

ALEKSANDRA ZAWADZKA
Biblioteka Uniwersytecka we Wrocławiu

Między lokalnością a potrzebami ministerialnymi — z doświadczeń rejestracji dorobku piśmienniczego Biblioteki Uniwersyteckiej we Wrocławiu w roku 2015

Dla wielu osób zajmujących się zawodowo rejestracją dorobku publikacyjnego pracowników uczelni wyższych i ośrodków badawczo-rozwojowych rok 2015 już na zawsze będzie stanowił istotną cezurę. Rok uruchomienia ogólnokrajowej bazy gromadzącej dorobek wielu jednostek naukowych pod nazwą Polska Bibliografia Naukowa (PBN) został zapewne wybrany nieprzypadkowo. Wdrożenie go ok. dwóch lat przed planowaną oceną parametryczną teoretycznie wydaje się okresem optymalnym: pracownicy wpisujący dane mają czas na zapoznanie się z systemem, pracownicy-autorzy mają czas na odpowiednie przygotowanie tekstów tak, aby spełniały wymagania nowego systemu raportującego, podobnie wydawcy. Każda zmiana zbliżonego typu zajmuje pewien czas: do nowych rozwiązań muszą się dostosować wszyscy uczestnicy tego procesu: autorzy, wydawcy, bibliografowie. Zmiany zaczęto wprowadzać od bibliografów, czyli od rejestracji. Rozwiązanie, zdawałoby się, praktyczne i sensowne, pod warunkiem jednakże, że wszystkie elementy tej zmiany będą jasno przedstawione i będzie czas na ich wprowadzenie.

Jednak sposób wdrożenia pierwszego w Polsce krajowego systemu gromadzącego dane o dorobku publikacyjnym jednostek naukowych daleki był od ideału. Nie można oprzeć się wrażeniu, że projektowi temu brakuje wyraźnej wizji, precyzyjnie obmyślnego i konsekwentnie realizowanego planu. Dowód? Według stanu z lipca 2016 r. istniały aż trzy wersje PBN. Tak nie powinno być. W każdym projekcie informatycznym w tym początkowym okresie stabilność rozwiązania, nawet nieidealnego, jest elementem wysoce pożądanym. Ma to działać na zasadzie: przetestujmy, zobaczmy, jak działa, poprawiajmy, ale nie wprowadzajmy

kilku konkurencyjnych rozwiązań, nawet jeśli funkcjonują one w jednej bazie danych. To wprowadza niepotrzebny chaos i zamieszanie, powodując negatywną ocenę projektu. Nowe dane plus multiplikujące się nowe sposoby ich przekazywania stanowią potencjalny punkt zapalny na linii programiści — pracownicy helpdesku — osoby wprowadzające dane. Tak naprawdę wszyscy uczymy się na żywym organizmie, nieustannie zmieniającym się. Pewne modyfikacje wprowadzane już w trakcie, w czasie trwania projektu dają się oczywiście wytłumaczyć i zrozumieć. Jednakże nie wszystkie.

Wizję takiego systemu ogólnopolskiego przedstawiła już w 2003 r. Maria Garczyńska¹ i reprezentanci bibliotek akademickich zgadzają się, że taki system jest niezbędny. Autorka podkreśliła konieczność wielu ustaleń przed wdrożeniem projektu i koordynacji prac, jako koordynatora proponując Konferencję Dyrektorów Bibliotek Szkół Wyższych. Zadanie realizuje jednak 12 lat później Ministerstwo Nauki i Szkolnictwa Wyższego (MNiSW). W początkowym okresie działania projektu PBN wydawało się, że i zleceniodawcy z MNiSW, i kolejni realizatorzy z ICM UW, ICI (Index Copernicus) w kwestiach szczegółowych mają zaledwie powierzchowną wiedzę. Odsyłanie od jednego do drugiego decydenta nie jest rzadkością. W wielu sprawach jesteśmy zostawieni sami sobie, nie ma wypracowanych standardów, schematów, propozycji, choć wydawałoby się, że specyfikacja pól i podpól przekazana w schemacie wzorcowym PBN .xml powinna być wystarczającą wskazówką. Z informatycznego punktu widzenia tak jest, ale nie chodzi tu przecież tylko o pola i ich nazwy — istotna jest zawartość tychże pól. Rozumiejąc jednocześnie prototypowość rozwiązania, tak w sensie ogólnopolskiej bazy gromadzącej dane o dorobku jednostek naukowych, jak i brak doświadczenia Index Copernicus w tworzeniu wielod dziedzinowej bazy bibliograficznej, należy podkreślić, że na poziomie jednostki naukowej i biblioteki wielokrotnie wprowadzano dane, działając intuicyjnie, opierając się raczej na wieloletnim doświadczeniu własnym i literaturze przedmiotu niż na wskazówkach producenta oprogramowania.

Uczelnie korzystają z różnych systemów i z różnych sposobów opisywania zgłaszanych publikacji. Są bazy lokalne zorientowane bardziej bibliograficznie i takie, które skupiają się wyłącznie na raportowaniu danych do MNiSW, co w obu przypadkach przekłada się na sposób i zakres prezentowanych informacji, a nawet dobór materiału do bazy.

