyszukiwarce fasetowej tematy i określniki z tej samej kategorii informacji mogą być wyświetlane w jednej fasecie, np. tematy i określniki geograficzne. Nie wydaje się zasadne stosowanie jako punktów dostępu pełnych haseł rozwiniętych, ponieważ wyniki w fasetach sortowane są zwykle frekwencyjnie, a logicznym sposobem prezentacji haseł przedmiotowych jest indeks sortowany alfabetycznie, który jest czynnikiem skupiającym. Skutkiem prezentacji haseł rozwiniętych w fasecie byłoby niepożądane rozpraszenie informacji. Zakładając zastosowanie pierwszego scenariusza, należałoby

w systemie leksykalnym języka informacyjno-wyszukiwawczego, „Zagadnienia Informacji Naukowej” 1999, nr 1, s. 3-14; J. Woźniak, *Kategoryzacja: studium z teorii języków informacyjno-wyszukiwawczych*, Warszawa 2000, s. 31-32.

uwzględnić konieczne zmiany w metodyce JHP BN, ponieważ wiele trudności z osiągnięciem pożądanego efektu wyszukiwawczego wynika z aktualnych zasad opracowania przedmiotowego i reguł gramatycznych przeznaczonych do systemów manualnych. Problem ten dotyczy przede wszystkim JHP BN, w różnym stopniu też innych języków haseł przedmiotowych (KABA, LCSH, RAMEAU):

- uogólniające określniki w systemie fasetowym obniżają trafność wyników, co jest szczególnie widoczne w przypadku określnika geograficznego, który w JHP BN nie może stanowić jednostki mniejszej niż państwo. Uogólniające określniki geograficzne i wyszczególniające tematy geograficzne prezentowane w jednej fascie generują nieprawdziwą informację o liczbie publikacji na dany temat. Podobny problem pojawia się również w przypadku obszernych znaczeniowo określników ogólnych, np. historia, socjologia;
- uogólniający charakter określników generuje potrzebę stosowania opisu wielokrotnego. Wiąże się to z metodyką opracowania rzeczowego polegającą na stosowaniu w jednym opisie przedmiotowym tematów jednostkowych i właściwych im tematów ogólnych (stosowanie *haseł odwróconych*), co w systemie fasetowym może powodować zafałszowanie wyników wyszukiwania i często stanowi redundancję informacji zaprogramowanej w relacjach między hasłami na poziomie kartoteki;
- obecność dużej liczby haseł uogólniających, które często wskazują nie przedmiot, lecz nadmiernie uogólnioną klasę przedmiotów lub jego ujęcie, co zaciera przedmiotowy charakter katalogu. Szerzej o tym problemie pisała między innymi M. Błaziak w kontekście nowo wydanego słownika JHP BN autorstwa J. Trzcińskiej i E. Stepniakowej⁸⁵;
- głęboka prekoordynacja na poziomie jednostek leksykalnych, która przejawia się w doborze wyrażen złożonych semantycznie, łączących informacje z kilku faset, np. hasło *Powieść młodzieżowa angielska* łączy w jednym temacie informacje o formie, przeznaczeniu czytelniczym i cesze językowo-etnicznej, przez co zostaje ograniczona swoboda w kształtowaniu instrukcji wyszukiwawczej;
- niepełne nasycenie słownika JHP BN relacjami, które są istotnym elementem systemu wyszukiwania informacji.

85 M. Błaziak, *Hasło przedmiotowe a adekwatny opis przedmiotowy*, „Zagadnienia Informacji Naukowej” 1991, nr 1, s. 170-173; J. Trzcińska, E. Stepniakowa, J. Kossonoga, *Słownik języka haseł przedmiotowych Biblioteki Narodowej. Stan na dzień 30 czerwca 1986 r.*, Warszawa 1989.

Po analizie powyższej listy nasuwa się uzasadniona obawa, iż JHP BN w obecnym kształcie może wpłynąć negatywnie na trafność i kompletność wyników w wyszukiwaniu fasetowym. Obawa ta nie wynika z negatywnej oceny JHP BN, ale z funkcji, do jakiej był przeznaczony i jaką spełniał dobrze przez ponad 50 lat istnienia, mianowicie narzędzia manualnego katalogu przedmiotowego BN. Trzeba mieć też na względzie to, iż pewne rozwiązania JHP BN wynikały z ograniczeń katalogu kartkowego, w którym trzeba było pogodzić przedmiotowy i wyszczególniający charakter opracowania przedmiotowego z dbałością o nierozpraszenie informacji w zbyt szczegółowych hasłach i określnikach. Dodatkowo każda zmiana w obrębie hasła wiązała się z koniecznością melioracji, która w katalogu kartkowym jest czynnością czaso- i pracochłonną. Sytuacja ta zmienia się jednak w systemach zautomatyzowanych, w których te problemy nie występują i nie ma racjonalnych powodów, aby utrzymywać w nich rozwiązania przeznaczone do systemów manualnych. Odnosząc się do recepcji badań cranfieldzkich, z których wynika zbliżona efektywność wyszukiwawcza różnych języków informacyjno-wyszukiwawczych⁸⁶, warto rozważyć JHP BN z punktu widzenia nie tylko efektywności wyszukiwawczej, ale też i ekonomicznej. Zdaniem Jadwigi Woźniak to właśnie efektywność ekonomiczna języków informacyjno-wyszukiwawczych powinna być kluczowym argumentem wyboru narzędzia opracowania przedmiotowego, które dostosowane do możliwości i potrzeb konkretnej biblioteki stanowić będzie optymalne rozwiązanie⁸⁷. Rozpatrując aspekt ekonomiczny wykorzystania JHP BN w wyszukiwaniu fasetowym, warto ponownie zaznaczyć, że opiera się ono wyłącznie na indeksowaniu współrzędnym, co powoduje bezużyteczność struktur gramatyki pozycyjnej i instrukcji wyszukiwawczej sporządzonej przez katalogującego, ponieważ każda jednostka traktowana jest jako równorzędny punkt dostępu. Zakładając postulowany kierunek zmian w sposobie wyrażania tematów i określników JHP BN – co wydaje się trudne do osiągnięcia – o wiele lepszy efekt wyszukiwawczy można osiągnąć za pomocą deskryptorów indeksujących poszczególne kategorie informacji o zawartości, ale przy znacznie mniejszym nakładzie pracy. Zdaniem autora to właśnie tworzenie instrukcji wyszukiwawczej, a nie poprawna analiza zawartości dokumentu jest procesem najbardziej pracochłonnym i sprawiającym najwięcej trudności katalogującym⁸⁸. Wynika to

86 J. Woźniak, *Dobry język informacyjno-wyszukiwawczy...*, s. 32.

87 *Ibidem*, s. 32.

88 T. Głowacka, *Analiza dokumentu i jego opis przedmiotowy*, Warszawa 2003, s. 9-10; D. Domalewska, *Rzeczowe opracowanie zbiorów w komputerowych katalogach ekonomicznych – aspekt informacyjny*, „Praktyka i Teoria Informacji Naukowej i Technicznej” 2006, t. 14, s. 26.

między innymi z tego, iż proces przekładu informacji na hasła JHP BN wymaga redukcji tej informacji zgodnie z zasadami opracowania JHP BN. Sprowadza się to do subiektywnej oceny tej informacji według wytycznych poradnika JHP BN, opartego głównie na regułach pragmatycznych w postaci schematów opracowania poszczególnych rodzajów piśmiennictwa i grup zagadnień⁸⁹. W przypadku opracowania deskryptorowego katalogujący nie tworzy instrukcji wyszukiwawczej, lecz skupia swoją uwagę na indeksowaniu istotnych cech publikacji, a sama instrukcja wyszukiwawcza jest tworzona przez użytkownika w trakcie wyszukiwania w systemie fasetowym. Następstwem takiego sposobu opracowania jest możliwość odstąpienia od stosowania haseł rozwiniętych, które już teraz w grupie haseł 150, 151, 155 stanowią 84% wszystkich rekordów wzorcowych. Od kilku lat obserwujemy także postępujący przyrost rekordów wzorcowych dla haseł rozwiniętych⁹⁰. W 2014 roku do bazy Sierra wprowadzono 45 428 rekordów wzorcowych dla haseł rozwiniętych i 45 757 rekordów dla haseł nierozwiniętych różnych typów⁹¹. Wśród wszystkich rekordów dla haseł nierozwiniętych aż 92% stanowią hasła z grupy 100, 110, 111, 130 i 100a/t, a tylko 4% to słownikowe hasła nierozwinięte. W ujęciach miesięcznych przyrost rekordów wzorcowych poszczególnych typów kształtował się następująco:

Tabela 1

Przyrost rekordów wzorcowych dla haseł przedmiotowych⁹²

Miesiąc	Hasła rozwinięte	Hasła nierozwinięte	Hasła słownikowe
1	2	3	4
Styczeń	48% (2113)	52% (2272)	11% (466)
Luty	55% (1939)	45% (1627)	4% (155)
Marzec	48% (2016)	52% (2160)	3% (145)
Kwiecień ⁹³	55% (4587)	45% (3728)	2% (187)
Maj	53% (4186)	47% (3778)	3% (273)
Czerwiec	54% (3827)	46% (3231)	2% (160)

89 *Język Haseł Przedmiotowych Biblioteki Narodowej: poradnik*, Warszawa 2012 – <http://fabryka-jezyka.bn.org.pl/?p=1255> [08.02.2015].

