

Tezaurus

„Bezpieczeństwo pracy i ergonomia” w systemie bibliotecznym ALEPH w Bibliotece CIOP-PIB

Rozbudowa zasobów informacyjnych biblioteki CIOP-PIB stwarza konieczność doskonalenia efektywności procesów wyszukiwawczych oraz kontroli słownictwa w dziedzinie bezpieczeństwa pracy. Aplikacja testowej wersji tezausa w strukturze systemu wspomaga procesy dokumentowania oraz wyszukiwania informacji.

The "Occupational safety and ergonomics" thesaurus in the ALEPH system of the library of CIOP-PIB

An increase in informational resources in the library of CIOP-PIB necessitates an improvement in the efficiency of searching and controlling vocabulary in the area of occupational safety. The use of a test version of the thesaurus supports the processes of content description and searching.

Wstęp

Tezaurus¹ jest słownikiem języka informacyjno-wyszukiwawczego, tzn. języka „sztucznego”, stworzonego specjalnie do celów indeksowania dokumentów (czyli opisu zawartości treści, tematyki dokumentów, np. książek, artykułów z czasopism), a także do wyszukiwania dokumentów w zbiorach. Pozwala on na połączenie w jeden zbiór słownictwa używanego przez autorów, osoby klasyfikujące i użytkowników poszukujących informacji oraz podnosi wskaźnik zbieżności słownictwa używanego przez te osoby. Tezaurus bazuje na słownictwie języka naturalnego, ale nie jest z tym językiem tożsamy – wykorzystuje wyrażenia pochodzące z języka naturalnego, które mogą być zastosowane do opisu i wyszukiwania informacji. Aplikacja tezausa w systemie informacji wspomaga funkcjonalność systemu w zakresie kontroli słownictwa wykorzystywanego w procedurach opisu dokumentów oraz ich wyszukiwania. Możliwości doskonalenia procedur opisu i wyszukiwania z zastosowaniem tezausów są wykorzystywane w wielu międzynarodowych, a także krajowych systemach informacyjnych, bibliotekach, bazach danych, witrynach internetowych i innych zasobach informacyjnych.

Zastosowania tezausów – przykłady

Rozbudowa zasobów informacyjnych oraz ich udostępnianie w postaci coraz obszerniejszych baz danych stwarza konieczność doskonalenia procesów wyszukiwania infor-

¹ Według *Słownika terminologicznego informacji naukowej* tezaurus to zbiór semantycznie i hierarchicznie powiązanych terminów (deskryptorów i askryptorów) z określonej dziedziny wiedzy, stosowany do przekładu języka naturalnego na terminy sformalizowanego języka informacyjnego [1]

macji. Przykładowo, komputerowy system biblioteczny ALEPH, funkcjonujący w Bibliotece CIOP-PIB, zawiera ok. 29 tys. rekordów, opisujących zasoby informacyjne. Biblioteki wielu polskich instytucji, m.in. uczelni, uczestniczą w konsorcjach, zapewniając swoim użytkownikom dostęp do różnych baz danych. Bazy czasopism elektronicznych, udostępnianych ze stanowisk komputerowych Biblioteki CIOP-PIB, umożliwiają dostęp do ponad 16 tys. tytułów czasopism specjalistycznych. Przewidywany ciągły przyrost dokumentów wskazuje na konieczność stosowania narzędzi ułatwiających skuteczne dotarcie do szukanych informacji. Przykładowo, baza czasopism elektronicznych PROQUEST daje rozbudowane możliwości wyszukiwawcze. Znacznym ułatwieniem dla użytkownika są tezaury, pozwalające na precyzyjny dobór słownictwa. Baza wykorzystuje tezaurus *PROQUEST* oraz kilka innych tezausów, np. *PsycINFO* dla czasopism z dziedziny psychologii.

Jako przykład polskiej biblioteki wykorzystującej tezaury można wskazać Bibliotekę Sejmową, której szeroki zakres tematyczny obejmuje praktycznie wszystkie dziedziny działalności człowieka. System Tezausów Biblioteki Sejmowej (STEBIS) obejmuje zbiór tezausów, wyodrębnionych przede wszystkim na podstawie kryteriów tematycznych. Każdy z tezausów zawiera mikrotezaury. Niektóre z nich reprezentują dziedziny nawiązujące tematycznie do szeroko pojmowanych zagadnień ochrony pracy, np. *Mikrotezaurus „Praca”*, zawarty w *Tezaurusie handlu, transportu i pracy*. Z kolei każdy z mikrotezausów może mieć kilka działów. Przykładowo, w mikrotezaurusie *Praca* wyodrębniono takie działy, jak: *Prawo pracy*, *Stosunki pracy*, *Zatrudnienie* i *Status zawodowy*. Tezaurus zagadnień społecznych zawiera kilka mikrotezausów, m.in. *Zdrowie*,