Głosy dotyczące funkcjonowania samej PBN to, jak dotąd, raczej opinie użytkowników niż naukowe analizy. Głównie wypowiedzi na blogach². Najnowszym i równocześnie rzetelnym głosem jest opublikowany w maju 2016 r. 165 numer

¹ M. Garczyńska, *Bibliografie publikacji pracowników państwowych szkół wyższych w Polsce w świetle badań ankietowych*, „Przegląd Biblioteczny” 2003, z. 1, s. 74.

² Np. E. Kulczycki, *Lepiej jak jest więcej, czyli Polskie Bibliografie Naukowe*, http://ekulczycki.pl/warsztat_badacza/lepiej-jak-jest-wiecej-czyli-polskie-bibliografie-naukowe/ [dostęp: 12.02.2017]; E. Rozkosz, *Nowy moduł sprawozdawczy Polskiej Bibliografii Naukowej. Część 1*,

„Biuletynu EBIB”. Przyniósł on bardzo ciekawe teksty dotyczące aktualnego stanu raportowania dorobku do PBN i dalej do MNiSW. Swoje problemy, a także ich rozwiązania przedstawiły osoby związane z uczelniami sprofilowanymi: technicznymi (PWr³ i PW⁴) oraz uczelnią medyczną (GUMed⁵). Zabrakło głosu uczelni humanistycznej bądź wielodyscyplinarnej.

PBN w Uniwersytecie Wrocławskim

Baza Bibliografia publikacji pracowników i doktorantów Uniwersytetu Wrocławskiego (BPPiD) wchodziła w projekt pn. PBN bez doświadczenia sprawozdawczego jako przekaznik danych do MNiSW. Dla tak dużej uczelni, jaką jest Uniwersytet Wrocławski, jedynym sposobem dającym szansę jak najpełniejszej realizacji nowego obowiązku sprawozdawczego był import masowy. Baza ogólnouczelniana funkcjonuje w formie elektronicznej od roku 2004, od 2005 w wersji on-line. Dorobek aktualizowany jest najczęściej w odstępach tygodniowych, choć zdarzają się też dłuższe przerwy. Operację przekazywania danych do PBN zdecydowano się przeprowadzić własnym sumptem bez odwoływania się do systemów zewnętrznych przy maksymalnym zaangażowaniu osób tworzących bazę i informatyka zatrudnionego w Bibliotece Uniwersytetu Wrocławskiego (BUWr).

Aby dokładnie uzmysłowić skalę problemu, z jakim musiał zmierzyć się zespół Sekcji Prac Bibliograficzno-Dokumentacyjnych BUWr, konieczne jest przedstawienie danych dotyczących uczelni i jej dorobku publikacyjnego. Uniwersytet Wrocławski jest dużą jednostką: według stanu na 29 lutego 2016 r. uczelnia zatrudnia 1882 nauczycieli akademickich i 1499 osób niebędących nauczycielami akademickimi. Doktoranci kształcący się w uczelni to 1558 osób (stan na 31 grudnia 2015 r.)⁶. Uczelnia składa się z dziesięciu wydziałów: Wydział Biotechnologii, Wydział Chemii, Wydział Filologiczny, Wydział Fizyki i Astronomii, Wydział Matematyki i Informatyki, Wydział Nauk Biologicznych, Wydział Nauk

<http://www.edukacjainformacyjna.pl/2016/08/nowy-modul-sprawozdawczy-polskiej-bibliografii-naukowej-czesc-1/> [dostęp: 12.02.2017].

³ A. Komperda, *Import publikacji Politechniki Wrocławskiej do Modułu Sprawozdawczego PBN — sprawozdanie z realizacji zadania*, „Biuletyn EBIB” 2016, nr 165, s. 1–8, <http://open.ebib.pl/ojs/index.php/ebib/article/view/434> [dostęp: 12.02.2017].

⁴ M. Miller-Jankowska, M. Lewandowska-Tranda, *Przekazywanie danych do Systemu Informacji o Nauce poprzez uczelnianą bazę wiedzy. Z doświadczeń Politechniki Warszawskiej*, „Biuletyn EBIB” 2016, nr 165, s. 1–9, <http://open.ebib.pl/ojs/index.php/ebib/article/view/435> [dostęp: 12.02.2017].

⁵ B. Siek, *Polska Bibliografia Naukowa a modyfikacje indeksu autorskiego bazy Bibliografia GUMed*, „Biuletyn EBIB” 2016, nr 165, s. 1–8, <http://open.ebib.pl/ojs/index.php/ebib/article/view/437> [dostęp: 12.02.2017].

⁶ Dane za: Biuletyn Informacji Publicznej Uniwersytetu Wrocławskiego, <http://bip.uni.wroc.pl/183/55/informacje-ogolne.html> [dostęp: 12.02.2017].

Historycznych i Pedagogicznych, Wydział Nauk Społecznych i Wydział Prawa, Administracji i Ekonomii. Jednostki te reprezentują połowę obszarów nauk funkcjonujących w Polsce (4/8), tj. obszar humanistyczny, społeczny, przyrodniczy i ścisły. Każdy z nich ma swoją specyfikę w zakresie sposobów publikowania: dla nauk humanistycznych, społecznych i prawniczych najbardziej powszechną formą publikowania są książki i rozdziały, dla nauk z zakresu STM — artykuły. Teksty wieloautorskie to domena nauk ścisłych i technicznych. Nasi autorzy publikują w wielu krajach: od Stanów Zjednoczonych, przez Wielką Brytanię, Polskę, Japonię, Niemcy, aż po Singapur, Węgry i Rumunię. Również w zakresie języków publikacji panuje ogromna różnorodność: oprócz tekstów w językach angielskim i polskim można odnaleźć prace napisane w alfabecie cyrylicy czy np. tłumaczenia z języka francuskiego, włoskiego, łaciny, starogreckiego, jidysz. W tytułach prac wykorzystywane są znaki literowe z różnych języków, ale także matematyczne i fizyczne, wymagające specjalnego zakodowania. Taka różnorodność zasobu nie ułatwiała pracy nad projektem PBN.