90 Zob. „Sprawozdania Biblioteki Narodowej” za lata 2007-2013 – www.bn.org.pl/o-bn/sprawozdanie-bn [08.02.2015].

91 Dane z systemu bibliotecznego Sierra wg stanu na dzień 1 II 2015.

92 Dane z systemu bibliotecznego Sierra za okres od 1 I do 30 VI wg stanu na 30 VI 2014, za okres od 1 VII do 31 XII 2014 wg stanu na 31 XVII 2014.

93 Przyrost nowych rekordów wzorcowych w kwietniu związany jest z włączeniem do systemu Sierra Bibliografii Zawartości Czasopism.

1	2	3	4
Lipiec	55% (4480)	45% (3725)	3% (249)
Sierpień	51% (3295)	49% (3136)	4% (228)
Wrzesień	50% (3959)	50% (4024)	2% (194)
Październik ⁹⁴	47% (4385)	53% (4934)	2% (200)
Listopad	46% (4741)	54% (5532)	1% (112)
Grudzień	44% (5946)	56% (7617)	0,5% (74)

Z powyższych statystyk wynika, że nakład pracy włożony w tworzenie i kontrolę poprawności rekordów dla haseł rozwiniętych jest niewspółmierny z korzyściami dla użytkowników, którzy w bardzo małym stopniu wykorzystują przeglądanie indeksu haseł przedmiotowych. Alarmująca jest też rekordowo niska liczba nowo utworzonych rekordów dla haseł słownikowych w stosunku do ogółu utworzonych rekordów wzorcowych dla haseł JHP BN, których ilość w grudniu 2014 r. pierwszy raz w historii spadła poniżej 1%. Trudno jest także uzasadnić z ekonomicznego punktu widzenia tworzenie struktur, które nie wpływają ani negatywnie, ani pozytywnie na efektywność wyszukiwawczą w systemie fasetowym.

Dyskusja na temat uproszczenia opracowania przedmiotowego i dostosowanie go do współczesnych oczekiwań użytkowników trwa w Bibliotece Narodowej od kilku lat. Jej pierwszym efektem była decyzja z 2009 roku o rezygnacji ze stosowania określników formalnych na rzecz tematów formalnych i jej realizacja w latach 2010-2011⁹⁵. Problematyka ta została szerzej omówiona przez Annę Stolarczyk i Wandę Klenczon w trakcie XII i XIII Ogólnopolskich Warsztatów JHP BN i UKD, które odbyły się w latach 2010-2011⁹⁶. Wśród postulowanych zmian w metodyce JHP BN wskazywano

94 Przyrost nowych rekordów od października 2014 związany jest z rozpoczęciem katalogowania zawartości naukowych prac zbiorowych – www.bn.org.pl/aktualnosci/786-artykuly-i-rozdzialy-publicacji-zbiorowych-uwzglednione-w-katalogu-bn.html [08.02.2015].

95 Decyzja podjęta przez Zespół ds. Rozwoju Języka Haseł Przedmiotowych BN, *Zmiany w Języku Haseł Przedmiotowych Biblioteki Narodowej* – www.bn.org.pl/download/document/1266914664.doc [18.02.2015]; *Język haseł przedmiotowych Biblioteki Narodowej (JHP BN)*, „Sprawozdanie Biblioteki Narodowej za rok 2010”, s. 31-32 – www.bn.org.pl/download/document/1307358487.pdf [08.02.2015].

96 W. Klenczon, A. Stolarczyk, *Zmiany w słownictwie i metodyce JHP BN – założenia, realizacja i perspektywy* – <http://www.bn.org.pl/download/document/1278510106.ppt>; W. Klenczon, A. Stolarczyk, *JHP BN stan prac nad słownictwem i metodyką opracowania* – www.bn.org.pl/download/document/1308143417.pdf [08.02.2015].

wówczas potrzebę uporządkowania słownictwa i uproszczenia gramatyki pozycyjnej, w tym ograniczenie stosowania określników chronologicznych. Odrębnym zagadnieniem był też problem pre- i postkoordynacji jednostek jiw w wyszukiwaniu informacji oraz zestawienie wad i zalet obu modeli wyszukiwania. Do głównych zalet modelu postkoordynacyjnego zaliczono: łatwość posługiwania się zarówno przez użytkowników, jak i katalogujących, krótsze i bardziej efektywne szkolenie personelu oraz szybsze i tańsze opracowanie. Z drugiej strony prelegentki zwróciły uwagę, że postkoordynacyjny model tworzenia charakterystyk przedmiotowych umożliwia wyszukiwanie tylko w opcji *searching* i generuje grupę wyników nierelwantnych, powstałych głównie w wyniku *fałszywej koordynacji*. Ze wstępnych analiz wykonanych w 2014 r. przez Pracownię JHP BN na reprezentatywnej grupie publikacji ogłoszonych w „Przewodniku Bibliograficznym” w latach 2013-2014 wynika, iż problem fałszywej koordynacji może dotyczyć 6% tematów i określników ogólnych oraz 2% tematów i określników geograficznych z charakterystyk przedmiotowych publikacji⁹⁷. Zagadnienie to wymaga osobnego opracowania, w którym zostałyby przeanalizowane poszczególne rodzaje publikacji (publikacje naukowe z poszczególnych dziedzin, poradniki, beletrystyka itp.). W ogólnej ocenie następstw fałszywej koordynacji należałoby także zwrócić uwagę na problem relewancji i pertynencji wyników wyszukiwania oraz przeanalizować ich znaczenie dla pozyskiwania przez użytkownika informacji potencjalnie inspirującej. Zdaniem Remigiusza Sapy współczesne katalogi OPAC w niewielkim stopniu umożliwiają dostęp do takiej właśnie informacji i nastawione są na uzyskanie maksymalnej liczby wyników zgodnych z instrukcją wyszukiwawczą. Informacja o publikacjach, które nie spełniają tego warunku, pojawia się jedynie w przypadku braku wyników na zapytanie. Zdaniem autora dążenie do efektywności i precyzji wyszukiwania w katalogach bibliotecznych nie powinno przesłaniać innych aspektów pozyskiwania informacji⁹⁸. W przypadku języków prekoordynowanych, w tym JHP BN, których zaletami są precyzyjność i niemal 100% trafność w wyrażaniu informacji o przedmiocie dokumentu, w trakcie warsztatów zwrócono także uwagę na ich poważne wady, między innymi hasła trudne dla katalogujących i niezrozumiałe dla użytkowników, konieczność długie-

97 K. Trzaska – badania własne przeprowadzone w marcu 2014 r. na zlecenie Pracowni Języka Haseł Przedmiotowych BN.

98 R. Sapa, *Potencjał bibliotecznych katalogów online w zakresie sprzyjania zjawisku przypadkowego pozyskiwania informacji o literaturze naukowej*, „Przegląd Biblioteczny” 2008, t. 76, z. 1, s. 103-104.

go i ciągłego szkolenia personelu, pracochłonne i kosztowne opracowanie⁹⁹. Biorąc pod uwagę stale zmniejszający się procent wyszukiwania przedmiotowego w katalogach, trudności w dostosowaniu charakterystyk przedmiotowych do aktualnych standardów metadanych i kwestie ekonomiki pracy, Biblioteka Narodowa rozpoczęła prace nad projektem uproszczenia JHP BN i przekształcenia go w Deskryptory Biblioteki Narodowej. W czerwcu 2014 r. została powołana Pracownia Deskryptorów Biblioteki Narodowej, która koordynuje prace metodyczne i dyskusje nad kształtem opracowania formalnego i przedmiotowego z wykorzystaniem deskryptorów.