mgr AGNIESZKA MŁODZKA-STYBEL

Centralny Instytut Ochrony Pracy
– Państwowy Instytut Badawczy

Spółczeństwo oraz *Ochrona Środowiska*, które swoim zakresem tematycznym obejmują wiele zagadnień nawiązujących do tematyki bezpieczeństwa pracy. System tezausów Biblioteki Sejmowej jest wykorzystywany do indeksowania dokumentów w bazach zautomatyzowanego katalogu ALEPH – BS. Jest dostępny w sieci Internet na stronie Biblioteki Sejmowej, jako baza wzorcowa w komputerowym systemie bibliotecznym (rys. 1.).

Wszystkie tezaury systemu STEBIS są tezausami wielojęzycznymi, odpowiedniki obcojęzyczne mogą służyć jednocześnie jako tablice przejścia do tezausa Parlamentu Europejskiego EUROVOC, z którym system STEBIS jest w znacznym stopniu kompatybilny. Poszczególne mikrotezaury są także publikowane w postaci wydawnictw książkowych [5].

Zakres słownictwa Tezausa Biblioteki Sejmowej odzwierciedla więc w znacznej mierze polską wersję językową Tezausa EUROVOC.

Prace nad Tezausem „Bezpieczeństwo pracy i ergonomia” (CIOP-PIB)

Podobnie jak w przypadku tezausów Biblioteki Sejmowej, Tezaurus „Bezpieczeństwo pracy i ergonomia”, utrzymywany aktualnie przez CIOP-PIB, został opracowany na podstawie tezausa udostępnionego przez Międzynarodowe Centrum Informacji o Bezpieczeństwie i Higienie Pracy CIS/ILO (*International Occupational Safety and Health Information Centre*) z siedzibą w Genewie [2]. Polski tezaurus z dziedziny bezpieczeństwa pracy jest tezausem dziedzinowym. Opracowany został w 1991 roku w CIOP, jako tłumaczenie i adaptacja Tezausa CIS z roku 1976.

Rozwój słownictwa z dziedziny bezpieczeństwa pracy, a także potrzeby w zakresie szerszego wykorzystywania tezausa w procesach wyszukiwania oraz dokumentowania piśmiennictwa dziedziny wskazywały na konieczność jego aktualizacji.

Tezaurus „Bezpieczeństwo pracy i ergonomia” w wersji po aktualizacji w 2007 roku został rozbudowany o ponad 5 tys. deskryptorów w języku polskim, uzupełnionych na podstawie aktualizacji Tezaurusu CIS oraz zgłoszeń potrzeb związanych z terminologią. Wprowadzono 3 dodatkowe wersje językowe: angielską, francuską i hiszpańską (ponad 12 tys. deskryptorów w każdym z języków). Tak więc Tezaurus „Bezpieczeństwo pracy i ergonomia” jest tezaurusem wielojęzycznym, z możliwością wprowadzenia kolejnych wersji językowych. Dodatkowo rozbudowano zbiór askryptorów i relacji. Cały Tezaurus jest podzielony na 20 działów odzwierciedlających tematykę bezpieczeństwa pracy:

- Bezpieczeństwo i higiena pracy – kategorie ogólne
- Fizyka, chemia i zjawiska przyrody
- Chemia – związki nieorganiczne
- Chemia – związki organiczne
- Enzymy
- Substancje i produkty naturalne
- Produkty przemysłowe
- Sprzęt, stanowiska pracy i czynności
- Praca i organizacja przemysłu
- Pożary, wybuchy
- Patologia zawodowa
- Fizjologia, anatomia, psychologia
- Pomiar i badania
- Profilaktyka medyczna i leczenie
- Technika bezpieczeństwa i higieny pracy
- Środki ochrony indywidualnej
- Organizacja ochrony pracy
- Kategorie pracowników
- Dziedziny gospodarki
- Zagadnienia ogólne

Deskryptory w każdym dziale uporządkowane są hierarchicznie. Każdemu deskryptorowi przyporządkowany jest kod literowy, wskazujący miejsce w hierarchii. Na przykład deskryptor *Wybuch pyłu węglowego* (kod literowy: Lefo) zlokalizowany jest w dziale „Pożary i wybuchy” [L] na czwartym poziomie hierarchii:

Pożary i wybuchy (L) > Wybuch (Le) > Wybuch pyłu (Lef) > Wybuch pyłu węglowego (Lefo).