PBN oficjalnie zaanonsowano w Rozporządzeniu MNiSW z dnia 29 czerwca 2015 r. w sprawie Systemu Informacji o Nauce, dając czas na wgranie danych za lata 2013–2015 do 30 października 2015 r., czyli w ciągu czterech miesięcy. Sam system ruszył w sierpniu 2015 r. Do 1 października 2015 r. wprowadzono do bazy BPPiD UW r 13 469 rekordów za lata 2013–2015. Mając zatem do opracowania ponad 10 tys. rekordów w terminie sierpień–październik 2015 r., konieczne stało się przeanalizowanie istniejących już rekordów w celu odpowiedzi na pytania:

1) jak uzupełnić rekordy o dane wymagane w Rozporządzeniu MNiSW z dnia 29 czerwca 2015 r. w sprawie Systemu Informacji o Nauce: dopisanie pól i podpól, uruchomienie nieużywanych dotąd pól i podpól istniejących w schemacie MARC, konieczność dopisywania informacji tekstowej zamiast stosowanej dotychczas informacji w postaci liczbowej lub kodowanej (008 w – 1 --> true)?

2) jak przełożyć używany w bazie BPPiD UW r format MARC na wymagany przez PBN .xml?

Prace dostosowawcze

Należy podkreślić, że w roku 2015 tak naprawdę funkcjonowały dwa zupełnie różne systemy raportowania dorobku o nazwie PBN: pierwszy do 8 sierpnia 2015 r., tzw. z ICM (repozytoryjny) i drugi po tej dacie, tzw. z ICI (sprawozdawczy). Każdy z nich korzystał z innego umocowania prawnego. Co więcej, ostateczne Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego ws. Systemu Informacji o Nauce w sposób zasadniczy różniło się od projektu będącego pod obradami sejmowej Komisji Edukacji, Nauki i Młodzieży. Z niektórych elementów, jak np. informacji prawach autorskich zrezygnowano, inne (szczegółowa infor-

macja o *open access*) dodano. A to właśnie projekt był podstawą pierwszych prac programistycznych, natomiast aplikacją typowaną jako System Informacji o Nauce wydawał się PBN-ICM. W związku z kompletną zmianą filozofii i operatora systemu po sierpniu 2015 r. okazało się, że wszelkie działania podjęte do tej daty były w dużym stopniu bezużyteczne.

Koronnym przykładem będzie przywołanie sprawy identyfikacji autorów. Problem polegał na tym, że w bazie BPPiD hasła osobowe identyfikujące autorów wyprowadzane są w sposób bibliograficzny, tj. na podstawie zapisów w publikacjach. Powiązany z ówczesną wersją PBN POL-on dysponuje natomiast danymi z dowodów osobistych, które w wielu przypadkach nakazywały rozszerzenie nazwy pracownika, np. o dodatkowe imię. Każdy pracownik dostawał też inne ID niż istniejące w kartotece osobowej prowadzonej w Sekcji. W systemie PBN-ICM konieczne było zatem powiązanie numeryczne i literowe opisu pobieranego z systemu lokalnego z danymi POL-on. Aby tego dokonać, pracownicy Sekcji Dokumentacji i Sekcji Komputeryzacji BUWr wystąpili do Działu Kadr UWr z prośbą o udostępnienie danych osobowych przekazywanych do POL-on celem stworzenia list weryfikacyjnych, które proces identyfikacji autorów pozwoliłyby zautomatyzować, przekładając wersję nazwiska stworzoną w SPBD na zapis stosowany w POL-on. Takiej zgody nie uzyskano, dlatego trzeba było wystąpić o nadanie stosownych uprawnień do systemu POL-on i pobranie ich na własną odpowiedzialność. Przygotowane dużym nakładem pracy informatyka listy pracowników Sekcja otrzymała pod koniec sierpnia 2015 r., kiedy okazało się, że w nowym systemie (PBN-ICI) zrezygnowano z takiego połączenia.

Wobec dużych zmian i niepewności co do przyszłego kształtu informacji w zakresie identyfikacji osobowej zdecydowano o całkowitym wyodrębnieniu tych danych w rekordach — powstało nowe, powtarzalne pole Osoba PBN. Odnotowujemy w nim:

- dane osobowe — obecnie w brzmieniu bibliograficznym,
- opis roli osoby (wybór między wartościami: „author” i „editor” (w wersji z pierwszej połowy 2015 r. był jeszcze „translator”),
- opis afiliacji (wybór między wartościami „true” i „false”),
- opis zatrudnienia (wybór między wartościami „true” i „false”).

Istnieją również podpola umożliwiające m.in. wpisanie dokładnej afiliacji podanej przez autora w publikacji (wykorzystywane tylko dla naszych autorów), oznaczenia numerycznego dla autora i jednostki.