Indeksowanie encji opisu bibliograficznego

Zakres zmian w sposobie opisu bibliograficznego wynikający z encyjnego podejścia do informacji bibliograficznej można ująć w trzy grupy zagadnień.

1. Wydzielenie encji opisu bibliograficznego i ich atrybutów uwzględniających w szerszym niż dotychczas stopniu model FRBR¹⁰⁰, które w sposób jednoznaczny rozgraniczają dotychczasowe elementy opisu formalnego i rzeczowego. Wiąże się to z koniecznością zwiększenia liczby kontrolowanych punktów dostępu, szczególnie dla tych elementów opisu formalnego, które zawierają informację zapisaną analogowo.
2. Utworzenie jednej bazy deskryptorów, kontrolującej wszystkie encje opisu bibliograficznego, ich atrybuty i relacje pomiędzy instancjami encji. Związane jest to między innymi z ujednoczeniem i integracją kartotek wzorcowych haseł formalnych i kartoteki Języka Haseł Przedmiotowych Biblioteki Narodowej, co zostanie szerzej omówione w trzecim rozdziale artykułu.
3. Przeniesienie większości atrybutów encji z poziomu rekordu bibliograficznego do rekordu wzorcowego właściwych instancji.

Zgodnie z tymi założeniami przyjęto następujący układ encji opisu bibliograficznego, które w zależności od potrzeb powinny być uszczegóławiane lub uogólniane na poszczególnych poziomach.

Encje – cechy formalne publikacji

Encje: oznaczenie odpowiedzialności

Encja przyjmuje wartość w postaci nazwy bytu, który stworzył, współtworzył lub brał udział w powstaniu, realizacji lub opublikowaniu dzieła (autor, tłumacz, instytucja, przy której afiliowany jest autor, wydawca itp.). Z uwagi

⁹⁹ W. Klenczon, A. Stolarczyk, *JHP BN stan prac nad słownictwem i metodyką opracowania...*

¹⁰⁰ *Functional Requirements for Bibliographic Records. Final Report.*

na zróżnicowanie funkcji jakie mogą pełnić poszczególne byty w powstaniu publikacji, wskazane jest uszczegółowienie encji przez określenie rodzaju współpracy na poziomie rekordu bibliograficznego.

Encja: oznaczenie dzieła

Z uwagi na wysoki poziom abstrakcji eferberowskiej definicji dzieła, o którym wspomniano w części *Model FRBR jako model związków encji* (s. 157), wskazane jest przyjęcie utylitarnej interpretacji, w której dane o poszczególnych atrybutach dzieła są przejmowane z pierwszego oryginalnego wydania lub innych źródeł, jeżeli pierwsze oryginalne wydanie jest niedostępne. Z tych samych względów wydaje się zasadne indeksować tylko część atrybutów dzieła i traktować jako encje opisu bibliograficznego: tytuł dzieła / identyfikator, formę dzieła, czas powstania dzieła, funkcję / przeznaczenie czytelnicze, obsadę wykonawczą itp. Pełne wdrożenie encji „oznaczenie dzieła” jest też uzależnione od globalnego modelu metadanych dla identyfikacji dzieła i jego atrybutów, które dopiero się krystalizuje¹⁰¹. Przyjęcie na tym etapie rozwiązań *explicite* z modelu FRBR, odmiennych niż w większości bibliotek na świecie, niesie ze sobą ryzyko niekompatybilności podczas wymiany danych.

Encja: cechy fizyczne publikacji

Encja przyjmuje wartość pojęcia charakteryzującego fizyczną postać publikacji (typ publikacji, datę publikacji, liczbę stron, technikę wykonania, typ nośnika, formę fizyczną, czas trwania itp.), a także cechy związane ze sposobem zapisu i odczytu informacji zawartej w publikacji (typ zawartości, typ medium, format zapisu). Instancje poszczególnych encji przejmowane będą z bazy deskryptorów. Mogą także przyjmować wartość binarną lub liczbową.

Encje – cechy zawartości publikacji

Encja: przedmiot publikacji

Przedmiot jest główną cechą zawartości publikacji i wyraża element rzeczywistości (byt) w niej przedstawiony. Element ten istnieje realnie lub potencjalnie, niezależnie od publikacji, w której został przedstawiony¹⁰². Encja przy-

¹⁰¹ Zob. m.in. eksperymentalny projekt OCLC, w którym klastry dla poszczególnych dzieł w większości wypadków są na etapie tworzenia; *OCLC WorldCat Work Entity* – www.oclc.org/developer/develop/linked-data/worldcat-entities/worldcat-work-entity.en.html [08.02.2015].

¹⁰² Zob. A. Łysakowski, *Katalog przedmiotowy*, cz. 1, *Teorja*, Wilno 1928, s. 41-42, por. wyd. 2, wstęp J. Woźniak, Warszawa 2002 – www.bbc.uw.edu.pl/publication/650 [08.02.2015].

muje zatem wartość w postaci wyszczególnionej nazwy (znaku) pojęcia bytu przedstawionego w publikacji, niezależnie od ujęcia, w jakim został przedstawiony. Przez wyszczególnienie rozumieć należy zindeksowanie elementu na takim samym poziomie szczegółowości jak w opracowywanej publikacji.

Encja: klasa treściowa

Encja ta jest drugą co do ważności cechą zawartości publikacji i jest ściśle związana z jej przedmiotem. Klasa treściowa jest zredukowaną treścią przedstawienia przedmiotu¹⁰³, a więc jego immanentnym ujęciem, np. dziedziną wiedzy, z punktu widzenia której przedstawiono przedmiot¹⁰⁴. Warto tu zauważyć, iż jeden przedmiot może być przedstawiony w wielu ujęciach, a jedno ujęcie może dotyczyć wielu przedmiotów. Z tego powodu zarówno uogólniona treść, jak i wyszczególniony przedmiot powinny być od siebie ściśle oddzielone w osobnych encjach i uzupełniać się wzajemnie w trakcie wyszukiwania. W przeciwieństwie do przedmiotu publikacji, nie jest zasadne wyszczególnianie treści, lecz tworzenie klas treściowych, które przejmują część funkcji określników JHP BN, ale stanowią samodzielny punkt wyszukiwawczy. Praktyka taka uwzględnia wszystkie, nawet najbardziej zróżnicowane potrzeby wyszukiwawcze użytkowników. Umożliwia wyszukiwanie publikacji na temat dowolnego przedmiotu bez względu na treść, publikacji o dowolnej treści bez względu na przedmiot oraz publikacji o określonej treści i przedmiocie.

Encja: zakres chronologiczny

Encja wskazuje lokalizację w czasie przedmiotu lub treści publikacji. W wyszukiwaniu współrzędnym lokalizacja ta stanowi autonomiczny punkt dostępu. Z uwagi na wielość zakresów chronologicznych, jakie mogą wystąpić w publikacji, a także ich różną przydatność wyszukiwawczą, zasadne jest uwzględniać jedynie główne zakresy i redukować je do ustalonych przedziałów lat.

Encja: zakres geograficzny

Encja wskazuje lokalizację w przestrzeni przedmiotu lub treści publikacji. W wyszukiwaniu współrzędnym lokalizacja ta stanowi autonomiczny punkt dostępu. Istnieją również przypadki, gdy zakres geograficzny pokrywa się z przedmiotem publikacji, np. historia miejscowości. Z tego względu wskazane

103 Zob. *ibidem*, s. 48-49; zob. też K. Twardowski, *O treści i przedmiocie przedstawień*, w: K. Twardowski, *Wybrane pisma filozoficzne*, Warszawa 1965, s. 3-8.

104 Rozdzielenie przedmiotu od treści jest jednym z podstawowych założeń teoretycznych katalogu przedmiotowego BN; zob. A. Łysakowski, *Katalog przedmiotowy...*, s. 43-46.

jest, aby zakres geograficzny wyrażać na takim samym stopniu szczegółowości jak w opracowywanej publikacji, a poszczególne jednostki geograficzne wiązać relacjami podrzędności, nadrzędności i asocjacyjnymi na poziomie instancji w kartotece deskryptorów.

