

I. ROZPRAWY, BADANIA, MATERIAŁY

JĘZYKI INFORMACYJNO-WYSZUKIWAWCZE WCZORAJ, DZIŚ... CZY JUTRO?

Bożenna Bojar
Uniwersytet Warszawski

*Informacja naukowa, języki informacyjno-wyszukiwawcze,
system informacyjno-wyszukiwawczy*

Na dzisiejsze spotkanie z Państwem¹ zostałam zaproszona przez organizatorów dlatego, że głoszone przeze mnie, na dwóch zorganizowanych w Zakopanem przez Polskie Towarzystwo Informatyki konferencjach, poglądy na temat przyszłości informacji naukowej w ogóle, a języków informacyjno-wyszukiwawczych w szczególności, wywołały ożywioną dyskusję. Dyskusja była przede wszystkim kulturalowa, bo z publikowanymi głosami dyskusyjnymi się nie spotkałam. Poruszony przeze mnie problem był widocznie istotny, gdyż do postawionej przeze mnie – i, mam nadzieję, należycie uargumentowanej – tezy, że tradycyjnie rozumiana informacja naukowa znajduje się na końcu drogi swojego rozwoju, wracano jeszcze po dwóch latach, kiedy to w następnym referacie odniosłam się do problemu przyszłości, a właściwie już tylko przeszłości języków informacyjno-wyszukiwawczych.

Zdaję sobie sprawę z tego, że tak ostro sformułowana teza, w środowisku ludzi zajmujących się informacją naukową i językami informacyjno-wyszukiwawczymi, mogła budzić sprzeciw, ale przecież i ja swoje życie naukowe w połowie związałam z tą dziedziną (druga połowa to lingwistyka teoretyczna). Przez wiele lat uczyłam w Instytucie Informatyki i Studiów Bibliologicznych (dawniej: Instytut Bibliotekoznawstwa i Informatyki) Uniwersytetu Warszawskiego lingwistycznych podstaw informacji naukowej, prowadziłam w Katedrze Lingwistyki Formalnej UW seminarium poświęcone językom informacyjno-wyszuki-

¹ Skróty tego tekstu wygłoszony został w Bibliotece Narodowej, 3 czerwca 2009 r., na XI Ogólnopolskich warsztatach zorganizowanych przez Komisję Opracowania Rzeczowego zbiorów przy ZG Stowarzyszenia Bibliotekarzy Polskich, na sesji: *Języki informacyjno-wyszukiwawcze – teraźniejszość i przyszłość* i stanowi powtórzenie tez przedstawionych przeze mnie we wcześniejszych wystąpieniach.

wawczym, na którym powstało sporo prac doktorskich i habilitacyjnych z informacji naukowej, współpracowałam z różnymi instytucjami zajmującymi się informacją naukową, byłam i jestem redaktorem jedynego polskiego czasopisma poświęconego tej problematyce – Zagadnień Informacji Naukowej. A teraz wystąpiłam w roli Kasandry, wieszczącej koniec informacji naukowej i języków informacyjno-wyszukiwawczych – mam jednak nadzieję, że nie podzielę jej losu i mnie Państwo uwierzą, a właściwie podzielą moje zdanie, do czego postaram się przekonać Państwa rzetelną argumentacją.

Ze sformułowanego, przed prawie już czterema laty, zdania się nie wycofuje, poglądów w tej sprawie nie zmieniłam, mogę więc tylko je jeszcze raz powtórzyć i przytoczyć poprzednio już podane argumenty, bo nic nowego właściwie nie mam w tej sprawie do powiedzenia. Czynię to na wyraźną prośbę organizatorów, bo podobno nie zapomniano mi tego, co wtedy powiedziałam, ale co dokładnie powiedziałam, słyszało niewiele osób (o końcu informacji naukowej mówiłam na sesji poświęconej pamięci profesora Eugeniusza Ścibora), a ani materiały z zakopiańskiej konferencji, ani księga pamiątkowa dedykowana profesorowi Ściborowi dotąd się nie ukazały drukiem. Odpowiadając na zgłaszane zapotrzebowanie środowiska, zebrałam wprawdzie moje poglądy w artykule *Informacja naukowa – czy to już koniec?*, opublikowanym w Biuletynie Polskiego Towarzystwa Informacji Naukowej, ale jakoś nikt na to nie zwrócił uwagi. Tak więc, jeszcze raz, ab ovo.