Dużym wyzwaniem okazały się wymagania dotyczące bardzo dokładnego rozpisania afiliacji i zatrudnienia autorów. Praktyka osób wpisujących dane do BPPiD UWr do sierpnia 2015 r. była taka, że parametr „praca afiliowana” traktowano bardzo ogólnie — na zasadzie, czy którykolwiek z autorów wskazał Uniwersytet Wrocławski w afiliacji. Nowe aplikacje PBN wymogły wpisywanie bardzo szczegółowego opisu wymagającego podania

nazwiska i co najmniej inicjału imienia dla każdego z autorów (dla autora własnego musi być to nazwisko i pełne imię) oraz dzielącego autorów według zatrudnienia i według afiliacji, przy czym poprzez afiliację rozumie się wydział, a nie uczelnię. Dla prac jednoautorskich udało się takie dane wprowadzić operacją grupową, kopiując po prostu do odpowiedniego pola zapis nazwiska i imienia, dopisując określenie „author”, dane o zatrudnieniu i afiliacji pobierając z innych pól, przy czym wymagało to przełożenia zapisu kodowanego na wymaganą przez PBN-ICI wartość tekstową „true” albo „false”. Problem jednak pozostawał poważny dla prac wieloautorskich (ok. 5019 rekordów). Stosowana konsekwentnie od lat zasada rozpisywania nazwisk i pełnych imion kolejnych współautorów w polach osobowych bibliograficznych pozwoliła chociaż na przeprowadzenie operacji kopiowania tych danych do pól, które miały posłużyć do przekazania informacji do PBN. Natomiast wskaźniki dotyczące afiliacji („true”/„false”) i dopisanie parametru „other contributors”, polegającego na zliczeniu autorów nieafiliowanych, musiało być wykonane ręcznie.

Kolejny problem dotyczył rozdziałów, czyli części niesamoistnych książek (ok. 4519 rekordów). W praktyce bibliograficznej nie ma potrzeby wpisywania do opisu rozdziału danych osobowych redaktorów pracy zbiorowej w formie hasła osobowego — istnieje tylko zapis pobrany z samej pracy zbiorowej. Taki wymóg jednak postawiło przed nami PBN. Dodatkowo wymóg ten początkowo był obowiązkowy dla wszystkich rozdziałów, potem jednak metodą dopytywania się w helpdesku uzyskano pozwolenie na dopisywanie tej informacji wyłącznie dla rozdziałów z prac redakcyjnych, których redaktorami byli nasi pracownicy i/lub nasi afiliujący. Oczywiście w mocy pozostawał wymóg dzielenia redaktorów na zatrudnionych/niezatrudnionych i afiliowanych/nieafiliowanych na konkretnych wydziałach UW. W tym zakresie prace objęły przede wszystkim wydziały filologiczne i społeczne.

Rozdzielenie ról autora i redaktora zaowocowało koniecznością dopisywania, na ogół ręcznego, angielskich wersji nazw tych ról. Wykorzystanie języka obcego znacząco utrudniło tę pracę — bardzo łatwo o pomyłkę, zwłaszcza w przypadku wartości „author”.

Zupełną nowością w zakresie wpisywanych i raportowanych danych był wymóg rozliczania prac z wyliczeniem ich objętości w arkuszach wydawniczych. Dane te były wcześniej wymagane przez MNiSW, ale dopiero na etapie składania wniosku o ocenę parametryczną. Brak jednoznacznej definicji arkusza wydawniczego oraz konieczność wliczania do niego również ilustracji i map czyniła ten element szczególnie niepewnym i pracochłonnym. Napotkano bardzo duży opór środowiska akademickiego przed podawaniem tej wartości. Niejasne jest też, kto ma ją znać i wyliczać: autor czy też pracownik zgłaszający pracę do bazy BPPiD UW (na ogół bibliotekarz). Wydawcy, posiadający najłatwiejszy dostęp do wyliczenia arkuszy, też nie są skłonni do dzielenia się tą wiedzą. Wobec dużej

liczby rekordów wymagających uzupełnienia o ten parametr zdecydowano o jego fakultatywności, szczególnie podczas przekazywania tzw. pierwszej transzy, tj. do końca października 2015 r. W miarę upływu czasu bibliotekarze nauczyli się wyliczać arkusze, sami autorzy zgłaszający pracę podają tę wartość, ale ciągle nie jest to zjawisko powszechne.

Prowadząc prace dostosowawcze bazy w pierwszym kwartale 2015 r., zdecydowano się dopisywać do istniejących rekordów artykułów informacje o danych czasopisma na liście ministerialnej — oczywiście, o ile czasopismo się na tej liście znalazło. W zakres tej informacji wchodziły następujące dane:

- oznaczenie jest/nie jest na liście ministerialnej,
- maksymalna przewidziana liczba punktów za publikację w danym tytule,
- ISSN ministerialny (zdarzało się, że był inny od rzeczywistego, wydrukowanego na czasopiśmie),
- tytuł ministerialny (odstępstwa od tytułu bibliograficznego stosownego w bazie ISSN czy NUKAT),
- oznaczenie listy.