Atrybuty encji

Atrybuty encji¹⁰⁵, które wyrażają wyłącznie istotę lub przypadłości (cechy akcydentalne) bytu, oznaczane są na poziomie instancji w kartotece deskryptorów, np. *narodowość, zawód, siedziba ciała zbiorowego, obszar działalności*, itp. Wartości atrybutów encji przejmowane są z bazy deskryptorów i tworzą sieć powiązań między encjami o tych samych atrybutach, np. grupa deskryptorów osób o atrybucie *lekarze*, grupa deskryptorów instytucji o atrybucie *działalność charytatywna*, itp.

Indeksowanie Deskryptorami Biblioteki Narodowej

Proces indeksowania Deskryptorami Biblioteki Narodowej sprowadza się do nadawania wartości poszczególnym encjom opisu bibliograficznego, czyli wyrażania ich instancji. Instancje encji przejmowane są z bazy deskryptorów, a całość procesu indeksowania można wyrazić za pomocą poniższego schematu.

Przy założeniu, że istnieje jedna rzeczywistość r , a publikacja przedstawia pewne jej elementy $E_{(r)}$; $X_{(K)}$ stanowi dowolną encję opisu bibliograficznego publikacji K ; $Y_{(r)}$ stanowi deskryptor dowolnego elementu rzeczywistości; proces indeksowania za pomocą Deskryptorów BN można zdefiniować jako akt dwóch idiogenicznych¹⁰⁶ i prawdziwych sądów pozostających ze sobą w stosunku implikacji:

Dla każdej publikacji K przedstawiającej elementy rzeczywistości $E_{(r)}$ zachodzi implikacja P : Jeżeli publikacja K nie posiada encji $X_{(K)}$ o wartości pustej, to deskryptor $Y_{(r)}$ denotuje $E_{(r)}$, który stanowi instancję encji $X_{(K)}$ dla publikacji K .

Przykład:

Publikacja: Ogłoszenie do Mieszkańców m.st. Warszawy [Inc.:] Ospa jest chorobą b. ciężką, często sprowadza śmierć, a prawie zawsze upośledzenie zdrowia i oszpe-

105 Należy odróżnić atrybut encji od atrybutu bytu w ujęciu metafizycznym.

106 K. Twardowski, *O idio- i allogetycznych teoriach sądu*, w: K. Twardowski, *Rozprawy i artykuły filozoficzne*, zebrali i wydali uczniowie, Lwów 1927, s. 418-420 – <http://ebuw.uw.edu.pl/publication/104686> [08.02.2015]; por. D. Tennerówna, *Istnienie jako „treść” sądu i sądu. Kilka uwag na marginesie teorii sądu prof. Twardowskiego*, „Przełęcz Filozoficzny” 1914, t. 17, nr 4, s. 465 – www.wbc.poznan.pl/dlibra/publication/106805 [08.02.2015].

cenie ciała [...] : Warszawa, w kwietniu 1938 r. / Zarząd Miejski w m. st. Warszawie. Wydział Opieki Społecznej i Zdrowia Publicznego. – Warszawa, 1938.

I. Encje nieposiadające wartości pustej:

przedmiot publikacji, zakres geograficzny, zakres chronologiczny, klasa treściowa, forma dzieła.

II. Poszczególne elementy rzeczywistości wyrażone Deskryptorami BN stanowią instancje encji w opisie bibliograficznym publikacji.

Deskryptor *Ospa (choroba)* denotuje encję *przedmiot publikacji*.

Deskryptor *Warszawa (województwo mazowieckie)* denotuje encję *zakres geograficzny*.

Deskryptory *1901-2000, 1918-1939* denotują encję *zakres chronologiczny*.

Deskryptor *Epidemiologia* denotuje encję *klasa treściowa*.

Deskryptor *Afisz* denotuje encję *forma dzieła*.

Deskryptor *Warszawa. Zarząd Miejski. Wydział Opieki Społecznej i Zdrowia Publicznego* denotuje encję *autor*¹⁰⁸.

3. Zintegrowane kartoteki wzorcowe

Zagadnienia związane z modelowaniem baz danych i ich funkcji w globalnych systemach wymiany informacji skłaniają środowisko bibliotekarzy do podejmowania kolejnych wyzwań związanych z ułatwieniem wyszukiwania informacji przez użytkowników. Od wieków to właśnie w bibliotekach tworzone różne formy wykazów bibliograficznych, które pomagały w wyszukiwaniu dokumentów na dany temat. Wpływ na uporządkowanie i usystematyzowanie wiedzy w bibliotece w dużej mierze mają nadal działania podejmowane przez bibliotekarzy. Jednym z większych wyzwań w obecnym czasie jest takie zorganizowanie własnych zasobów, aby zapewnić przystępny i szybki dostęp do informacji. Współczesne standardy sieciowe umożliwiają zorganizowanie powiązań danych w taki sposób, że niezależnie od miejsca przechowywania danych o publikacji, można stosunkowo prosto i łatwo uzyskać informację o jej cechach i lokalizacji poszczególnych egzemplarzy¹⁰⁷. Aby stało się to normą, należy przyjąć taką koncepcję, która zagwarantuje „standaryzację publikowanych danych pozwalających na wielokrotne ich wykorzystanie, możliwość

107 Szczegółowość klas treściowych jest uzależniona od konwencji przyjętej przez instytucję sporządzającą opis.

108 A. S. Chandel, R. Vijay Prasad, *Journey of catalogue from Panizzi's principles to Resources Description and Access*, „Journal of Library & Information Technology” 2013, t. 33, nr 4, s. 314-322 – <http://publications.drdo.gov.in/ojs/index.php/djlit/article/view/File/5007/2913> [08.02.2015].

odwoływania się do istniejących kolekcji oraz integrację i współdziałanie”¹⁰⁹ wielu podmiotów, które specjalizują się w przechowywaniu i przetwarzaniu danych.

W związku z tym podejmowane są działania mające na celu dostosowanie struktury bibliograficznych baz danych do modelu ER¹¹⁰. Związane to jest z optymalizowaniem rozwiązań technologicznych i wymogami, jakie przed światem bibliotekarskim stawiają użytkownicy, którzy na wzór poszukiwań w Google oczekują prostych i intuicyjnych narzędzi wyszukiwania informacji. Systemy biblioteczne powinny stwarzać takie możliwości wyszukiwania, do jakich przyzwyczajony jest każdy użytkownik Internetu. Realizacja tych założeń w systemach komputerowych nie jest możliwa bez ścisłego rozdzielenia poszczególnych elementów opisu bibliograficznego – wydzielenia encji opisu bibliograficznego. Obserwacja działań poszczególnych bibliotek na świecie i kierunków rozwoju wytycznych IFLA¹¹¹ skłania do stwierdzenia, że istnieje uzasadniona potrzeba tworzenia bibliograficznych baz danych zgodnych ze współczesnymi standardami sieciowymi. Zagadnienie to zostało szczegółowo omówione w pierwszej części artykułu. W tej części zostaną poruszone problemy związane z praktycznym wymiarem ujednoczenia i scalania kartotek wzorcowych w bazach bibliotecznych.

Proces integracji kartotek wymaga zarówno dużej znajomości problemów związanych z metodyką tworzenia haseł, jak i wiedzy dotyczącej zintegrowanych systemów bibliotecznych oraz ich możliwości technicznych. Library of Congress, w zależności od funkcji, jakie pełni dane hasło w opisie bibliograficznym, wprowadza je do kartoteki haseł formalnych lub przedmiotowych. Hasła osobowe, korporatywne i tytułowe, które mogą być stosowane zarówno w opisie formalnym, jak i przedmiotowym, przejmowane są z kartoteki haseł formalnych. Ten przydział decyduje o przyporządkowaniu numeru kontrolnego z odpowiedniej kartoteki¹¹². Dla użytkownika jest to o tyle ważne, że niezależnie, w którym indeksie będzie poszukiwał rekordu wzorcowego dla *Harvard College (1636-1780)* – czy to przez indeks haseł przedmiotowych (*Subject Authority Headings*), czy

109 M. Roszkowski, *Od MARC 21 do Semantic Web: reprezentacja metadanych bibliograficznych w środowisku sieciowym*, w: *Bibliografi@: źródła, standardy, zasoby*, pod red. J. Franke, Warszawa 2013, s. 27.