Odpowiedź na pytanie o przyszłość języków informacyjno-wyszukiwawczych wymaga przede wszystkim odpowiedzi na pytanie, czym są języki informacyjno-wyszukiwawcze, choćby po to, żeby wykluczyć używanie w dyskusji argumentów dotyczących innego rozumienia przedmiotu dyskusji. Argumenty na uzasadnienie własnej tezy wymagają z kolei pokazania:

- jak funkcjonowały języki w różnego typu systemach informacyjnych,
- jak się zmieniały w zależności od rozwijających się potrzeb informacyjnych użytkowników informacji,
- jaka refleksja teoretyczna tym zmianom towarzyszyła.

Czym są języki informacyjno-wyszukiwawcze – problemy definicyjne

Nikogo chyba nie trzeba przekonywać, że przez wiele lat języki informacyjno-wyszukiwawcze znajdowały się w centrum zainteresowań badawczych informacji naukowej. Działo się tak dlatego, iż informacja naukowa zarówno w swej części praktycznej, jak i teoretycznej zajmuje się systemami informacyjno-wyszukiwawczymi, każdy system informacyjno-wyszukiwawczy, aby mógł spełniać swą rolę wymaga narzędzia. Takim „intelektualnym” narzędziem, służącym do reprezentacji obiektów znajdujących się w systemie (dokumentów) jest język informacyjno-wyszukiwawczy, a narzędziem fizycznym, dokonującym operacji w systemie jest obsługujący go człowiek lub maszyna.

Sam termin *język informacyjno-wyszukiwawczy* pojawił się w użyciu wraz z wyodrębnieniem się informacji naukowej jako dziedziny wiedzy obejmującej refleksją teoretyczną szeroko rozumiane systemy informacyjno-wyszukiwaw-

cze, a więc dopiero w drugiej połowie XX w. Wcześniej, gdy systemami informacyjnymi były przede wszystkim biblioteki, nie mówiono ani o informacji naukowej, ani o systemach informacyjno-wyszukiwawczych, ani o językach informacyjno-wyszukiwawczych – były biblioteki, katalogi, bibliografie, było księgoznawstwo (w szerokim znaczeniu terminu tego używał na przykład Świderski), bibliologia i bibliotekoznawstwo, były różnego rodzaju klasyfikacje biblioteczne i biblioteczno-bibliologiczne. Był także język haseł przedmiotowych – to chyba pierwszy raz użyto tu terminu, *język* – ale nie nazywano go wtedy językiem informacyjno-wyszukiwawczym. Dopiero wtedy, gdy w odpowiedzi na zapotrzebowanie na coraz bardziej wyspecjalizowaną informację, obok tradycyjnych bibliotek zaczęły funkcjonować specjalistyczne systemy informacyjne, w których ważne stały się takie parametry wyszukanej informacji jak kompletność i relewancja, coraz więcej uwagi zaczęto poświęcać mogącym sprostać tym nowym wymaganiom zbiorom wyszukiwawczym tworzonym przez informację pochodną, a więc i umożliwiającym zapewnienie tej informacji odpowiedniej jakości językom. Wtedy to zaczęto mówić o informacji naukowej i o językach informacyjno-wyszukiwawczych.

Języki informacyjno-wyszukiwawcze definiowane były przede wszystkim z uwzględnieniem ich funkcji w systemie informacyjno-wyszukiwawczym – jako przeznaczone do odtwarzania podstawowej treści dokumentu i (lub) kwerendy w celu wyszukania ze zbioru informacyjnego tylko tych dokumentów, które odpowiadają na kwerendę, tzn. których charakterystyka wyszukiwawcza pokrywa się z instrukcją wyszukiwawczą, ale też z uwzględnieniem sposobu ich powstawania, a mianowicie określane były jako języki sztuczne (STIN).