Baza Tezaurus, poza hierarchicznym zestawieniem deskryptorów, ma także inne cechy „naprowadzające” czytelnika na określone słownictwo. Wyświetlając rekord deskryptora uzyskujemy informacje o terminach położonych wyżej lub poniżej w hierarchii, a więc szerszych lub węższych znaczeniowo (BT, NT), a także w przypadku niektórych deskryptorów – o terminach związanych znaczeniowo z danym



Rys. 1. Tezaurus Biblioteki Sejmowej na stronie internetowej biblioteki, w komputerowym systemie bibliotecznym ALEPH-BS

Fig. 1. The Sejm Library's thesaurus on the library website, as a part of the ALEPH library computer system

deskryptorem (RT). Do wielu deskryptorów dołączone są wyrażenia askryptorowe (UF), będące często synonimami. System „naprowadza” czytelnika (lub bibliotekarza), który zastosował wyrażenie będące askryptorem, na odpowiadający temu askryptorowi deskryptor. Na przykład: rekord deskryptora *młode osoby* (dział – *Kategorie pracowników*[W]) zawiera 2 askryptory (UF – *młodzież, młodzi pracownicy*), termin związany (RT – *limit wieku*), termin szerszy (BT – *kobiety, dzieci i osoby młode*), a także trzy dodatkowe wersje językowe deskryptora:

Termin Faseta TT	<ul style="list-style-type: none"> • młode osoby • Wem • [W] - KATEGORIE PRACOWNIKÓW
UF	<ul style="list-style-type: none"> • młodzi pracownicy • młodzież
BT	<ul style="list-style-type: none"> • kobiety, dzieci i osoby młode
RT	<ul style="list-style-type: none"> • limit wieku
FT (Eng)	<ul style="list-style-type: none"> • young persons
FT (Fre)	<ul style="list-style-type: none"> • adolescents
FT (Spa)	<ul style="list-style-type: none"> • jóvenes
Sys. No.	000012260

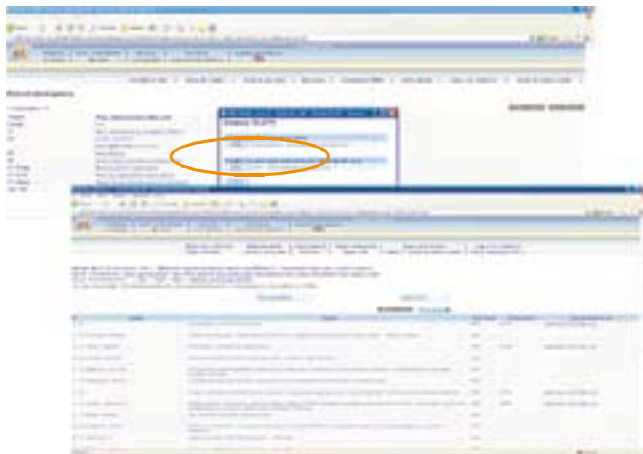
Wpisanie askryptora podczas wyszukiwania powoduje automatyczne „przełączenie” na odpowiedni deskryptor, np. wpisując jako wyrażenia wyszukiwawcze wyraz *młodzież* (askryptor) uzyskamy zestawienie wyników wyszukiwania zawierające publikacje, w których opisano zastosowanie deskryptora *młode osoby* (deskryptor). Będą to udostępniane w bibliotece pozycje piśmiennictwa, dotyczące problematyki ludzi młodych, a w szczególności młodych pracowników.

Tak więc baza Tezaurus, udostępniona w systemie, daje możliwość bezpośredniego jej wykorzystywania w procedurach opisu dokumentacyjnego (przez bibliotekarzy) oraz w procedurach wyszukiwania (przez czytelników) – dzięki powiązaniom z bazami bibliograficznymi systemu, odzwierciedlającymi zasoby Biblioteki CIOP-PIB.