110 Model ER został omówiony w rozdziale *Bazy danych bibliograficznych a relacyjne bazy danych*, s. 152-157.

111 *National Bibliographies in the Digital Age: Guidance and New Directions*, München 2009; *Guidelines for Subject Access in National Bibliographies...*

112 M. Nasiłowska, *Typologia haseł wzorcowych stosowanych w jhp KABA* – www.ebib.pl/publikacje/matkonf/kaba/nasilowska.php [08.02.2015].

indeks hasel formalnych (*Names Authority Headings*), zostanie odesłany do tego samego rekordu o numerze kontrolnym 85268664. Podobnymi zasadami kieruje się Bibliothèque nationale de France. Przy wyszukiwaniu hasła *Universidad de Salamanca* z poziomu *autorités BnF* oraz RAMEAU, użytkownik znajdzie ten sam rekord *Universidad de Salamanca (Espagne) – Capilla* (FRBNF16769473). Na podstawie powyższych przykładów widać, że ujednoczenie form hasel przedmiotowych i formalnych dla tych samych bytów jest koniecznością, ponieważ zapobiega to rozpraszaniu informacji o tym samym przedmiocie poszukiwań. Większość bibliotek opracowujących zbiory zgodnie z wytycznymi *Functional Requirements for Authority Data (FRAD)* oraz *Functional Requirements for Subject Authority Data (FRSAD)* dostosowała do nich swoje kartoteki, czasami pozostawiając jednak pewien margines rozbieżności, ponieważ uwarunkowania, jakimi kierowano się przy tworzeniu tych hasel, mają swoją historię i czasami trudno jednoznacznie podjąć takie decyzje na korzyść jednej opcji, np. w *autorités BnF* nadal nie rozwiązano problemu ulokowania hasel dla nazw geograficznych i tytułów czasopism¹¹³.

Biblioteka Narodowa, decydując się na przekształcenie JHP BN w Deskryptory BN, które są bardziej przyjazne i zrozumiałe dla współczesnego użytkownika, powinna wykonać takie działania, które umożliwią stosowanie deskryptorów jako elementów wspólnej kartoteki obsługującej wszystkie encje opisu bibliograficznego. W pierwszej kolejności należy więc ujednoczyć formy i zasady doboru hasel w kartotece wzorcowych hasel formalnych (kwhf) oraz w kartotece wzorcowych hasel przedmiotowych (kwhp). To zadanie jest niezbędne, aby przejść do etapu następnego, czyli fizycznej integracji obu kartotek. Przed przekształceniem JHP BN na Deskryptory BN należy dokonać analizy istniejących hasel i zasad ich doboru. Niezależnie od tego, czy wyrażane są za pomocą JHP BN, czy Deskryptorów BN, ich budowa powinna być identyczna w stosunku do hasel znajdujących się w kartotece kwhf. W związku z zaleceniami IFLA w Bibliotece Narodowej od kilku lat następuje proces ujednoczania hasel i w zasadzie pozostał niewielki margines zagadnień do uporządkowania. Niewątpliwie korzystne dla użytkowników poszukujących informacji na temat określonej osoby, instytucji itp. byłoby, gdyby nie napotykali oni w bazie BN na rozbieżności wynikające z braku konsekwencji w doborze form dla hasel jednostkowych.

Ujednoczenie hasel oraz fizyczna integracja obu kartotek zmienia całkowicie sposób podejścia do uzyskania informacji dotyczącej osób, instytucji, przedmiotu publikacji itp. Współczesne systemy informacyjne pozwalają na wyszukiwanie uwzględniające atrybuty deskryptora oraz funkcje, które

113 *Ibidem*.

poszukiwany deskryptor pełni w opisie bibliograficznym, np. *Lanckorońska, Karolina (1898-2002)* – przedmiot lub autor publikacji; *Lanckorońska, Karolina (1898-2002)* jako wynik zapytania o osoby, których zakresem działalności była historia sztuki. Informacje te powinny być zdefiniowane w taki sposób, aby umożliwić systemom komputerowym zinterpretowanie tych danych w sposób jednoznaczny.

Przykład 1

*Rekord wzorcowy dla hasła osobowego*¹¹⁴

024 7_ |a <http://viaf.org/viaf/76324812> |2 uri
 024 7_ |a 0000000079780281 |2 isni
 043 __ |c pl
 046 __ |f 1898 |g 2002
 100 1_ |a Lanckorońska, Karolina |d (1898-2002)
 368 __ |d Doktorzy honoris causa |s 1990
 370 __ |a Austria |b Rzym (Włochy)
 372 __ |a Historia sztuki
 372 __ |a Służby specjalne
 372 __ |a Nauka
 372 __ |a Mecenat
 373 __ |a Armia Krajowa
 373 __ |a Polskie Towarzystwo Naukowe na Obczyźnie |s1948
 373 __ |a Fondation Lanckoronski |s 1967
 373 __ |a The De Brzezie Lanckoronski Foundation
 374 __ |a Historycy sztuki
 375 __ |a 2 |2 iso5218
 377 __ |a ita |a ger |a pol
 400 1_ |a Lanckorońska, Carolina
 400 1_ |a Lanckorońska, Charlotte
 510 2_ |a Fondation Lanckoronski
 510 2_ |a The De Brzezie Lanckoronski Foundation

Uzupełnienie rekordów wzorcowych o dodatkowe atrybuty związane z osobą, instytucją itp. znacząco zwiększa możliwości wyszukiwawcze. Informacje zapisane w grupie pól 3XX stanowią dodatkowe punkty dostępu.

114 Przykład poglądowy, w którym zastosowano między innymi nowe pola z grupy 3XX.

Hasło osobowe – tematy osobowe

Zarówno w hasłach osobowych formalnych, jak i przedmiotowych powinny obowiązywać te same zasady redakcji i doboru haseł, lecz w praktyce występują pewne istotne różnice związane z odmienną tradycją katalogu przedmiotowego. Zapis hasła osobowego powinien być zgodny z przepisami, jakie stosuje bibliografia narodowa i przy opracowaniu należy kierować się normą PN-N-01229¹¹⁵ oraz wynikającymi z tego konsekwencjami, czyli stosowaniem odpowiednich dopowiedzeń. Jadwiga Sadowska¹¹⁶ przeprowadziła w latach 90. analizę ewolucji zapisu tematów osobowych i ich zbliżenie się do hasła osobowego formalnego na przestrzeni kilkudziesięciu lat. Poniższy przykład pokazuje, w jaki sposób zmieniały się koncepcje zapisu dopowiedzeń.

Przykład 2

- Franciszek z Asyżu św. – do połowy lat 80.
- Franciszek z Asyżu, św. – od połowy 80. do połowy 90.
- Franciszek z Asyżu (św.) – 1996-1997
- Franciszek z Asyżu (św. ; 1181?-1226) – od 1998¹¹⁷

W momencie przyjęcia normy w 1998 roku „Przewodnik Bibliograficzny” wprowadził zmiany w zapisie hasła osobowego. Dotyczyło to między innymi zastosowania przecinka, który oddziela nazwisko od imienia, oraz dodania dat życia.

Przykład 3

- Konopnicka Maria
- Konopnicka, Maria (1841-1910)¹¹⁸

Norma, według której zaczęto redagowanie haseł w BN, była zgodna z przepisami międzynarodowymi i światową tendencją w zakresie normalizacji bibliograficznej. W tym czasie zarówno Library of Congress, jak i Bibliothèque nationale de France podjęły pracę nad ujednoczeniem swoich kartotek oraz tworzeniem jednego rekordu wspólnego dla kartoteki formalnych haseł osobowych i kartoteki haseł przedmiotowych. Dotychczasowe zasady redakcji i wyboru haseł według JHP BN przewidują, że jedna osoba może wystąpić tylko

115 PN-N-01229: 1998. *Hasło opisu bibliograficznego. Hasło osobowe.*

116 J. Sadowska, *Hasła przedmiotowe w teorii Adama Łysakowskiego i praktyce „Przewodnika Bibliograficznego”*, Warszawa 2000.

117 *Ibidem*, s. 113.

118 *Ibidem*.

pod jedną nazwą, a wszystkie pozostałe są traktowane jako formy odrzucone. Norma PN-N-01229 dopuszcza jednak użycie kilku form hasła wyrażającego tę samą osobę i skojarzenie ich ze sobą odsyłaczem uzupełniającym. Zasada ta jest praktykowana w hasłach formalnych.

Przykład 4

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
100 1_ a Słomczyński, Maciej d (1920-1998) 400 1_ a Alex, Joe 400 1_ a Kwaśniewski, Kazimierz 667 ___ a pisarz, tłumacz, publikował pod pseudonimami: Joe Alex i Kazimierz Kwaśniewski	100 1_ a Słomczyński, Maciej d (1920-1998). 400 1_ a Słomczyński, M. 400 1_ a Słomczynski, Maciej d (1920-1998). 500 1_ a Kwaśniewski, Kazimierz d (1920-1998). 500 1_ a Alex, Joe d (1920-1998). 500 1_ a Cooper, Jack M. d (1920-1998). 667 ___ a Pisarz, tłumacz. Publikuje również pod pseudonimami: Joe Alex, Kazimierz Kwaśniewski.