Tak skonstruowana definicja była na tyle pojemna, że wyznaczana przez jej konotację denotacja obejmowała swym zakresem wszystkie dotąd funkcjonujące klasyfikacje, języki używane w katalogach, bibliografiach i wszelkich innych zbiorach informacji dokumentacyjnej – wszystkie one pełniły określoną definicją funkcję i wszystkie były specjalnie dla pełnienia tej funkcji tworzone, były więc – w opozycji do języków naturalnych, w których zapisana była treść opisywanych przez nie dokumentów – językami sztucznymi. Warunki definicyjne spełniały nawet te klasyfikacje, które służyły do ustawiania książek na półkach bibliotecznych i ich odszukiwania, i tam bowiem odwzorowana była w jakimś stopniu ich treść (choćby najogólniej rozumiany temat książki – bezpośrednio, lub przynależność do określonej dziedziny wiedzy – pośrednio), była też i kwerenda, umieszczona choćby w pamięci poszukującego.

Dla pełnej jasności wyводу odnieść się należy również do tego, jak w wypadku języków informacyjno-wyszukiwawczych pojmowana jest cecha sztuczności. W odniesieniu do języków informacyjno-wyszukiwawczych cecha ta przypisywana jest nie tylko wszystkim językom informacyjno-wyszukiwawczym – jako kryterium odróżniające je od języków naturalnych, ale też włączające je do szerszej klasy języków sztucznych (sztucznych systemów semiotycznych), czyli takich, w wypadku których konstituowanie systemu poprzedza użycie, a więc realizowanie się w konkretnych aktach komunikacyjnych. Cecha ta służy również jako kryterium wewnątrzsystemowe, do podziału języków informacyjno-wyszukiwawczych na tzw. języki sztuczne i tzw. języki paranaturalne, a ściślej: na języki o notacji sztucznej i języki o notacji paranaturalnej. W tych pierwszych albo plan

wyrażania generowany jest przez zbiór wyrażen elementarnych o formie innej niż przynależna językowi naturalnemu – mogą to być na przykład cyfry rzymskie lub arabskie lub inne dowolne znaki, albo konstytuowany jest przez wyrażenia o formie przynależnej jakiemuś językowi naturalnemu, ale desygnujące inny niż w języku naturalnym plan treści (inne znaczenie) – mogą to być na przykład wyrażenia składające się z liter jakiegoś alfabetu używanego w jakimś języku naturalnym, ale tworzące wyrażenia w nim niewystępujące. W tych drugich słownik stanowi podzbiór wyrażen języka naturalnego: identycznej formie przypisane jest znaczenie skorelowane ze znaczeniem odpowiedniego wyrażenia w języku naturalnym. Do języków informacyjno-wyszukiwawczych nazywanych sztucznymi ze względu na notację należy większość klasyfikacji, do języków paranaturalnych języki słów kluczowych, języki haseł przedmiotowych i języki deskryptorowe. Wszystkie te języki są jednak, zgodnie z definicją, językami sztucznymi, specjalnie konstruowanymi dla potrzeb bibliotek lub innych systemów informacyjnych, mają właściwą im gramatykę i semantykę określoną na zbiorze dokumentów, do odwzorowywania których służą.

