

Ewa Włostowska, Jacek Głębocki
Główna Biblioteka Lekarska
Splendor Systemy Informacyjne

"Polska Bibliografia Lekarska" Głównej Biblioteki Lekarskiej w roku swojego trzydziestolecia

Abstrakt

W artykule przedstawiono historię „Polskiej Bibliografii Lekarskiej” (PBL), która w postaci komputerowej bazy danych, początkowo w oprogramowaniu własnym Głównej Biblioteki Lekarskiej, na sprzęcie Singer 1500, istnieje od 1979 r. Omówiono pokrótce metodologię opracowywania danych i realizację w PBL zasad indeksowania i wyszukiwania przyjętych w US National Library of Medicine dla systemu MEDLARS/Medline. Jako wybrane zagadnienie przedstawiono procedury związane z prowadzeniem i wykorzystaniem w PBL polskiej wersji amerykańskiego tezausa Medical Subject Headings (MeSH), która pod nazwą TezMeSH funkcjonuje w GBL również od 1979 r. Zaprezentowano wersję internetową bazy PBL oraz nowoczesne metody jej przeszukiwania, wykorzystujące nie tylko różne jednostki leksykalne języka informacyjno-wyszukiwawczego TezMeSH w wersji 2009, ale także jego hierarchiczną strukturę. Omówiono znaczenie bazy dla użytkowników na podstawie oceny jej zawartości przez specjalistów z Cochrane Collaboration oraz przedstawiono możliwości rozbudowy i kierunki rozwoju bazy zmierzające do dostarczenia użytkownikowi nie tylko streszczeń (dołączanych od 1991 r.), lecz także pełnych tekstów publikacji.

Słowa kluczowe

naukowa informacja medyczna, bazy danych, bibliografie medyczne

Abstract

Appreciation of the Polish Medical Bibliography created by the Central Medical Library (GBL) on its Thirtieth Anniversary. The paper offers a comprehensive presentation of the history of the “Polish Medical Bibliography” (PBL). PBL was launched by the Central Medical Library in 1979 and it was originally run on Singer 1500 terminals using proprietary software created specially for the Library. The paper briefly surveys the methodology of data processing and the realisation of the principles of indexing and searching introduced by the US National Library of Medicine for MEDLARS/Medline system in PBL. The paper also presents procedures connected with maintaining and utilising of the Polish version of the American Medical Subject Headings thesaurus (MeSH) in the Polish Medical Bibliography. The Polish version of MeSH, called TezMeSH, functions in the Central Medical Library since 1979. Furthermore, the paper presents the Internet version of the PBL database and the state-of-the-art search methods, that use not only lexical units of the updated 2009 version of the TezMeSH controlled vocabulary thesaurus, but also take into account the hierarchic structure of the thesaurus. The paper also offers an attempt at evaluation of the usefulness of the database for end users, based on the contents evaluation performed by The Cochrane Collaboration experts and finally it presents the perspectives for the future development of the database, and the visions for its expansion. In the future the database should evolve in the direction of supplying the users not only with abstracts (available for texts published since 1991), but also with full text publications.

Keywords

medical information science; database; bibliography of medicine

Komputerowa baza danych „Polska Bibliografia Lekarska” (PBL) jest nowoczesną formą kontynuacji prac nad rejestracją polskiego piśmiennictwa medycznego, rozpoczętych w drugiej połowie lat 40. ubiegłego stulecia przez założyciela Głównej Biblioteki Lekarskiej (GBL), lekarza i bibliografa – prof. dr. med. Stanisława Konopkę. Baza obejmuje polskie piśmiennictwo z zakresu nauk medycznych (podstawowych i klinicznych), nauk pokrewnych

i organizacji ochrony zdrowia. Zawiera dane o wydawnictwach zwartych, artykułach z czasopism, niepublikowanych rozprawach doktorskich, materiałach konferencyjnych i listach. Rejestruje też polonica.