Podobny problem pojawia się w wypadku, kiedy kilka osób występuje pod jednym pseudonimem.

Przykład 5

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
100 1_ a Safjan, Zbigniew d (1922-2011) 400 1_ a Zbych, Andrzej 667 ___ a o 667 ___ a prozaik, pseud. Andrzej Zbych wspólnie z Andrzejem Szypulskim 100 1_ a Szypulski, Andrzej d (1936-2011) 400 1_ a Janowicz, Remigiusz 400 1_ a Zbych, Andrzej 667 ___ a o 667 ___ a pisarz, scenarzysta, dziennikarz, Andrzej Zbych: pseudonim wspólnie ze Zbigniewem Safjanem	100 1_ a Zbych, Andrzej. 500 1_ a Safjan, Zbigniew d (1922-2011). 500 1_ a Szypulski, Andrzej d (1936-2011). 667 ___ a Pseud. zbiorowy Zbigniewa Safjana i Andrzeja Szypulskiego.

Library of Congress w takich sytuacjach stosuje pole 500. W niektórych przypadkach wydaje się to zasadne, ponieważ autor publikujący część swoich dzieł pod pseudonimem chciał wyraźnie zaznaczyć ich odrębność. Tworzenie tylko jednego rekordu wzorcowego dla autora i traktowanie pseudonimu jako formę odrzuconą uniemożliwia wydzielenie tej części jego twórczości, którą

z jakichś powodów wydzielił sam autor. Dobrym przykładem jest amerykańska pisarka Jayne Ann Krentz, która używa kilku pseudonimów w zależności od tego, czy powieść była o tematyce historycznej, futurystycznej czy współczesnej. Jest to o tyle interesujące, że czytelnik poszukujący jedynie romansów historycznych tej autorki znajdzie je wszystkie pod hasłem *Quick Amanda*. Biblioteka Narodowa natomiast przyjęła zasadę, że cała twórczość tej autorki jest zgromadzona pod hasłem *Krentz, Jayne Ann (1948-)*. W bazie ISNI stosuje się rozwiązanie, w którym część pseudonimów tej samej osoby otrzymuje osobny identyfikator.

Przykład 6

ISNI: 0000 0001 1447 1100
Name: Glass, Amanda
Dates: 1948-
Related identities: Krentz, Jayne Ann (other identity, same person)
Related names: Bentley, Jayne (1948- ; see also from)
Castle, Jayne (1948-; see also from)
James, Stephanie (1948-; see also from)
Krentz, Jayne Ann (1948-; see also from)
Krentz, Jayne Ann (see also from)
Quick, Amanda (1948-; see also from)
Taylor, Jayne (1948-; see also from)
Sources: VIAF BNE LC NKC BNF

ISNI: 0000 0000 7829 8020
Name: Castle, Jayne

ISNI: 0000 0001 1579 6193
Name: Amanda Quick

W tym wypadku można by przyjąć takie rozwiązanie, w którym dla zaprezentowania twórczości Joe Alexa lub opracowań historycznoliterackich na temat jego powieści kryminalnych, użyty zostałby pseudonim, natomiast w przypadku omawiania całej twórczości lub jego działalności jako tłumacza, wskazane byłoby użycie formy *Słomczyński, Maciej*. Jednak należy podkreślić, że w przypadku polskich autorów nie jest to duża skala problemu (dotyczy jeszcze np. Jana Pawła II oraz wcześniejszej formy Karol Wojtyła). W związku z tym należy przyjąć takie rozstrzygnięcie, które jest wygodne dla użytkownika. Trudno wskazać rozwiązanie idealne. Jeśli zastosowanych zostanie kilka rekordów wzorcowych powiązane polem 500, to informacje o jednej osobie będą rozproszone w bazie danych.

Hasła / tematy korporatywne

Hasło korporatywne według normy PN-N-01230 wyraża nazwę ciała zbiorowego według definicji: „Ciało zbiorowe to organizacja, grupa osób i/lub organizacji, występujące pod własną nazwą, jednoznacznie je identyfikującą”¹¹⁹. W JHP BN nie było ono stosowane w takiej samej formie jak w przypadku kwhf. Wynika to z odmiennego podejścia do haseł korporatywnych. Adam Łysakowski uważał, że hasło korporatywne jest bytem sztucznym i w zasadzie wyszukiwanie przez instytucje nie jest niezbędne. Tam, gdzie wydawało się zasadne zastosowanie nazwy instytucji, zalecał stosowanie nazw uproszczonych i powszechnie znanych, np. ZHP. Hasła dla instytucji były traktowane przez Łysakowskiego jako podgrupa haseł rzeczowych¹²⁰. Z czasem w katalogach alfabetycznych i przedmiotowych zaczęto stosować hasła korporatywne¹²¹, ale w wielu przypadkach sposób ich doboru nie przekłada się na postanowienia normy PN-N-01230. Oczywiście przez lata i w tej grupie haseł następowało pewne zbliżenie do haseł formalnych, nadal jednak różnice w doborze i redakcji haseł są znaczne. Wydaje się zasadne, aby w trakcie ujednolicania ogół haseł występujących w obu kartotekach podzielić na kilka rodzajów i w ten sposób eliminować punkty sporne w formie zapisu. Nie jest celem tego artykułu szczegółowa analiza porównawcza wszystkich rodzajów haseł. Istotne jest wskazanie stanu obecnego i rozbieżności, a w konsekwencji zakresu prac do wykonania podczas ujednolicania i fizycznej integracji obu kartotek.

Nazwy ciał zbiorowych międzynarodowych w przypadku kwhf nie są zapisywane konsekwentnie. Niektóre hasła mogą mieć formę zapisu w języku angielskim, np. *International Labour Organization* lub polskim, np. *Organizacja Narodów Zjednoczonych*, a czasami również występują w formie akronimu, np. UNESCO. Hasła w kwhf występują w języku polskim, np. *Międzynarodowa Organizacja Pracy* lub w formie akronimów, np. ONZ, UNESCO. Oczywiście norma dopuszcza obie formy zapisu, jednak obecnie należałoby przyjąć jedno z tych rozwiązań.

W przypadku ciał zbiorowych terytorialnych (związanych z władzą ustawodawczą, wykonawczą, sądowniczą, samorządową itp.) tj. działających na terenie określonej jednostki administracyjnej (państwo, województwo, miasto, etc.) w kartotece formalnej hasło zawsze rozpoczyna się od tej jednostki, a w drugiej kolejności zapisuje się nazwę własną danej instytucji.

119 PN-N-01230 : 2001. *Hasło opisu bibliograficznego. Hasło korporatywne.*

120 A. Łysakowski, *Katalog przedmiotowy: podręcznik*, Warszawa 1946, s. 75-84.

121 I. Gruchała, *Nazwa ciała zbiorowego jako hasło główne i dodatkowe opisu bibliograficznego*, „FIDES – Biuletyn Bibliotek Kościelnych” 2002, nr 2, s. 62-63 – <http://digital.fides.org.pl/publication/1527> [08.02.2015].

Przykład 7

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
110 2_ a Ministerstwo Kultury i Dziedzictwa Narodowego	110 1_ a Polska. b Ministerstwo Kultury i Dziedzictwa Narodowego.
410 2_ a MKiDN	410 2_ a Ministerstwo Kultury i Dziedzictwa Narodowego.
410 2_ a Ministerstwo Kultury	410 2_ a MKiDN.
410 2_ a Ministerstwo Kultury i Sztuki	510 1_ w a a Polska. b Ministerstwo Sztuki i Kultury.
410 2_ a Ministry of Culture and National Heritage	510 1_ w a a Polska. b Ministerstwo Kultury i Sztuki.
	510 1_ a Polska. b Ministerstwo Kultury.
	665 __ a Utworzone VII 1944 jako Resort Kultury i Sztuki PKWN, powołane w celu sprawowania opieki nad literaturą, teatrem, muzyką i sztukami pięknymi; podlega mu szkolnictwo artyst. i ochrona zabytków. Od 31.12.1944 nazwa Ministerstwo Kultury i Sztuki. Od października 1999 nazwa: Ministerstwo Kultury i Dziedzictwa Narodowego. Od 13.11. 2001 nazwa: Ministerstwo Kultury. 31 października 2005 zmiana nazwy na: Ministerstwo Kultury i Dziedzictwa Narodowego

Różnica w stosowaniu haseł dla ciał zbiorowych terytorialnych występuje na każdym szczeblu jednostki administracyjnej. O ile jednak w JHP BN dla jednostek centralnych z terenu Polski nie stosuje się dopowiedzeń lokalizujących, to dla jednostek regionalnych z terenu Polski w dopowiedzeniu wyraża się miejscowość będącą siedzibą / centrum regionu działalności. Dla instytucji spoza terytorium Polski w dopowiedzeniu wyraża się nazwę państwa lub miejscowość. Natomiast kwhf konsekwentnie stosuje hasła dla ciał zbiorowych terytorialnych w takiej samej formie, na każdym szczeblu jednostki administracyjnej.