Cecha sztuczności pozostała cechą definicyjną języków informacyjno-wyszukiwawczych do dziś – w najnowszym kompendium wiedzy o informacji naukowej, opublikowanym przez Wydawnictwo Stowarzyszenia Bibliotekarzy Polskich w 2006 r. podręczniku *Informacja naukowa. Rozwój – metody – organizacja* przygotowanym pod redakcją Zbigniewa Żmigrodzkiego oraz Wiesława Babika i Diany Pietruch-Reizes, w rozdziale *Języki informacyjno-wyszukiwawcze – charakterystyka funkcjonalno-strukturalna i metodologiczna* czytamy, iż „Języki informacyjno-wyszukiwawcze (jiw) są sztucznymi systemami semiotycznymi, stworzonymi w celu zapewnienia realizacji określonego typu procesów informacyjnych. Są to wyspecjalizowane sztuczne systemy językowe przeznaczone do opisu treści i podstawowych cech formalnych dokumentów” (IN, s.191) i dalej: „Językami informacyjno-wyszukiwawczymi lub językami informacyjnymi we współczesnej literaturze z zakresu informacji naukowej nazywa się wyspecjalizowane sztuczne systemy językowe tworzone jako narzędzia opisu i wyszukiwania treści oraz podstawowych cech formalnych informacji przedstawionej w dokumentach oraz pytaniach użytkowników” (IN, s.194) oraz „Jiw – jako języki formalne, których konstrukcja poprzedza użycie, wyspecjalizowały się w realizacji tylko pewnych określonych funkcji, a mianowicie funkcji metainformacyjnej, polegającej na reprezentowaniu treści i formy dokumentów, oraz funkcji wyszukiwawczej, polegającej na wyrażaniu kierowanych do systemu informacyjno-wyszukiwawczego pytań informacyjnych i kształtowaniu procesu wyszukiwania informacji” (IN, s.194). Podobna definicja, podkreślająca sztuczny charakter języka informacyjno-wyszukiwawczego znalazła się w przygotowanym pod moją redakcją *Słowniku encyklopedycznym informacji, języków i systemów informacyjno-wyszukiwawczych* (SBP 2002), która mówi, że jest to „język sztuczny, którego wyspecjalizowaną funkcją jest funkcja metainformacyjna polegająca na odwzorowywaniu cech informacji i/lub cech nośników informacji, na których są one utrwalone, przez tworzenie ich reprezentacji zwanej charakterystyką wyszukiwawczą dokumentu oraz funkcja wyszukiwawcza polegająca na wyrażaniu zapytań informacyjnych i umożliwieniu odnalezienia w zbiorze wyszukiwawczym tych informacji, które spełniają warunki wyrażone w zapytaniu”.

W dalszych rozważaniach o przyszłości języków informacyjno-wyszukiwawczych zastanowimy się nad tym,

- jaki wpływ na rozwój języków informacyjnych wywarły możliwości urządzeń używanych w realizowanych przez system procesach informacyjnych,
- jaka refleksja teoretyczna towarzyszyła wyżej wymienionym czynnikom, a więc jak rozwijała się teoria języków informacyjno-wyszukiwawczych.

Języki informacyjno-wyszukiwawcze wczoraj

Najważniejszym elementem systemu informacyjno-wyszukiwawczego, w dużej mierze decydującym o jego funkcjonalnej wartości dla użytkownika, jest język informacyjno-wyszukiwawczy. Od jego struktury zależy zarówno wartość zbioru wyszukiwawczego systemu, wyznaczana przez możliwość wyrażenia w języku informacyjno-wyszukiwawczym informacji zawartej w dokumentach odwzorowywanych w zbiorze wyszukiwawczym, jak i możliwości wyszukiwawcze systemu, wyznaczone przede wszystkim przez strukturę paradygmatyczną i syntagmatyczną języka informacyjno-wyszukiwawczego.

Języki informacyjno-wyszukiwawcze istniały, choć wtedy nikt o nich nie tylko nie mówił, ale nawet nie myślał w tych kategoriach, od kiedy funkcjonowały pierwsze systemy informacyjno-wyszukiwawcze, a były nimi już starożytne biblioteki, żeby wspomnieć chociażby bibliotekę aleksandryjską. Nie miejsce tu na przypomnienie ich historii, którą Państwo zapewne znają lepiej niż ja.

Rozwój języków informacyjno-wyszukiwawczych wyznaczany był zawsze przez trzy najważniejsze czynniki:

- przyrost informacji koniecznej do przetworzenia,
- rozwój i specjalizację potrzeb informacyjnych użytkowników informacji (systemów informacyjnych),
- możliwości systemu w zakresie przetwarzania informacji.