Krótką historia

Już na początku lat 70. ubiegłego wieku rozpoczęła się w GBL automatyzacja procesów biblioteczno-informacyjnych. Zakupiono sprzęt minikomputerowy Singer 1500, przystąpiono do opracowywania programów odpowiadających potrzebom bazy bibliograficznej; nawiązano współpracę z wiodącymi wówczas ośrodkami naukowej informacji medycznej w Europie i Stanach Zjednoczonych, która dała możliwość korzystania z ich doświadczeń.

Podjęto decyzję, iż polska baza bibliograficzna opracowywana będzie przy użyciu języka informacyjno-wyszukiwawczego stosowanego dla amerykańskiej bazy szeroko znanego światowego systemu naukowej informacji medycznej MEDLARS/medline US National Library of Medicine (NLM), Bethesda, Mary Land. W systemie tym do sporządzania charakterystyk treściowych dokumentów stosowano i stosuje się nadal deskryptory tezauryś Medical Subject Headings (MeSH), według ściśle określonych zasad indeksowania. GBL uzyskała zgodę autorów oryginału i przystąpiono do tłumaczenia tezauryś MeSH na język polski. Już w 1978 r. gotowa była baza Tezaurus, pierwsza w Polsce komputerowa baza danych zawierająca polski tezaurus medyczny (polską wersję MeSH) – narzędzie do opracowywania polskiej medycznej bazy bibliograficznej.

Polskie jednostki leksykalne zaczerpnięto z języka fachowego piśmiennictwa z zakresu medycyny i nauk pokrewnych. Słownictwo tezauryś dzieliło się na dwie grupy: zasadniczą i pomocniczą. Na zasadniczą grupę składały się deskryptory, czyli terminy przyjęte do indeksowania (np. układ krążenia) oraz askryptory (terminy równoważne z deskryptorami, ale odrzucone, np. układ sercowo-naczyniowy). Słownictwo pomocnicze tworzyły modyfikatory (terminy wyrażające pewien aspekt, służące do zawężania znaczenia deskryptorów, np. układ krążenia – wpływ środków chemicznych) i tzw. deskryptory obowiązkowe, czyli terminy, których użycie jest obowiązkowe przy analizie każdego dokumentu, ponieważ wskazują na pewne cechy formalne dokumentu (np. że jest to praca doświadczalna) lub niektóre aspekty treściowe (np. że była przeprowadzona na szczurach płci męskiej).

Pierwsza baza Tezaurus składała się z 14 000 rekordów, z których każdy zawierał deskryptor amerykański, deskryptor polski, przynajmniej jeden kod semantyczny, askryptory polskie, łacińskie, rosyjskie i ekwiwalent *Słownika hasel z zakresu medycyny i nauk pokrewnych* służącego do tematowania druków zwartych do katalogu przedmiotowego. Zasady stosowania jednostek leksykalnych tezauryś zawarte były w amerykańskim podręczniku indeksowania – *Medlars indexing manual, 1976*.

W 1979 r. zaczęto opracowywać i wprowadzać dane do bazy PBL. Każdy rekord bazy zawierał opis bibliograficzny dokumentu oraz jego dokładną charakterystykę treściową sporządzaną z wykorzystaniem terminologii polskiego tezauryś medycznego GBL. Łącznie w bazie bibliograficznej w 1979 r. znalazło się 5881 opisów.

W latach 1979–1990 „Polska Bibliografia Lekarska”, podobnie jak wszystkie bazy własne GBL, prowadzona była w unikatowym oprogramowaniu opracowanym w drugiej połowie lat 70. ubiegłego wieku specjalnie dla potrzeb tych konkretnych organizujących się baz komputerowych, z uwzględnieniem ówczesnych możliwości sprzętowych. Po zakupieniu

nowego sprzętu firmy IBM oraz starannym przetestowaniu pakietu CDS/MicroISIS UNESCO, w 1991 r. wszystkie bazy własne GBL zostały przełożone do nowego systemu. Nowe możliwości programowe spowodowały, że rozpoczęto wprowadzanie do bazy PBL odredakcyjnych streszczeń artykułów w językach polskim i angielskim, co znacznie zwiększyło możliwości wdrożonego wówczas programu wyszukiwawczego.