Przykład 8

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
110 2_ a Rada Miejska (Żyrardów)	110 1_ a Żyrardów. b Rada Miejska.
451 __ a Żyrardów (woj. mazowieckie) x Rada Miejska	410 2_ a Rada Miejska m. Żyrardowa.

Warto też odnotować za Marią Nasiłowską i Małgorzatą Zarudzką, że „w Bibliotece Kongresu hasła rekordów dla nazw geograficznych wykorzystywanych

w opisie przedmiotowym zostały ujednoczone z nazwami geograficznymi wykorzystywanymi do budowania haseł korporatywnych (pierwszy element hasła zawierający nazwy organów władzy terytorialnej) i zakwalifikowane do podzbioru haseł bibliograficznych (*Names*)¹²². Przy opracowaniu haseł korporatywnych podstawowe rozbieżności między kartotekami są związane z wcześniejszymi i późniejszymi formami nazw instytucji. W hasle przedmiotowym używana jest nazwa ostatnia danego ciała zbiorowego, o ile nie nastąpiła jego faktyczna zmiana obejmująca przekształcenie struktury organizacyjnej, siedziby, organu prowadzącego itp. W kwłhf przesłanką do utworzenia nowego hasła jest zmiana nazwy ciała zbiorowego.

Przykład 9

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
110 2_ a Szkoła Główna Handlowa (Warszawa)	110 2_ a Szkoła Główna Handlowa w Warszawie.
410 2_ a Miejska Szkoła Handlowa w Warszawie	410 2_ a Szkoła Główna Handlowa (Warszawa).
410 2_ a SGH	410 2_ a Warsaw School of Economics.
410 2_ a SGPiS	410 2_ a SGH.
410 2_ a Szkoła Główna Planowania i Statystyki w Warszawie	510 2_ w a a Szkoła Główna Planowania i Statystyki w Warszawie.
410 2_ a Warsaw School of Economics	oraz
410 2_ a Wyższa Szkoła Handlowa w Warszawie	110 2_ a Szkoła Główna Planowania i Statystyki (Warszawa).
410 2_ a Szkoła Główna Warszawska	410 1_ a The Higher Course in National Economic Planning.
451 __ a Warszawa x Szkoła Główna Handlowa	510 2_ w b a Szkoła Główna Handlowa (Warszawa).
451 __ a Warszawa x Szkoła Główna Planowania i Statystyki	
510 2_ a Stowarzyszenie Skarbowców Wychowanków Szkoły Głównej Handlowej w Warszawie	
510 2_ a Szkoła Handlowa Leopolda Kronenberga (Warszawa)	
510 2_ a Towarzystwo Wyższej Szkoły Handlowej w Warszawie	

122 M. Nasiłowska, M. Zarudzka, *Hasła przedmiotowe rozwinięte w katalogu centralnym NUKAT. Krok ku postkoordynacji*, „Przegląd Biblioteczny” 2004, t. 72, z. 3/4, s. 209 – www.bbc.uw.edu.pl/dlibra/publication/165 [08.02.2015].

Opracowując przedmiotowo daną publikację w LCSH uwzględnia się w hasle nazwę ciała zbiorowego funkcjonującą w określonym czasie¹²³. To rozwiązanie ma swoje zalety, ponieważ hasło przedmiotowe odsyła nas do odpowiedniej nazwy w danym okresie działalności instytucji. Stosowanie jednego hasła w postaci ostatniej nazwy instytucji z jednej strony skupia informacje o tym samym bycie, z drugiej zaś uniemożliwia wydzielenie opisów dla okresów, w których występowała ona pod innymi nazwami.

Rozbieżności w zasadach doboru formy występują również w przypadku ciał zbiorowych kościelnych zarówno na poziomie centralnym, jak i lokalnym.

Przykład 10

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
110 2_ a Konferencja Episkopatu Polski. b Komisja Charytatywna	110 2_ a Ecclesia Catholica. b Konferencja Episkopatu Polski. b Komisja Charytatywna.
410 2_ a Konferencja Episkopatu Polski. b Komisja ds. Duszpasterstwa Dobroczynnego	410 2_ a Komisja Charytatywna Episkopatu Polski.
410 2_ a KCEP	410 2_ a Episkopat Polski. b Komisja Charytatywna.
410 2_ a Komisja Charytatywna Konferencji Episkopatu Polski	410 2_ a Kościół Katolicki. b Episkopat Polski. b Komisja Charytatywna.
410 2_ a Komisja ds. Duszpasterstwa Dobroczynnego Konferencji Episkopatu Polski	410 2_ a Kościół Katolicki. b Konferencja Episkopatu Polski. b Komisja Charytatywna.
665 __ a powstała w 1940 r.; zmiana nazwy z Komisji d.s. Duszpasterstwa Dobroczynnego na Komisję Charytatywną Episkopatu Polski w 1981 r.	

Przykład 11

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
151 __ a Łowicka, diecezja (katol.)	110 2_ a Ecclesia Catholica. b Diecezja Łowicka.
451 __ a Diecezja Łowicka	410 2_ a Diecezja Łowicka.
550 __ w g a Diecezje x katolicyzm z Polska	410 2_ a Kościół Katolicki. b Diecezja Łowicka.
667 __ a eryg. 1992 r.	665 __ a Diecezja Łowicka erygowana przez Jana Pawła II w 1992 i włączona do metropolii warszawskiej.

123 M. Nasiłowska, *Integracja kartotek haseł wzorcowych NUKAT, ibidem*, s. 220.

Nazwy własne kościołów, wyznań, zakonów mają również odmienne formy. Ujednolicenie ich jest o tyle trudne, że w zasadzie w poszczególnych krajach przyjęto odmienne rozwiązania. W kartotekach Biblioteki Narodowej stosuje się różne formy dla tego typu haseł.

Przykład 12

Kartoteka wzorcowych haseł przedmiotowych (JHP BN)	Kartoteka wzorcowych haseł formalnych
110 2_ a Jezuici	110 2_ a Societas Jesu.
410 2_ a SI	410 2_ a Jezuici.
410 2_ a SJ	410 2_ a Towarzystwo Jezusowe.
410 2_ a Societas Iesu	410 2_ a SJ.
410 2_ a Societas Jesu	410 2_ a Compagnie de Jésus.
410 2_ a TJ	410 2_ a Compagnia di Gesù.
410 2_ a Towarzystwo Jezusowe	410 2_ a Księża Towarzystwa Jezusowego.
	410 2_ a Księża Tow. Jezusowego.
	410 2_ a Gesellschaft Jesu.
	410 2_ a Księża Jezuici.
	665 ___ a Zakon założony w 1540 r. przez Ignacego z Loyoli. Do Polski sprowadził jezuitów kardynał S. Hozjusz w 1564 r.

Bibliothèque nationale de France przyjmuje, że przy jednostkach terytorialnych zakonu podaje nazwę własną, np. *Compagnie de Jésus. Prowincja polska*, ale *Jésuites – Éducation*. Library of Congress podaje hasło *Jesuits*.

Przedstawione powyżej przykłady wskazują, że ujednolicenie haseł nie jest problemem trudnym do pokonania. Opierając się na tych samych normach i zasadach, należy rozważyć, które rozwiązania są korzystniejsze dla użytkownika, i uświadomić sobie, jaki jest cel tworzenia kartotek wzorcowych.