Przez tysiące lat istnienia systemów informacyjnych waga dwóch pierwszych czynników zwiększała się nieustannie:

- zbiory informacji ciągle się powiększały, bo ludzie ciągle tworzyli nowe informacje (znana teza o przyroście wykładniczym informacji) i wprowadzali je do systemów informacyjnych;

– coraz bardziej powiększał się zbiór użytkowników systemów informacyjnych, rosła też świadomość użytkowników informacji w zakresie potrzeb informacyjnych: chcieli uzyskać możliwie szybko jak najwięcej informacji na interesujący ich temat i tylko takich (problem kompletności i trafności wyszukiwania, a więc relewancji i pertynencji informacji);

– możliwości systemów informacyjnych w zakresie przetwarzania informacji prawie się nie zmieniały, a narzędziem do przetwarzania informacji i wyszukiwania był człowiek, przede wszystkim bibliotekarz, który, na początku, język informacyjno-wyszukiwawczy miał w głowie: wiedział, gdzie, w jakiej szafie, na jakiej półce znajduje się książka potrzebna czytelnikowi. Później pamięć bibliotekarza wspomagały różnego rodzaju języki informacyjno-wyszukiwawcze w postaci różnego rodzaju spisów lub klasyfikacji bibliotecznych. Urządzenia wspomagające wyszukiwanie pojawiły się bardzo niedawno: tzw. mała mechanizacja

w systemach informacyjno-wyszukiwawczych to właściwie dopiero połowa ubiegłego wieku, a komputery to ostatnie dziesięciolecie.

Bibliotekarz trwał na swoim posterunku, a tymczasem dokumentów przybywało. Ich zbiór, zwłaszcza po wynalezieniu druku, powiększał się tak szybko, że pamięć bibliotekarza przestała ogarniać, nie tylko ich zawartość, ale również przynależne im w zbiorze miejsce. Zadanie, a właściwie misja zawodu bibliotekarza, misja zadowolenia czytelnika, dostarczenia mu jak najszybciej dokumentu zawierającego potrzebną mu informację, zaczęło przerastać jego psychiczne i fizyczne możliwości.

Jedyną możliwością zapewnienia jakiejś takiej sprawności systemów informacyjnych było doskonalenie języków informacyjno-wyszukiwawczych. Pierwszym etapem w ich rozwoju było, wobec wzrostu liczby dokumentów, wprowadzenie klasyfikacji służących do ich fizycznego porządkowania, drugim wprowadzenie spisów dokumentów, a więc stworzenie zbioru dokumentów pochodnych, co umożliwiło przeprowadzanie operacji wyszukiwania nie na zbiorze dokumentów prymarnych (co stawało się fizycznie coraz trudniejsze), ale na zbiorze odwzorowujących je dokumentów pochodnych – tak powstały klasyfikacje biblioteczne i katalogi. Dalszym, ważnym krokiem, było oddzielenie informacji o dokumentach od konkretnej, fizycznej ich kolekcji – tak powstały bibliografie i języki opisujące ich zawartość, a więc klasyfikacje bibliograficzne.

Oddzielenie informacji o dokumencie od dokumentu było wprowadzie przyczyną nowych problemów zarówno natury praktycznej, jak i teoretycznej: problem tworzenia zbiorów dokumentów pochodnych i wyszukiwania informacji w tych zbiorach, umożliwiło jednak rozszerzenie kręgu użytkowników informacji, którzy nie musieli już fizycznie znajdować się w miejscu posadowienia zbioru dokumentów prymarnych. To problem skomplikowało dodatkowo, bowiem wobec utrudnienia użytkownikowi konfrontacji informacji pochodnej (informacji o treści dokumentu) z informacją prymarną, zawartą w dokumencie, co umożliwiłoby mu wprowadzenie ewentualnych korekt lub uzupełnień, przed obsługą systemu informacyjnego postawiło trudne zadanie odpowiedniego, a więc adekwatnego do treści dokumentu, ale i do potrzeb użytkownika informacji, opracowania informacji pochodnej. Najważniejszy stał się więc opis treści dokumentu (opracowanie rzeczowe), do czego tradycyjne języki informacyjno-wyszukiwawcze nie były przygotowane. Problem komplikował fakt, że w tak poszerzonym zbiorze użytkowników informacji znaleźli się użytkownicy władający różnymi językami naturalnymi, aby więc przełamać barierę językową, trzeba było odpowiednio przystosować język informacyjno-wyszukiwawczy. Rozwiązania problemu szukano na dwa sposoby:

- starano się stworzyć język niezależny od jakiegokolwiek języka naturalnego, a więc język o notacji sztucznej – takim językiem była na przykład Klasyfikacja Dziesiąta Deweya i Uniwersalna Klasyfikacja Dziesiąta. Ta niezależność była pozorna, bo odpowiedniki słowne symboli były w języku naturalnym, co w pewien sposób wyznaczało konotację, a więc znaczenie symboli;

- starano się stworzyć języki informacyjno-wyszukiwawcze zrozumiałe w swej formie dla użytkowników władających danym językiem naturalnym, a więc języki o notacji paranaturalnej – takimi językami były już biblioteczne języki haseł przedmiotowych, a później języki słów kluczowych i języki deskryptorowe.