Stan obecny

Baza „Polska Bibliografia Lekarska” opracowywana i prowadzona jest w Dziale Bibliografii Medycznych (poprzednio Zakład Naukowej Informacji Medycznej) GBL od 1979 r. na bieżąco i aktualizowana raz na kwartał. Charakterystyki wyszukiwawcze dokumentów sporządza zespół złożony głównie z osób z wykształceniem medycznym lub biomedycznym, które przeszły wewnętrzne szkolenie w zakresie indeksowania. Podstawową regułą charakteryzowania treści dokumentu jest zasada maksymalnej szczegółowości: główny temat dokumentu musi być wyrażony najbardziej szczegółowymi deskryptorami tezaury. Używana do tego celu terminologia pochodzi z bazy polskiego tezaury medycznego, która obecnie ma nazwę TezMeSH i jest polską wersją amerykańskiego MeSH-a z bieżącego roku.

Baza TezMeSH, prowadzona w systemie Expertus, aktualizowana jest co roku zgodnie z najnowszą wersją bazy źródłowej. Coroczne zmiany dotyczą od kilkuset nawet do kilku tysięcy rekordów: dodawane są nowe terminy powstające wraz z dynamicznym rozwojem zarówno nauk medycznych, jak i pokrewnych, zmieniane lub usuwane są przestarzałe deskryptory. Obecnie baza TezMeSH 2009 zawiera 25 342 terminy źródłowe, tyle samo ich odpowiedników polskich i 14 980 askryptorów. Każdy rekord zawiera ponadto wykaz modyfikatorów dopuszczonych do użycia z danym dekskryptorem, a duża część ma definicję w języku polskim i wskazówki pomocne w procesie indeksowania. Terminologia TezMeSH zgrupowana jest w 16 kategoriach oznaczonych literami:

- A – anatomia,
- B – biologia (organizmy),
- C – choroby,
- D – leki i substancje chemiczne,
- E – techniki analityczne, diagnostyczne i terapeutyczne oraz sprzęt medyczny,
- F – psychologia i psychiatria,
- G – zjawiska i procesy fizjologiczne,
- H – dyscypliny i zawody,
- I – antropologia, edukacja, socjologia,
- J – technologia, przemysł, rolnictwo,
- K – humanistyka,
- L – informatyka,
- M – dane o ludności,
- N – opieka zdrowotna,
- V – typy publikacji,
- Z – nazwy geograficzne.

Każdy deskryptor ma przynajmniej jeden alfanumeryczny kod semantyczny wyznaczający jego miejsce w strukturze tezaury. Przyjęta zasada polihierarchiczności pozwala na występowanie jednego deskryptora w kilku miejscach drzewa deskryptorowego. Baza TezMeSH udostępniana jest z serwera Głównej Biblioteki Lekarskiej bezpłatnie po podpisaniu umowy.

Baza „Polska Bibliografia Lekarska” zawiera blisko 410 000 rekordów. Każdy z nich to informacja o jednym dokumencie zawierająca jego opis bibliograficzny, charakterystykę treściową oraz streszczenia odredakcyjne. W 2008 r. objęto indeksowaniem blisko 350 tytułów czasopism polskich i obcych. Średni roczny przyrost bazy, to 19 000 rekordów. PBL, choć prowadzona nadal w systemie MiroISIS, udostępniana jest w Internecie w systemie Expertus firmy Splendor S.A.