Po zmianie systemu w sierpniu 2015 r. okazało się, że i ta praca była niepotrzebna, bo MNiSW wycofało się z wymogu pobierania tych danych z PBN. Ale w rozporządzeniu w sprawie Systemu Informacji o Nauce wprowadzono konieczność wpisywania numerów ISSN i ISBN, co było kolejną zupełną nowością dla osób współpracujących z BPPiD UW. Dotychczas parametr ten nie był wpisywany do rekordu. Aby przyspieszyć prace, skopiowano zatem już wpisane ISSN-y ministerialne do odpowiedniego, wybranego pola, a braki uzupełniano ręcznie. Dopisywanie ISBN-ów odbywało się wyłącznie ręcznie lub poprzez operacje kopiowania z baz wydawniczych i bibliotecznych.

Konieczność przekazywania danych o wersji elektronicznej można uznać za awans dla tej formy publikowania. W zasadzie zrównane zostały w swoim znaczeniu publikacje typu *born print* i *born digital*. Wydaje się jednak, że przyjęto tutaj założenie: najpierw praca drukowana, potem wersja elektroniczna.

W rekordach istniejących w BPPiD UW opis adresu dostępowego do wersji elektronicznej istnieje już od 2006 r. Dopisywano tak adresy klasyczne <http://www...>, jak i DOI. Jednakże ten ostatni zapis musiał przejść modernizację. Wobec braku interpretera DOI w systemie obsługującym bazę BPPiD UW numer ten zapisywano jako adres internetowy, prowadzący do interpretera i dalej do konkretnego artykułu czy abstraktu. W schemacie PBN .xml przewidziano zupełnie osobne pole na DOI bez żadnego poprzedzającego go klasycznego adresu internetowego. Podczas przygotowywania tzw. pierwszej transzy, tj. do końca października 2015 r., z braku czasu udało się tylko niektóre DOI przenieść do postaci skróconej, uznając, że wymóg wskazania wersji elektronicznej będzie spełniony adresem rozszerzonym.

W roku 2015 doczekaliśmy się także jako środowisko ministerialnej polityki *open access* (OA). Bezpośrednim wyrazem uznania dla tej formy upubliczniania

(pełny tekst pracy w nielimitowanym dostępie) jest konieczność przekazywania danych o sposobie udostępniania wersji elektronicznej dokumentu, przy czym założono, że jeśli dokument udostępniany jest w wersji licencjonowanej, należy przekazać wyłącznie adres dostępowy. W przypadku *open access* konieczne jest podanie bardzo szczegółowych informacji. Pomimo prowadzenia od kilku lat tak szeroko zakrojonych akcji jak Open Access Week czy konferencji z tego zakresu, znajomość problematyki OA nadal pozostaje w środowisku akademickim niska. Nie mają wiedzy ani sami autorzy, ani bibliografowie. Problematyczność przekazywania danych o wersjach elektronicznych, w tym o OA, podkreślano już w 165 numerze „Biuletynu EBIB”⁷. Tutaj chciałam zwrócić uwagę, że wobec istnienia licencji krajowych na wiele baz pełnotekstowych osobom wprowadzającym i korygującym niekiedy jest bardzo trudno bez wątpliwości wskazać, który tekst jest OA, a który nie. Mając do nich dostęp przez licencję krajową, niepotrzebne są logowania i opłaty, bo już zostało to uregulowane ogólnie. Powszechna jest niezajomość licencji, na których upubliczniane są wersje elektroniczne dokumentów. Oznaczenia, jakie należy wpisywać do pól (np. CC-BY, CC-BY-SA), nie dość, że są inne niż zapis samego licencjodawcy — oryginalnie brak myślnika między CC i dalszą nazwą licencji — to z racji swego angielskojęzycznego brzmienia nie „mówią”.

W kontekście OA dziwią niektóre rozwiązania zastosowane przez twórców schematu PBN .xml. Bardzo wiele wątpliwości budzą listy słownikowe, tj. konkretne zapisy, spośród których należy wybierać. Odnotowując artykuł OA dla parametru „sposób udostępnienia” mamy wybór: „OPEN_JOURNAL”, „OPEN_REPOSITORY” i „OTHER” (pisownia oryginalna wersalikami). W wielu przypadkach artykułów jest to wybór niewystarczający. Istnieją bowiem czasopisma o pełnym OA, hybrydowym OA, częściowym OA. I dla tych wszystkich przypadków mamy wpisywać „OPEN_JOURNAL”? Dużo czytelniejsze byłoby OPEN_ARTICLE, skoro i Rozporządzenie MNiSW ws. Systemu Informacji o Nauce i późniejsze tzw. rozporządzenie parametryzacyjne⁸ mówią o linku do pełnej wersji artykułu. Artykułu, nie czasopisma (*journal*). W obecnej sytuacji wartość słownikowa OTHER będzie dużo obszerniejsza niż inne kategorie. Poza tym każda jednostka będzie wpisywała według własnego odczucia. W konsekwencji czasopismo z częściowym OA będzie raz zgłoszone jako OPEN_JOURNAL, a z innej instytucji jako OTHER. Jeśli wchodzimy w promowanie OA, a tak należy odebrać upublicznienie przez MNiSW polityki *open access*, informacje pozyskane przez PBN będą z gruntu fałszywe.

⁷ E. Rozkosz, A. Drabek, *Moduł sprawozdawczy Polskiej Bibliografii Naukowej*, „Biuletyn EBIB” 2016, nr 165, s. 1–18, <http://open.ebib.pl/ojs/index.php/ebib/article/view/438/596> [dostęp: 12.02.2017]; A. Komperda, *op. cit.*, s. 7.