Wyszukiwanie w zasobach informacyjnych Biblioteki CIOP-PIB z wykorzystaniem tezaurusu

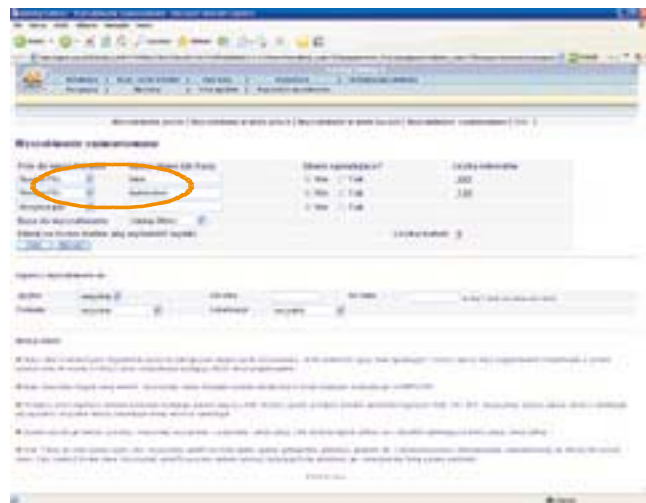
Biblioteka CIOP-PIB (<http://www.ciop.pl/395.html>) jest biblioteką specjalistyczną, dysponującą bogatym księgozbiorem z dziedziny bezpieczeństwa i ochrony zdrowia człowieka w środowisku pracy. Zasoby informacyjne biblioteki reprezentują niemal wszystkie dziedziny nauki, zarówno humanistyczne, jak i techniczne, a także obszerne zbiory o charakterze uniwersalnym.

System ALEPH jest wyposażony w standardowe funkcje wyszukiwawcze, a dodatkową możliwość daje zastosowanie Tezaurusu. Przeglądanie bazy możliwe jest zarówno w wersji alfabetycznej, jak i hierarchicznej, odzwierciedlającej strukturę Tezaurusu. Wersja alfabetyczna zawiera listę deskryptorów uporządkowanych alfabetycznie, np. po wpisaniu frazy „praca” otrzymujemy fragment listy: *praca a choroby, praca biurowa, praca chałupnicza* itd. z możliwością przejścia do kolejnych stron listy. Przeglądanie bazy w wersji hierarchicznej (systematycznej) umożliwia zapoznanie się ze zbiorem deskryptorów uporządkowanych hierarchicznie.



Rys. 2. Przejście od Bazy Tezaurus do baz bibliograficznych odzwierciedlających zasób biblioteki oraz przykładowa lista publikacji związanych z wybranym deskryptorem (*Rola organizacji pracowniczych*)

Fig. 2. The connection between the Thesaurus database and bibliographic databases reflecting the library's resources and a sample list of publications related to a selected descriptor (*"Role of workers' organizations"*)



Rys. 3. Przykład wyszukiwania dla dwóch wyrażen z bazy Tezaurus (*hałas i budownictwo*)

Fig. 3. A sample search for two expressions from the Thesaurus database (*noise and construction industry*)

Po przeglądzie i wyborze słownictwa z bazy Tezaurus można wykorzystać je we wszystkich dostępnych opcjach wyszukiwania. System pozwala także na bezpośrednie przejście z bazy Tezaurus, po zaznaczeniu wybranego wyrażenia, do list publikacji dotyczących danego tematu. Na przykład: szukamy publikacji dotyczących organizacji pracowniczych – przeglądamy bazę Tezaurus (wersja alfabetyczna lub hierarchiczna), wybieramy deskryptor *Rola organizacji pracowniczych* z działu *Organizacja ochrony pracy*. Przechodzimy do baz bibliograficznych i otrzymujemy wyniki wyszukiwania publikacji, którym przypisano wybrany deskryptor – jako odzwierciedlający treść publikacji (rys. 2.).

Deskryptory, wybierane w wyniku przeglądania bazy Tezaurus możemy także wykorzystywać we wszystkich udostępnianych przez system procedurach wyszukiwawczych, m.in. w wyszukiwaniach złożonych z zastosowaniem operatorów logicznych, np. dwa wyrażenia z Tezaurus: *hałas i budownictwo* (rys. 3.).

Wybór pól daje także możliwość wykorzystania wybranego wyrażenia z Tezaurus w połączeniu z innymi danymi, np. wyborem określonego autora publikacji (pole *Autor*), frazy z tytułu itd.