„Polska Bibliografia Lekarska” jest dostępna poprzez przeglądarki WWW z serwera Głównej Bibliografii Lekarskiej. Korzystanie z bazy podlega autoryzacji poprzez kontrolę numeru IP lub hasła. Menu wyszukiwawcze daje liczne i różnorodne możliwości przeszukiwania bazy PBL: od wybranych lub wszystkich pól opisu, poprzez typ dokumentu i szerokie hasło klasyfikacyjne (wskazujące na przynależność dokumentu do konkretnej dyscypliny medycznej, np. kardiologia, pielęgniarstwo) aż do wyszukiwania zaawansowanego z użyciem nie tylko pojedynczych deskryptorów bazy TezMeSH, ale również struktury hierarchicznej tezaury (trele structure), czyli wybranej gałęzi z drzewa semantycznego MeSH.

Wyszukiwanie wspomagane tezaurem jest bardzo cenną opcją wyszukiwawczą. Dostęp do właściwego deskryptora nie wymaga dokładnej znajomości jego postaci. Użytkownik ma zapewniony dostęp do właściwego deskryptora nie tylko poprzez tradycyjne wyszukiwanie, lecz także poprzez askryptory i formy deskryptora używane poprzednio. Możliwe jest także wyszukiwanie według dowolnych słów definicji i adnotacji. Pomocną funkcją jest możliwość obejrzenia położenia deskryptora w strukturze hierarchicznej – drzewie tezaury. Oprócz wyszukiwania prostego, tj. wyszukiwania rekordów sklasyfikowanych danym deskryptorem, można posłużyć się jedną z kilku metod zapewniających poszerzenie zbioru odnalezionych rekordów, tj.:

- wyszukiwaniem rekordów, w których został użyty dany deskryptor, a także inne deskryptory o podanym początku;
- wyszukiwaniem *explode*, tj. wyszukiwaniem rekordów, w których użyto danego deskryptora lub dowolnego deskryptora położonego niżej w danej gałęzi drzewa tezaury, a więc deskryptora bardziej szczegółowego.

Przydatność funkcji *explode* zobrazować można następująco: w wyniku przeszukania bazy PBL z użyciem hasła „choroby serca” uzyskujemy wynik 1244 rekordów. Przeszukanie bazy z użyciem tego samego terminu, ale pobranego z bazy TezMeSH z zastosowaniem funkcji *explode* sprawia, że otrzymujemy wynik 16 879 rekordów. Tak duża różnica liczby wyszukanych rekordów spowodowana jest tym, że w drzewie deskryptorowym tezaury pod terminem „choroby serca” znajdują się aż 153 deskryptory węższe (w tym np.: „zawał serca” i „choroba wieńcowa”), które uruchomione zostały do przeszukiwania dzięki zastosowaniu funkcji *explode*. Wyszukiwanie można także dodatkowo zawęzić tylko do tych rekordów, w których podany deskryptor jest deskryptorem głównym (*focus*), a więc stanowi główny temat dokumentu. Obie operacje, tj. *explode* i *focus*, można stosować łącznie.

Baza PBL cieszy się niezmiennie dużym zainteresowaniem użytkowników. Jest udostępniana w Internecie do samodzielnego przeszukiwania, ale też służy w Głównej Bibliotece Lekarskiej do sporządzania tematycznych zestawień literatury. Usługa ta, choć wykonywana już od ponad 20 lat, nadal jest chętnie wykorzystywana zarówno przez indywidualnych użytkowników, jak i firmy farmaceutyczne monitorujące niepożądane działania leków i materiałów farmaceutycznych.

„Polska Bibliografia Lekarska”, jako największa w Polsce baza medyczna, była obiektem zainteresowania międzynarodowej niezależnej grupy badaczy z Cochrane Collaboration. Jest to międzynarodowa organizacja typu non-profit, która ułatwia podejmowanie trafnych decyzji dotyczących postępowania medycznego poprzez wyszukiwanie w bazach medycznych sprawdzonych faktów i opracowywanie ich zgodnie z zasadami medycyny opartej na dowodach (Evidence Based Medicine).