VIAF I ISNI

Jak ważne jest ujednolicanie haseł i jakie z tego wynikają konsekwencje globalne, można się przekonać na podstawie takich baz jak VIAF i ISNI, które agregują dane baz wzorcowych bibliotek i innych instytucji z całego świata. VIAF jest bazą międzynarodową prowadzoną przez OCLC, która gromadzi dane z 34 instytucji w 29 krajach (stan na lipiec 2014). Liczy obecnie ok. 45 milionów rekordów wzorcowych oraz 105 500 000 rekordów bibliograficznych¹²⁴.

124 OCLC – 2014 VIAF Annual Report to VIAF Council – www.oclc.org/content/dam/oclc/viaf/OCLC-2014-VIAF-Annual-Report-to-VIAF-Council.pdf [08.02.2015].

Dane te wyraźnie wskazują, jak olbrzymie jest zainteresowanie współpracą w zakresie globalnej kontroli haseł wzorcowych i wzajemnym wspomaganie się przy identyfikacji poszczególnych osób, instytucji itp. VIAF staje się ważnym źródłem kontroli haseł wzorcowych na poziomie międzynarodowym. Grupuje hasła osobowe, korporatywne itp. z różnych bibliotek w klastry, które otrzymują wspólny numer identyfikujący¹²⁵. Informacje do VIAF przesyłane są na ogół przez biblioteki narodowe poszczególnych krajów lub inne centrale bibliograficzne¹²⁶. Łączenie w jeden wspólny klaster odbywa się automatycznie poprzez identyfikację odpowiednich pól z zastosowaniem narzędzi stosujących różnorodne algorytmy. Jeżeli w jakichś przypadkach są one błędnie rozpoznane, dana osoba, instytucja, tytuł, impreza nie zostaną włączone do tego samego klastra.

ISNI – International Standard Name Identifier – jest globalnym numerem, który służy jednoznacznej identyfikacji tożsamości podmiotów zaangażowanych w tworzenie, produkcję, zarządzanie i dystrybucję zawartości w sektorze produkcji nośników. System ISNI ma też ułatwić identyfikację podmiotów odpowiedzialnych za twórczy wkład w powstanie dzieła. Standard został wprowadzony międzynarodową normą ISO 27729, a w Polsce normą PN-ISO 27729:2014-04P ustanowioną w 2014 roku¹²⁷.

Przykład 13

ISNI:	0000 0001 1638 6225
1	2
Name:	Federowicz, Grażyna
Dates:	1959-
Creation class:	cre Language material Text
Creation role:	author contributor editor redactor

125 A. Angjeli, A. Mac Ewan, V. Boulet, *ISNI and VIAF transforming ways of trustfully consolidating identities* – <http://library.ifla.org/985/1/086-angjeli-en.pdf> [08.02.2015].

126 Polskę reprezentuje zarówno BN, jak i NUKAT.

127 *Polska Norma dotycząca identyfikatora ISNI* – <http://porzadek-rzeczy.bn.org.pl/polska-norma-dotyczaca-identyfikatora-isni/> [08.02.2015].

1	2
Related names:	Biblioteka Narodowa Biblioteka Narodowa. Wydawnictwo Dombek, Ewa Kaczynska, Maria Machnik, Hanna Maziarz, Iwona Skwirowska, Stefania Stasiak, Urszula Sterzycka, Elzbieta
Titles:	Bibliografia podziemnych druków zwartych z lat 1976-1989 Polish bibliography 1901-1939
Notes:	
Sources:	VIAF LC NLP NSK NUKAT SUDOC BNF BOWKER TEL

VIAF i ISNI były inspirowane przez program Universal Bibliographic Control (UBC)¹²⁸, który wyrósł z potrzeby dzielenia się wiedzą i promowania różnorodności kulturowej przy jednoczesnym usprawnieniu w bieżącej pracy bibliotek na świecie. Przez wiele lat ograniczenia technologiczne uniemożliwiały pełne wykorzystanie potencjału zasad UBC, którego idea jest to, aby każda krajowa agencja bibliograficzna była odpowiedzialna za rejestrację dorobku wydawniczego swojego kraju. Powinna ona objąć rejestracją bibliograficzną wszystkie typy dokumentów. VIAF i ISNI, korzystając z nowych technologii, potwierdziły, że te pryncypia, które zostały nakreślone podczas kongresu w Paryżu w 1977 roku, nadal są aktualne¹²⁹. Dzięki wspieraniu innowacji technicznych i organizacyjnych, oba systemy odgrywają ogromną rolę w całej gospodarce produkcji danych, ich kontroli, rozpowszechnianiu i wykorzystaniu. ISNI i VIAF ożywiły zasady UBC, otwierając nową erę gromadzenia i przetwarzania danych autorytatywnych dla osób i ciał zbiorowych w skali międzynarodowej i międzyinstytucjonalnej. Doświadczenie pracy w otwartym globalnym środowisku udostępniania danych i wiedzy zapewne rozszerza zakres zasad UBC oraz stwarza możliwości nieprzewidziane przez jego twórców. To oczywiście stawia nowe wyzwania. Pierwszym zasadniczym celem jest poprawienie jakości danych

128 *IFLA Professional Statement on UBC* – www.ifla.org/publications/ifla-professional-statement-on-ubc [08.02.2015].

129 *International Congress on National Bibliographies (ICNBS) Recommendations* – www.ifla.org/publications/international-congress-on-national-bibliographies-icnbs-recommendations [08.02.2015].

źródłowych. Oczywiście wydaje się, że wyższa jakość danych źródłowych daje większe korzyści zarówno dla VIAF i ISNI przy konsolidacji danych, jak i dla ISNI, kiedy dane wiążą się ze sobą¹³⁰.

Zakończenie

Katalogowanie oparte na encjach, jako postulat nowego spojrzenia na kwestię wyrażania danych bibliograficznych, wychodzi naprzeciw zmieniającym się zachowaniom informacyjnym użytkowników, które powinny być uwzględniane przy projektowaniu systemów informacyjnych i struktury danych. W artykule wskazaliśmy potrzebę większej niż dotychczas kompatybilności struktur metadanych bibliotecznych z modelami bazodanowymi. Kompatybilność danych umożliwia lepszą ich prezentację i szerszą obecność we współczesnych systemach informacyjnych bibliotek i globalnych agregatorów danych. Specyfika wyszukiwania fasetowego, w którym użytkownik samodzielnie kształtuje instrukcję wyszukiwawczą, wymaga od naszych baz danych jednostek prostych semantycznie i oznaczania w jak najszerszym zakresie relacji istniejących między desygnatami.

Dostosowanie danych do standardów wyszukiwania fasetowego wymaga między innymi:

1. Ścisłego rozdzielenia encji danych opisu formalnego i przedmiotowego. Wskazane jest wprowadzenie nowych pól formatu MARC 21, które umożliwią rozgraniczenie tych elementów, m.in. przeniesienie informacji o formie dzieła, czasie jego powstania, cesze językowo-etnicznej, przeznaczeniu czytelniczym i in. do pól grupy 3XX.
2. Stopniowego zwiększania kontrolowanych punktów dostępu dla elementów opisu formalnego i ograniczania analogowego sposobu zapisu informacji.
3. Ujednolicenia zasad doboru haseł z grupy pól 100-130 i scalenia kartotek kwhf i kwhp w jedną Bazę Deskryptorów BN.
4. Przeniesienia dużej części atrybutów encji z poziomu rekordu bibliograficznego, do rekordów wzorcowych poszczególnych instancji encji.
5. Odstąpienia od stosowania gramatyki pozycyjnej w opisie przedmiotowym i tworzenie Deskryptorów BN stosowanych w adekwatnych polach grupy 6XX.
6. Powrotu do źródeł katalogu przedmiotowego, stosowania wyszczególniającej wykładni dla przedmiotu i uogólniającej wykładni dla treści publikacji, odzwierciedlenia rzeczywistości przedstawionej w dokumencie w sposób logiczny i w języku, którym posługują się użytkownicy.

130 A. Angjeli, A. Mac Ewan, V. Boulet, *op. cit.*

Summary

Marta Cichoń, Jakub Kalinowski, Grażyna Federowicz *Cataloging based on entities*

The article presents the basis for changes in the scope of formal and subject matter descriptions. The main purpose of these changes is greater adherence of descriptions to FRBR and bringing the structure of bibliographic databases closer to relational databases. In this model, individual elements of bibliographic description are entities which are controlled instances of a common database of National Library Descriptors for both the formal and subject matter description of elements. An important element of the planned changes is to simplify the language of subject headings, including: elimination of positional grammar, revision of existing denotative-connotative concepts in information-retrieval languages, the use of natural vocabulary, and an increase in the number and role of logical relationship descriptors in the database.