W sytuacji, gdy przeglądanie bazy Tezaurus nie prowadzi bezpośrednio do wyboru deskryptorów, np. szukamy publikacji na temat wieloaspektowy i trudno jest wybrać odpowiedni (dla tego tematu) dział Tezaurus do przeszukiwania, wygodnie jest w takiej sytuacji zastosować funkcję maskowania – wykorzystujemy znak zapytania („?”) dodany na końcu wyrażenia, np. interesują nas zagadnienia społeczne związane z pracą, wprowadzamy wyrażenie wyszukiwawcze *społ?*. W takim przypadku konieczna jest jednak staranna selekcja wyników, ponieważ tylko niektóre wyniki są zgodne z naszymi oczekiwaniami, np. spośród kilku proponowanych deskryptorów (*aspekt społeczny, służby pomocy*

społecznej, grupa środowiskowa, klimat społeczny, światło spolaryzowane, Międzynarodowe Stowarzyszenie Zabezpieczenia Społecznego), deskryptor *światło spolaryzowane* powinien być pominięty przy tworzeniu zestawu wyników jako nie odzwierciedlający danej tematyki.

Więcej informacji na temat wyszukiwania w systemie, m.in. z wykorzystaniem Tezaurus *„Bezpieczeństwo pracy i ergonomia”* zawiera instrukcja dla użytkowników, udostępniona w systemie ALEPH [3].

Podsumowanie

Obserwowany w ostatnich latach wzrost zainteresowania tezaurusami wynika m.in. z intensywnego rozwoju zasobów informacyjnych oraz związanego z tym faktem, coraz częściej obserwowanego, braku satysfakcji z wyników wyszukiwania tekstowego. Konieczność doskonalenia efektywności procesów wyszukiwawczych oraz kontroli słownictwa, m.in. przez aplikację tezaurusu w systemie, dostrzegana jest przez wiele instytucji i organizacji utrzymujących i rozwijających systemy informacyjne. Tezaurus „spina” procedury opisu dokumentacyjnego oraz wyszukiwania. Jego rola jest szczególnie istotna w przypadku potrzeby znacznej precyzji wyszukiwania. Pośrednio, tezaurus pełni także rolę słownika pojęciowego, który pomaga zidentyfikować terminy reprezentujące dany temat. Prace nad tezaurusami prowadzi wiele instytucji w kraju i na świecie. Technologie informacyjne umożliwiają rozwój zarówno baz danych stosowanych do utrzymywania tezaurusów dostępnych w sieci Internet, na płytach

CD oraz innych nośnikach danych, jak i ich modułów w systemach obsługujących biblioteki. Współpraca międzynarodowa stymuluje rozwój tezaurusów wielojęzycznych, stanowiących istotny element europejskich i globalnych systemów informacyjnych.

Aktualizacja Tezaurusu „Bezpieczeństwo pracy i ergonomia” (CIOP-PIB) i jego aplikacja w bibliotecznym systemie komputerowym stanowi ważny krok w rozwoju systemu. Tezaurus, funkcjonujący w systemie, wspomaga użytkownika przy wyszukiwaniu dokumentów, wskazuje, jakie terminy wyszukiwawcze – oprócz lub zamiast terminu już znalezionego – może wykorzystać dodatkowo, do poszerzenia lub zawężenia pytania. Pośrednio, może pełnić także rolę bazy zasobów terminologicznych z dziedziny bezpieczeństwa pracy, udostępnianej w czterech wersjach językowych.

PIŚMIENNICTWO

- [1] *Słownik terminologiczny informacji naukowej*. Pod red. M. Dembowskiej. PAN, Wydawnictwo Ossolineum, Wrocław 1979
- [2] CIS Worldwide Networking, International Occupational Safety and Health Information Centre, Geneva, 2005, <http://www.ilo.org/public/english/protection/safework/cis/intro/>, (dokument elektroniczny, dostęp 10.12.07)
- [3] *Instrukcja wyszukiwania w bazie CIOP-PIB* http://aleph.ciop.pl/exlibris/aleph/u16_1/alephe/www_f_pol/icon/instr_wyszukiwania.pdf, (dokument elektroniczny, dostęp 10.12.07)
- [4] A. Młodzka-Stybel *Tezaurusy z dziedziny bezpieczeństwa pracy*, „Bezpieczeństwo Pracy” 4(415) 2006
- [5] *Praca. Mikrotezaurus*. Oprac. Ewa Chmielewska-Gorczyca, Biblioteka Sejmowa, *Tezaurus handlu, transportu i pracy*. Wydawnictwo Sejmowe, Warszawa 2002

Publikacja przygotowana na podstawie wyników uzyskanych w ramach programu wieloletniego pn. „Dostosowywanie warunków pracy w Polsce do standardów Unii Europejskiej” dofinansowanego w latach 2005 - 2007 w zakresie zadań służb państwowych przez Ministerstwo Pracy i Polityki Społecznej. Główny koordynator: Centralny Instytut Ochrony Pracy – Państwowy Instytut Badawczy