Oczekiwania użytkowników informacji względem systemów informacyjnych, którzy chcieli uzyskiwać informację jak najlepiej wyselekcjonowaną, spowodowały rozwój języków specjalistycznych, przeznaczonych dla systemów dziedzinowych, a więc specjalistycznych klasyfikacji i dziedzinowych języków deskryptorowych. Przyczyniły się też do dalszej rozbudowy języków uniwersalnych, takich jak Uniwersalna Klasyfikacja Dziesiętna czy języki haseł przedmiotowych - tu jako przykład można wymienić używany w polskich bibliotekach uniwersyteckich język KABA. Wpłynęły też na rozwój paradygmatyki i gramatycznych środków, takich jak relacje hierarchiczne i asocjacyjne oraz wskaźniki więzi, roli i wagi w językach deskryptorowych.

Wszystkie te języki informacyjno-wyszukiwawcze były w systemie informacyjnym językami informacji pochodnej, informacji o zawartości dokumentów. Tak było nawet w systemach informacji faktograficznej, bo informacje gromadzone w systemie również musiały być opisywane w języku informacyjno-wyszukiwawczym. Przygotowanie zbioru informacji pochodnej, na którym przeprowadzana była operacja wyszukiwania, wraz z rozwojem języków informacyjno-wyszukiwawczych, powiększaniem się ich zasobów leksykalnych i komplikowaniem się systemów paradygmatycznych i gramatycznych stawało się coraz trudniejsze i wymagało coraz więcej czasu. Ambicja sprostania przez system coraz bardziej wysublimowanym oczekiwaniom użytkowników wymagała coraz dokładniejszego przedstawiania w charakterystykach wyszukiwawczych treści dokumentów, a więc coraz to głębszego indeksowania, co wymuszało rozbudowę leksyki, a zapewnienie kompletności wyszukiwania tak indeksowanych dokumentów wymagało wyszukiwania paradygmatycznego, a więc odwzorowywania szeroko rozumianych relacji paradygmatycznych (w tym mereologicznych i asocjacyjnych) w słownikach tych języków. Sprostać tym wymogom mogli tylko specjaliści dziedzinowi, bo zarówno tworzenie takich języków informacyjno-wyszukiwawczych, jak i posługiwanie się nimi w procesie analizowania i indeksowania dokumentów wymagało dogłębnej znajomości wiedzy w dokumentach prezentowanej.

Opracowanie rzeczowe zajmowało coraz więcej czasu i stawało się wąskim gardłem systemów informacyjno-wyszukiwawczych, ciągle bowiem wykonywane było przez pracowników systemu, którzy dokonywali analizy dokumentów, sporządzali ich charakterystyki słowne, a następnie przekładali je na język informacyjno-wyszukiwawczy systemu, tworząc w ten sposób charakterystyki wyszukiwawcze dokumentów. I choć samo wyszukiwanie z czasem odbywało się ze wspomaganiami różnego typu urządzeniami mechanicznymi czy automatycznymi, to operacje składające się na przygotowywanie informacji dla systemu długo odbywały się w sposób tradycyjny. Nie pomogły nawet pierwsze komputery, bo miały zbyt małą pamięć, pracowały zbyt wolno, a użycie ich do tworzenia charakterystyk wyszukiwawczych dokumentów wymagałoby nauczania ich najpierw przetwarzania tekstów języka naturalnego, co, mimo rozwoju lingwistyki komputerowej, długo pozostawało w sferze marzeń.