⁸ Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 27 października 2015 r. w sprawie kryteriów i trybu przyznawania kategorii naukowej jednostkom naukowym.

Kolejny problem to tzw. teksty *online first*. Jeśli intencją ministerstwa jest promowanie *open access*, należałoby zezwolić na zgłaszanie artykułów będących wersjami nieostatecznymi, tj. bez przyporządkowanego im rocznika, numeru i stron, ale już upublicznionych on-line, recenzowanych, w pełnej wersji, najczęściej z datą publikacji on-line, posiadających DOI i niekiedy już cytowanych. Jeśli przyjmujemy rozumienie słowa, „udostępnienie” bardzo szeroko, artykuł *online first* jest tekstem udostępnionym. Jednakże chcąc zgłosić taki tekst do PBN, trzeba wpisać do elementu „liczba miesięcy od momentu opublikowania do momentu udostępnienia w OA” wartość 0, a dowiedziono empirycznie, że system PBN-ICI takiej wartości nie przyjmuje. Problem ten rozciąga się też na wszystkie czasopisma z natychmiastowym OA — tekst jest udostępniony OA bez zwłoki w miesiącach, czy mamy podawać fałszywą wartość 1? Wydaje się, że przyjęto tutaj założenie: najpierw wersja drukowana, potem on-line. Niestety czasem jest odwrotnie, jednak nie przewidziano rozwiązania.

Zgodnie z interpretacją przepisów przez kolejnych konsultantów helpdesku wersję elektroniczną, w tym *open access*, zgłasza się według stanu na dzień raportu. Kłopot polega na tym, że niektórzy wydawcy stosują tzw. *moving wall* (embargo wydawnicze), czyli okres, kiedy tekst jest udostępniany wyłącznie na licencji, a po jego upływie jako *open access*. Jeśli *moving wall* kończy się przed końcem czteroletniego okresu objętego ewaluacją, niesprawiedliwym wydaje się niemożność poprawienia rekordu dotyczącego takiej publikacji. Stwierdzenie to budzi szczególny niepokój, jeśli założymy, że promowanie *open access* będzie miało przełożenie na wartości punktowe w jednym z obszarów ewaluacji, niekoniecznie publikacji naukowych.

Przygotowanie pliku XML

Równoległe z dostosowywaniem „wnętrza” rekordów trwały prace nad rozwiązaniem informatycznym pozwalającym pobierać dane w odpowiednim formacie z bazy BPPiD UW. W tym zakresie również pierwotny plan z początku roku 2015 zdecydowanie różnił się od ostatecznego rozwiązania zaimplementowanego we wrześniu 2015 r. Działanie to wymagało dużej cierpliwości i dokładności. Jako wzorca .xml użyto plik PBN .xml przedstawiony na stronie www PBN w dziale pomocy. W wyniku prac programistycznych został stworzony interpreter przekładający dane z formatu MARC na format PBN. Stopniowo udawało się informacje pobrane z bazy zapisywać w nowy sposób. Najpierw opracowano schemat dla artykułów, następnie dla książek i w końcu dla rozdziałów. W toku prac okazywało się, jak wielu szczegółów nie opublikowano na stronie www PBN bądź wręcz nie ustalono na linii Index Copernicus–MNiSW. Zabrakło informacji o sposobie odwzorowania alfabetów obcych: transliteracja czy transkrypcja, a może jednak oryginalna

cyrylica lub jidysz? W jaki sposób opisywać pojedyncze obce litery np. akcentowane „e” w języku francuskim, „s” basowe w języku niemieckim? Jak przekazywać dane o indeksach górnych i dolnych, znakach matematycznych i fizycznych w tytułach prac? Dopiero z maili do helpdesku wyłaniały się te ustalenia: tablice kodowania znaków, a także tablice kodowania nazw krajów i języków. Okazało się też, że w upubliczonym schemacie PBN .xml zabrakło elementu opisującego arkusze wydawnicze: tak jego nazwy, jak i sposobu wypełniania. Kłopotliwy był także element „pages”, który dopiero na prośby importerów rozszerzono do Nielimitowanej liczby znaków. Duże rozbieżności interpretacyjne budzi element „other contributors” — z zapisów rozporządzenia wynika, że powinno się podać tę wartość liczbowo, ale czy powinno się wliczać tu autorów w roli „editor”? Odpowiedź z helpdesku PBN-ICI nie była jednoznaczna⁹.

Funkcja PBN w bazie lokalnej

Mając ustalony sposób konwertowania danych z bazy BPPiD UW r, przygotowano tzw. funkcję PBN. Jej zadaniem jest przede wszystkim umożliwienie Importerom poszczególnych wydziałów eksportowania rekordów do PBN. W pierwszej połowie roku 2015 wydawało się, że instytucją przekazującą dane do PBN będzie Biblioteka Uniwersytecka jako organ prowadzący bazę lokalną. W sierpniu 2015 r. okazało się jednak, że sytuacja jest zupełnie inna — konieczne są osobiste upoważnienia od dziekanów. Importerami zostały wyznaczone osoby o różnym doświadczeniu bibliograficznym, również bez tego doświadczenia. Wobec takiego postawienia sprawy — to jednostka naukowa decyduje o tym, co przekazuje do PBN — zdecydowano o szerokim udostępnieniu funkcji PBN dla społeczności akademickiej UW r. Miało to nie tylko pomóc w testowaniu przedstawionego rozwiązania, lecz także umożliwić jednostkom ingerencję w wygenerowane pliki .xml w celu ich poprawienia czy uzupełnienia bez konieczności absorbowania pracowników Sekcji Prac Bibliograficzno-Dokumentacyjnych. Nie zdecydowano się na hasłowanie dostępu, pozostawiono swobodę wyboru rekordów (opisów).