Baza PBL z lat 1996–2006 została przeszukana pod kątem zawartych w niej danych o prowadzonych w Polsce randomizowanych badaniach klinicznych. Zebrane informacje zostały poddane zunifikowanej analizie, a następnie porównane z zawartością światowych baz medycznych Medline oraz Embase. Okazało się, że blisko 40% wyszukanych w PBL danych nie występuje w powszechnie znanych dużych bazach medycznych. Uzyskany wynik pozwolił autorom oceny na konstatację, iż „Polska Bibliografia Lekarska”, to cenne źródło medycznych faktów i dowodów naukowych nigdzie indziej poza tą bazą nierejestrowanych. Przedstawione wyniki badania świadczą o wysokim stopniu kompletności „Polskiej Bibliografii Lekarskiej”, co pozwala na szerokie upowszechnianie informacji o polskich badaniach naukowych i ich wynikach oraz osiągnięciach w zakresie medycyny i nauk pokrewnych.

Plany na przyszłość

„Polska Bibliografia Lekarska” jest bazą zmieniającą się bezustannie: stałemu unowocześnianiu podlega przede wszystkim interfejs wyszukiwawczy internetowej wersji bazy; systematycznie udoskonalana jest też integracja PBL z bazą TezMeSH. W trosce o aktualność bazy prowadzone są prace zmierzające do usprawnienia opracowywania danych w Głównej Bibliotece Lekarskiej. Służyć temu będzie nowy system rejestracji danych, którego testowanie rozpocznie się w niedalekiej przyszłości. System ten wymaga jednak wcześniejszej zasadniczej reorganizacji pracy w związku z koniecznymi zmianami w obiegu opracowywanych do bazy dokumentów. Przewidywana jest rozbudowa PBL poprzez rozszerzenie systemu prowadzenia bazy o moduł pobierania danych z bibliografii dorobku naukowego pracowników opracowywanych przez współpracujące biblioteki.

Mimo że bardzo cennym elementem PBL są odredakcyjne streszczenia artykułów w językach polskim i angielskim, to twórcy bazy zdają sobie doskonale sprawę z faktu, że oczekiwania użytkowników dotyczą możliwości szybkiego dostępu do pełnych tekstów prac. W związku z tym GBL podjęła rozmowy z wydawcami czasopism zmierzające do umożliwienia dostępu do pełnych tekstów dokumentów rejestrowanych w PBL. Niezależnie prowadzone są prace nad uzupełnieniem opisów bibliograficznych artykułów o linki na serwery wydawców czasopism.

Inną pożyteczną funkcją będzie udostępnienie na stronie internetowej Głównej Biblioteki Lekarskiej systemu przeszukiwania bazy Pubmed/Medline na podstawie Polskiego Tezaurusu Medycznego TezMeSH, co pozwoli na łączne przeszukiwanie amerykańskiej bazy medycznej o światowym zasięgu oraz jej polskiego odpowiednika.

Wszystkie wymienione wyżej zadania planowane i realizowane są w celu udostępniania użytkownikom naukowej informacji medycznej nowych funkcji służących zaspokajaniu ich potrzeb w zakresie szybkiego dostępu do żądanej informacji.