A jednak tym, co naprawdę przesądziło o rozwoju nie tylko języków informacyjno-wyszukiwawczych, ale całej informacji naukowej, było właśnie pojawienie się komputera, zupełnie nowego narzędzia, nazywanego najpierw elektroniczną maszyną liczącą. Fascynacja środowiska bibliotekarskiego jej możliwościami

dopingowała do podjęcia prac natury praktycznej i teoretycznej nad jak najlepszym wykorzystaniem komputerów w służbie bibliotek i ośrodków informacyjnych, którym już przestały wystarczać tradycyjne katalogi i wspomagające je instrumenty tzw. małej mechanizacji, operujące na kartach obrzeźnie lub przeziemie dziurkowanych.

Nowe narzędzie otwierało nowe perspektywy, ale też stawiało nowe wymagania. Podstawowym warunkiem automatyzacji operacji dokonywanych w systemach informacyjno-wyszukiwawczych była ich formalizacja – komputery przecież nie myślą, potrafią tylko przekształcać napisy, a ponieważ operacje w systemie informacyjnym wykonywane są na tekstach języka naturalnego, w jakim pisane są dokumenty i języka informacyjno-wyszukiwawczego używanego w systemie, zaczęła się era wytężonych prac nad językami informacyjno-wyszukiwawczymi i ich formalizacją.

Języki informacyjno-wyszukiwawcze są językami sztucznymi - bo tworzonymi specjalnie dla potrzeb najpierw systemów bibliotecznych, później dla wszelkich systemów informacyjno-wyszukiwawczych, ale nie zawsze są językami formalnymi, jeśli przez języki formalne rozumieć takie języki, których prezentacja spełnia warunki podania explicite:

- zbioru wyrażeń elementarnych danego języka (czyli słownika),
- reguł tworzenia wyrażeń złożonych – a więc gramatyki, czyli składni,
- reguł przyporządkowujących wyrażeniom elementów rzeczywistości pozajęzykowej, czyli reguł semantyki – w wypadku języków informacyjno-wyszukiwawczych rzeczywistość pozajęzykowa to zbiór informacji zawartych w dokumentach tworzących zbiór informacyjny systemu. Dla języków uniwersalnych, na przykład Uniwersalnej Klasyfikacji Dziesiętnej, rzeczywistość ta to cały potencjalny zbiór dokumentów źródłowych,
- relacji między wyrażeniami w systemie języka, czyli paradygmatyki języka,
- relacji między wyrażeniami tworzącymi wyrażenia złożone języka (zdania), czyli syntagmatyki języka,
- relacji zachodzących między komunikatami językowymi a sytuacją komunikacyjną, a więc w wypadku języków informacyjno-wyszukiwawczych między potencjalnymi charakterystykami wyszukiwawczymi i instrukcjami wyszukiwawczymi a odwzorowywanymi dokumentami oraz użytkownikami systemu. Są to reguły pragmatyki języków informacyjno-wyszukiwawczych. Tu właśnie powinny się znaleźć explicite sformułowane tak zwane reguły indeksowania, czyli tworzenia charakterystyk wyszukiwawczych dokumentów i instrukcji wyszukiwawczych, a nawet szerzej – opracowywania strategii wyszukiwania w zależności od potrzeb użytkowników, w tym tak zwane zasady relewancji pragmatycznej.

Języki sztuczne, a więc języki specjalnie projektowane, powinny być w taki właśnie sposób opisane – języki informacyjno-wyszukiwawcze nie spełniają jednak tych wymogów. Różne są tego powody. Najważniejsze z nich to:

- brak wiedzy lingwistycznej ich twórców w zakresie projektowania języków i opisu systemu językowego – języki informacyjno-wyszukiwawcze tworzyli przede wszystkim praktycy: bibliotekarze lub dokumentaliści, a nie lingwiści;
- niewykształcenie metajęzyka opisu języków informacyjno-wyszukiwawczych, a więc terminologii służącej do prezentacji tych języków. Stąd te same obiekty lingwistyczne nazywane były różnie w różnych językach lub ich typach,

a przede wszystkim w różnych tradycjach ich opisu. Nie zawsze kontrolowany rozwój istniejących już języków informacyjno-wyszukiwawczych, które w miarę użycia, aby sprostać potrzebom informacyjnym systemów i ich użytkowników, były rozbudowywane przez tworzących komunikaty metainformacyjne. Prowadziło to do powstawania różnych, znacznie różniących się odmian tych samych języków (na przykład języka haseł przedmiotowych w Polsce w odmianie używanej przez Bibliotekę Narodową i Bibliotekę Uniwersytetu Warszawskiego).

