

Blok tematyczny: Bibliografia językoznawstwa sławistycznego w przestrzeni cyfrowej (XVI Kongres Sławistów, Belgrad, 20-27.08.2018)

Zofia Rudnik-Karwatowa, Jakub Banasiak, Marcin Fastyn, Paweł Kowalski (Instytut Sławistyki PAN, Warszawa, Polska) **Zenon Mikos** (Biblioteka Sejmowa, Warszawa, Polska), *Metody i narzędzia optymalizacji systemu informacji bibliograficznej z zakresu językoznawstwa sławistycznego iSybislaw*

Streszczenie referatu wprowadzającego

Celem bloku Komisji Bibliografii Lingwistycznej jest, ogólnie mówiąc, przedyskutowanie najważniejszych problemów dotyczących nowoczesnej postaci bibliografii w środowisku sieciowym. Chodzi o podzielenie się doświadczeniami i wymianę opinii na temat znaczenia bibliografii cyfrowej dla językoznawczych badań naukowych oraz sposobów jej optymalizacji. Recesja tradycyjnych papierowych bibliografii otwiera perspektywy, które były dotąd niedostępne (lub wręcz niemożliwe do osiągnięcia) dla twórców nowoczesnych bibliografii. Wykorzystanie nowych technologii i metodologii umożliwiło powstawanie nowoczesnych systemów informacji bibliograficznej, stanowiących cenne narzędzie w pracy badawczej sławistów. Odpowiednie zarządzanie bardzo dużymi zasobami danych i informacji stwarza nieznane dotychczas możliwości analiz statystycznych, różnorodnych badań dotyczących zmian językowych (zmian w językach słowiańskich), a także ich prognozowania i in.

Zakładamy, że referaty zgłoszone do bloku tematycznego będą poświęcone zarówno kwestiom metodologicznym, jak i praktycznym, czyli skoncentrują się **na metodach tworzenia i udoskonalania** bibliograficznego systemu informacyjno-wyszukiwawczego, a także **efektywnego korzystania** z istniejących językoznawczych baz bibliograficznych, w tym bazy światowego językoznawstwa sławistycznego iSybislaw.

Referat wprowadzający przedstawia przyjęte w systemie **iSybislaw rozwiązania merytoryczne, a także techniczne i technologiczne, optymalizujące funkcjonalności tego systemu**. Gwoli ścisłości, skupimy się na nowych rozwiązaniach zastosowanych w ciągu 10 lat funkcjonowania systemu, jednak dla pełności obrazu przypomnimy też najważniejsze cechy innowacyjne zaimplementowanego w końcu 2007 r. systemu informacyjno-wyszukiwawczego.

Już na wstępie trzeba podkreślić, że efektywność systemu iSybislaw i jego optymalizacja są możliwe dzięki hipertekstowości, stanowiącej konstytutywną cechę nowoczesnych systemów informacyjno-wyszukiwawczych. Przykładowo, w systemie iSybislaw opartym na hipertekście wykorzystywane jest narzędzie intertekstualności, pozwalające m.in. dołączać, poprzez specjalne hiperłącza, pełne teksty dokumentów do ich opisów bibliograficznych. W konsekwencji baza iSybislaw zmienia swój charakter z bazy bibliograficznej na bibliograficzno-pełnotekstową. W ciągu

niespełna 4 lat wprowadzono do zbioru informacyjnego systemu ok. 1000 pełnych tekstów różnego typu dokumentów, przeważnie artykułów, a także recenzji, całych tomów czasopism, monografii.

Zdecydowana większość rozwiązań optymalizacyjnych systemu iw., który ma realizować oczekiwania i potrzeby informacyjne użytkowników, dotyczy polepszenia jakości języka informacyjno-wyszukiwawczego. W referacie przedstawimy najważniejsze problemy z tym związane:

- Język słów kluczowych jako podstawowy język iw. systemu iSybisław.
- Przemodelowanie języka klasyfikacyjnego, tak by stanowił skuteczniejsze narzędzie informacyjno-wyszukiwawcze odzwierciedlające najnowsze tendencje w językoznawstwie sławistycznym.

- Metoda indeksowania intelektualnego (vs. indeksowanie automatyczne).
- Odpowiednia szerokość i głębokość (szczegółowość) indeksowania zapewniająca odwzorowanie relewantnych dla użytkownika informacji zawartych w dokumentach. Wiąże się to z koniecznością wyznaczenia, opracowania i udostępnienia wszystkim indeksatorom jasnych kryteriów indeksowania za pomocą słów kluczowych i klasyfikacji odnoszących się do dokumentów z różnych subdyscyplin językoznawczych. Widzimy więc pilną potrzebę opracowania takiej instrukcji indeksowania dokumentów. Instrukcja taka mogłaby być wykorzystana także przez użytkowników. Punkt ten wiąże się również z koniecznością szczegółowych badań potrzeb użytkowników systemu, o czym dalej w tekście.

- Kontrola słów kluczowych (język słów kluczowych o słownictwie kontrolowanym). Uwzględnienie możliwości indeksowania swobodnego (tagowania) jako pomocniczego narzędzia dla indeksatorów.

- Rozwiązanie problemu językowego synonimii i bliskoźnaczności.
- Rozwiązanie problemu wieloznaczności.
- Ustanowienie klas ekwiwalencji międzyjęzykowych.
- Metodyka sporządzania abstraktów (streszczeń) i adnotacji (uwzględniająca potrzeby językowe i uwarunkowania/postęp technologiczny).

Przy współczesnych wymogach wydawniczych tworzenia do każdego tekstu naukowego abstraktów i streszczeń należy przyznać pierwszeństwo wykorzystywaniu i dołączaniu do rekordów poprzez linki udostępnianych abstraktów i streszczeń (rozwój hipertekstowości systemu).

Oddzielnie przedstawione będą aspekty pragmatyczne dotyczące zaspokojenia potrzeb informacyjnych z punktu widzenia użytkowników (relewancja pragmatyczna) - planowane jest przeprowadzenie badań ankietowych. Rozważane jest także zbadanie możliwości zaimplementowania bardziej zaawansowanych statystyk użytkownika strony, jak np. najczęściej odwiedzane artykuły, najczęściej wyszukiwane słowa kluczowe, mapa zachowań użytkowników. Chodzi o ustalenie istoty potrzeb użytkowników i uwzględnienie ich w charakterystykach wyszukiwawczych i instrukcjach wyszukiwawczych. I tutaj jakość języka

informacyjno-wyszukiwawczego musi odpowiadać wymaganiom i potrzebom użytkowników. Dlatego nieustannie prowadzone są prace badawcze nad zwiększeniem sprawności tego narzędzia informacyjnego systemu (por. koncepcja banku terminologii językoznawczej).

Coraz wyraźniejsze jest współdziałanie i integracja rozmaitych metod i narzędzi stosowanych w systemie w celu zapewnienia użytkownikom znalezienia relewantnej informacji i poszerzenia wiedzy niezbędnej do rozwiązania badanych przez nich problemów. Być może jest to tendencja, która wyznacza kierunek rozwoju (bibliograficznych) systemów informacyjno-wyszukiwawczych.

Obok wymienionej powyżej tendencji wewnętrznej do współdziałania oraz integracji metod i narzędzi wewnątrz systemu, zarysowuje się także tendencja do integracji zewnętrznej: włączanie do systemu informacji pochodzących z różnych zasobów sieciowych. Wiąże się to z rozwijaniem hipertekstowości w obrębie systemu, który stanie się jednym z elementów wieloelementowej sławistycznej infrastruktury badawczej.