Sama procedura polega więc na:

- wyborze wydziału (wymóg PBN, by podawać w plikach .xml PBN_ID wydziału),
- wskazaniu rekordów (pojedynczo lub grupowo: wszystkie, nowe, zmodyfikowane),
- wciśnięciu przycisku „Akceptuj”.

⁹ Sprawę jednoznacznie wyjaśniła interpretacja MNiSW ogłoszona na stronie www PBN w sierpniu 2016 r. i rozporządzenie MNiSW z dnia 12 grudnia 2016 r. w sprawie przyznawania kategorii naukowej jednostkom naukowym i uczelniom, w których zgodnie z ich statutami nie wyodrębniono podstawowych jednostek organizacyjnych (ogł. 27 grudnia 2016 r. z datą wejścia w życie 1 października 2017 r.).



Rys. 1. Fragment ekranu z interfejsu bazy danych BPPiD UWr z funkcją Polska Bibliografia Naukowa

Źródło: <http://www.bu.uni.wroc.pl/publikacje>.

Wygenerowany plik .xml należy zapisać na dysku i przygotować dla niego folder(y) oraz skompresować do formatu .zip przed wysłaniem do PBN¹⁰.

Nie ulegało jednak wątpliwości, że Sekcja Prac Dokumentacyjnych własnymi siłami nie uzupełni rekordów we wszystkich wymaganych zakresach. O potrzebach w zakresie danych wymaganych przez MNiSW w ramach raportowania dorobku publikacyjnego poinformowano zatem osoby dotychczas współpracujące z Sekcją oraz władze instytutów, wydziałów i uczelni. Aby je zobrazować i umożliwić ich zdobycie bezpośrednio na wydziałach, przygotowano w funkcji PBN tzw. ekrany pomocnicze pozwalające zorientować się w kompletności przekazywanych danych. Na ekranach dla poszczególnych zakresów informacji umieszczono tabele ze szczegółowym wyliczeniem koniecznych danych. Tabele te zostały odebrane jako bardzo pomocne.

Kolejny etap to przekazywanie danych do PBN. Po pokonaniu bariery rejestracyjnej (upoważnienia, loginy, hasła) dużym zaskoczeniem okazały się raporty o błędach występujących podczas importu. Aplikacja importująca PBN-ICI analizuje plik wskazany jako źródło danych i akceptuje lub odrzuca w całości, generując informację o numerze linii, w której znajduje się błąd. Należało nauczyć się analizować pliki .xml, a także interpretować nazwy błędów. Dla przykładu — określenie rodzaju błędu jako „boolean” jest kompletnie nieczytelne. Listy opisowej błędów nie ma. Dopiero metodą prób i błędów dowiedzieliśmy się, że

¹⁰ Po wdrożeniu kolejnej wersji aplikacji PBN Moduł Sprawozdawczy przygotowanej przez OPI w maju 2016 r. uproszczono procedurę importu masowego, rezygnując z konieczności kompresowania plików .xml.

„boolean” to taki błąd, gdzie w rekordzie brakuje informacji o zatrudnieniu, afiliacji autora lub nie podano wartości liczbowej dla „other contributors”. Bezpośrednim powodem mogło być przeniesienie danych w rekordzie oryginalnym do innego podpola lub faktycznie brak danych. Tak wyłapany błąd należało zidentyfikować w pliku xml (odnalezienie numeru linii) i bądź od razu poprawić w aplikacji wewnętrznej Sekcji lub pliku .xml pobranym z bazy lokalnej, bądź wrócić do funkcji PBN, „odznaczyć” ten konkretny zapis, ponownie wygenerować plik .xml i próbować dalej. Niestety aplikacja importowa PBN-ICI nie informowała o wszystkich błędach znalezionych w danym pliku, ale zatrzymywała import na konkretnej linijce, gdzie ten błąd wystąpił bez wczytywania poprawnych danych, które znajdowały się w tym samym pliku, co skutkowało niekiedy wielokrotnymi poprawkami tych samych zbiorów danych.

Dopiero próbne importy wskazały też, jakie dane są kluczowe do przyjęcia zgłoszenia przez PBN. Początkowo wydawało się, że będą to imiona i nazwiska autorów. W trakcie testów i próbnych dogrywek okazało się, że kluczowe są: zatrudnienie, afiliacja i „other contributors”.

Wspomaganie importerów przez Sekcję Prac Bibliograficzno-Dokumentacyjnych polegało na udostępnianiu szczegółowych instrukcji, udzielaniu informacji telefonicznych i mailowych. Skorzystano też z innowacyjnego rozwiązania wspólnej analizy plików przekazywanych do Sekcji za pośrednictwem chmury — umożliwiła to wdrożona w Uniwersytecie usługa Microsoft Office wraz z dyskiem wirtualnym OneDrive.