Bibliografia

1. ALMERIE, M.Q. i in. Searching the Polish Medical Bibliography (Polska Bibliografia Lekarska) for trials. *Health Information and Libraries Journal* 2007, Vol. 24, No. 4, s. 283–286. ISSN 1471-1834.
2. DUSIŃSKA, H., KOPYCKA, A., KUŹMIŃSKI, D. (red.) *Polska Bibliografia Lekarska za rok 1979*. Warszawa: Wydaw. Gł. Bibl. Lek., 1984, t. 1: A–G.
3. GŁĘBOCKI, J., PIETRUSZEWSKI, G. Polskie bibliografie medyczne w internecie udostępniane w systemie Expertus. *Biuletyn Głównej Biblioteki Lekarskiej* 2001, R. 47, nr 364, s. 47–52. ISSN 0373-174X.
4. GŁĘBOCKI, J., PIETRUSZEWSKI, G. Zastosowanie systemu bibliograficznego EXPERTUS do udostępniania Polskiej Bibliografii Lekarskiej w Internecie i na CD-ROM. *Biuletyn Głównej Biblioteki Lekarskiej* 2001, R. 47, nr 365, s. 101–110. ISSN 0373-174X.
5. GRODZKA, D. Rozwój koncepcji systemu Naukowej Informacji Medycznej w Głównej Bibliotece Lekarskiej (1971–1995). *Biuletyn Głównej Biblioteki Lekarskiej*, 1996. R. 42, nr 352, s. 21–33. ISSN 0373-174X.
6. GRODZKA, D., MALECKI, J. Problemy warsztatowe Polskiej Bibliografii Medycznej. *Biuletyn Głównej Biblioteki Lekarskiej* 1996, R. 42, nr 352, s. 3–8. ISSN 0373-174X.
7. GRODZKA, D. (red.) *Tezaurus medyczny polsko-angielski: polska wersja MeSH*. Warszawa: Wydaw. Gł. Bibl. Lek., 1995, t. 2, O–Ż, cz. 1: Układ alfabetyczny.
8. KUŹMIŃSKI, D. (red.) *Tezaurus z zakresu medycyny. Wersja polska MeSH*. Edycja 1978. Warszawa : Gł. Bibl. Lek., 1978, cz. 1: Część systematyczna.
9. TYLMAN-GĄDEK, W. Serwisy bibliograficzne Głównej Biblioteki Lekarskiej. Charakterystyka użytkowników i zgłaszanej tematyki. *Biuletyn Głównej Biblioteki Lekarskiej* 1979, R. 28, nr 7, s. 591–608. ISSN 0373-174X.
10. WIDY-WIRSKI, F. (red.) *Polska Bibliografia Lekarska za rok 1974*. Warszawa: Wydaw. Gł. Bibl. Lek., 1980, t. 1: A–O.
11. WŁOSTOWSKA, E., GŁĘBOCKI, J., CHROBAK, E. Nowoczesne metody samodzielnego przeszukiwania bazy danych Polska Bibliografia Lekarska: propozycja konspektu szkolenia użytkowników naukowej informacji medycznej. W: 25. *Jubileuszowa Konferencja Problemowa Bibliotek Medycznych. Kształcenie użytkowników naukowej informacji medycznej – koncepcje i doświadczenia. Lublin–Kazimierz Dolny, 12–14 czerwca 2006 roku*. Warszawa: Stowarzyszenie Bibliotekarzy Polskich, K[omisja] W[ydawnictw] E[lektronicznych], Redakcja „Elektronicznej Biblioteki” [on-line]. 2006. [Dostęp 21 kwietnia 2009]. EBIB Materiały konferencyjne nr 14. Dostępne w World Wide Web: http://www.ebib.info/publikacje/matkonf/25kpbm/wlostowska_glebocki_chrobak.php. ISBN 83-921757-5-1.

"Polska Bibliografia Lekarska" Głównej Biblioteki Lekarskiej w roku swojego trzydziestolecia / Ewa Włostowska, Jacek Głębocki // W: Bibliograficzne bazy danych : kierunki rozwoju i możliwości współpracy. Bydgoszcz, 27-29 maja 2009. - [Warszawa] : Stowarzyszenie Bibliotekarzy Polskich, K[omisja] W[ydawnictw] E[lektronicznych], Redakcja "Elektronicznej Biblioteki", 2009. - (EBIB Materiały konferencyjne nr 19). - ISBN: 83-921757-5-1. -Tryb dostępu : http://www.ebib.info/publikacje/matkonf/mat19/wlostowska_glebocki.php