Bezpośrednią konsekwencją tak dużego obciążenia zadaniem specjalnym PBN było zamknięcie bazy BPPiD UW, tj. brak jej aktualizacji w październiku 2015 r. Jednostki chciały zgłosić jak najwięcej dorobku, jednak ze względu na to, że czas opracowania rekordu znacznie się wydłużył, nie było możliwości równoczesnego wprowadzania dużych zmian w bazie i dopisywania czy korygowania nowych rekordów. Dane te oczywiście w większości zasilily bazę, ale w okresie późniejszym.

Prace w październiku 2015 r.

Kolejnym wyzwaniem okazało się takie zmodyfikowanie funkcji PBN, aby pokazywała informacje o dacie przekazania rekordu do PBN i ewentualnych modyfikacjach rekordu. Tuż przed zamknięciem drugiego „okienka transferowego” w styczniu 2016 r. otrzymałyśmy informację, jak wgrywać poprawki. Na szczęście modyfikacja algorytmu w funkcji PBN nie była zbyt kłopotliwa, a czasu było wystarczająco dużo na ich dogranie. Obecnie wszystkie rekordy od 2013 r. zamieszczone w bazie otrzymują dodatkową informację o dacie ich przekazania do PBN i ewentualnego zgłoszenia modyfikacji.

Nad projektem pracowało wiele osób. Chciałabym im wszystkim podziękować. Największy wkład miał zespół Sekcji Prac Bibliograficzno-Dokumentacyj-

nych w składzie: Anna Łobodziec, Teresa Haliniak, Wanda Rochowska i Aleksandra Zawadzka. Bez informatyk Barbary Palichleb nie byłoby funkcji PBN, tj. narzędzia interpretującego i pobierającego dane w formacie .xml oraz analizatora danych. Bez importerów z wydziałów przesyłających informacje o kłopotach, błędach, ale i te radosne o udanym imporcie 100, 300, a nawet 1000 rekordów nie byłoby widać efektów pracy. Mając za sobą trudne miesiące pracy koordynatora tego projektu w UW, mam nadzieję, że baza ministerialna w końcu okrzepnie i zamiast trzech systemów będzie jeden, stabilny.

Bibliografia

- Biuletyn Informacji Publicznej Uniwersytetu Wrocławskiego, <http://bip.uni.wroc.pl/183/55/informacje-ogolne.html> [dostęp: 12.02.2017].
- Garczyńska M., *Bibliografie publikacji pracowników państwowych szkół wyższych w Polsce w świetle badań ankietowych*, „Przegląd Biblioteczny” 2003, z. 1.
- Komperda A., *Import publikacji Politechniki Wrocławskiej do Modułu Sprawozdawczego PBN — sprawozdanie z realizacji zadania*, „Biuletyn EBIB” 2016, nr 165, <http://open.ebib.pl/ojs/index.php/ebib/article/view/434> [dostęp: 12.02.2017].
- Kulczycki E., *Lepiej jak jest więcej, czyli Polskie Bibliografie Naukowe*, http://ekulczycki.pl/warsztat_badacza/lepiej-jak-jest-wiecej-czyli-polskie-bibliografie-naukowe/ [dostęp: 12.02.2017].
- Miller-Jankowska M., Lewandowska-Tranda A., *Przekazywanie danych do Systemu Informacji o Nauce poprzez uczelnianą bazę wiedzy. Z doświadczeń Politechniki Warszawskiej*, „Biuletyn EBIB” 2016, nr 165, <http://open.ebib.pl/ojs/index.php/ebib/article/view/435> [dostęp: 12.02.2017].
- Rozkosz E., *Nowy moduł sprawozdawczy Polskiej Bibliografii Naukowej. Część 1*, <http://www.edukacjainformacyjna.pl/2016/08/nowy-modul-sprawozdawczy-polskiej-bibliografii-naukowej-czesc-1/> [dostęp: 12.02.2017].
- Rozkosz E., Drabek A., *Moduł sprawozdawczy Polskiej Bibliografii Naukowej*, „Biuletyn EBIB” 2016, nr 165, <http://open.ebib.pl/ojs/index.php/ebib/article/view/438/596> [dostęp: 12.02.2017].
- Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 27 października 2015 roku w sprawie kryteriów i trybu przyznawania kategorii naukowej jednostkom naukowym.
- Siek B., *Polska Bibliografia Naukowa a modyfikacje indeksu autorskiego bazy Bibliografia GUMed*, „Biuletyn EBIB” 2016, nr 165, <http://open.ebib.pl/ojs/index.php/ebib/article/view/437> [dostęp: 12.02.2017].

Between locality and governmental needs — experiences of outcome registration of Wrocław University Library in 2015

Summary

Legislative changes of Science Funding Law made in 2015 have imposed significant modifications in functioning of the database Bibliography of University of Wrocław Academic Staff Publications (<http://www.bu.uni.wroc.pl/publikacje>). It was founded in 2004 and since 2015 was only auxiliary

tool for searching records to report them with another application to the Polish Ministry of Higher Education. From 2015 we were enforced to be a source of raw data (xml) given to the nationwide application Polish Scholarly Bibliography (Polska Bibliografia Naukowa). The need to adapt big amount of already gathered data (about 10.000 records) to governmental guidelines and to prepare instructions to new data was a big challenge to the team of Section of Bibliographic-Documentation Duties and Computerization Department of Wrocław University Library. The text is a description of solutions used in the first period of cooperation the local database and nationwide one.

KEYWORDS: Bibliography of University of Wrocław Academic Staff's Publications, Polish Scholarly Bibliography.