Brak formalnego opisu języków informacyjno-wyszukiwawczych posługującego się tym samym metajęzykiem i opartego na tych samych zasadach prezentacji systemu ma daleko idące konsekwencje, gdyż:

- utrudnia opanowanie języka informacyjno-wyszukiwawczego przez jego użytkowników, a więc zarówno dokumentalistów, jak i użytkowników systemów informacyjnych;

- jest przyczyną szumu informacyjnego i strat informacji w systemie – zarówno na wejściu systemu, gdyż utrudnia stworzenie poprawnych charakterystyk wyszukiwawczych dokumentów i jest przyczyną błędów w indeksowaniu, jak i na wyjściu systemu, utrudniając sformułowanie poprawnych instrukcji wyszukiwawczych i obranie optymalnych strategii wyszukiwawczych;

- utrudnia ocenę poprawności budowy języków informacyjno-wyszukiwawczych, ocenę ich możliwości metainformacyjnych i porównywanie różnych języków.

Za istniejący stan rzeczy w prezentacji języków informacyjno-wyszukiwawczych nie można winić twórców tych języków. Rozwój metalingwistyki to właściwie dopiero druga połowa XX w., kiedy to, korzystając z metod opisu systemowego wypracowanych przez strukturalizm, zaczęto zajmować się modelami lingwistycznymi.

W wypadku języków informacyjno-wyszukiwawczych stosunkowo dobrze prezentowane były słowniki (rozumiane jako zbiory jednostek leksykalnych), i to niezależnie od typu języka - na przykład listy słów kluczowych dla języków słów kluczowych o kontrolowanej leksyce, tablice klasyfikacyjne podające symbole proste, jak w Uniwersalnej Klasyfikacji Dziesiętnej, listy tematów i listy określników w językach haseł przedmiotowych, układy deskryptorów w tezaurusach języków deskryptorowych. O wiele gorzej przedstawia się prezentacja relacji paradygmatycznych między tymi wyrażeniami. Językami, które z założenia prezentowały takie relacje, są języki deskryptorowe, prezentujące relacje paradygmatyczne w tezaurusie oraz klasyfikacje hierarchiczne. Językami, które z założenia nie prezentują relacji paradygmatycznych, są języki słów kluczowych. Językami, których prezentacje pokazują tylko wybrane relacje paradygmatyczne za pomocą systemów odsyłaczy, są języki haseł przedmiotowych.

Prezentacja relacji paradygmatycznych, zarówno hierarchicznych jak i asocjacyjnych, jest dla użytkownika języka bardzo ważna, pozwala bowiem na dokładniejsze odwzorowanie w charakterystyce wyszukiwawczej informacji zawartych w dokumentach, umożliwiając wybór spośród jednostek leksykalnych języka tych najbardziej relewantnych. Znacznie również usprawnia wyszukiwanie informacji i oszczędza wysiłek użytkownika systemu, umożliwiając automatyczne wyszukiwanie paradygmatyczne, a poprawiając parametry relewancji wyszukiwania, zmniejsza także szum informacyjny i straty informacji, podpowia-

dając możliwe transformacje obranej strategii wyszukiwawczej. Rozwój badań nad semantyką języków naturalnych, a szczególnie prace dotyczące konstruowania metajęzyków semantycznych, dały teorii języków informacyjno-wyszukiwawczych narzędzie do opisu systemów semantycznych tych języków i specyfikacji semantycznych relacji paradygmatycznych ich systemów leksykalnych, dały także asumpt do powstania nowego typu języków informacyjno-wyszukiwawczych, tak zwanych kodów semantycznych.

Pokazywanie explicite relacji paradygmatycznych języków informacyjno-wyszukiwawczych jest jednym ze sposobów częściowej prezentacji ich systemu semantycznego – częściowej, bo ograniczonej tylko do ukazania wybranych opozycji semantycznych, nie zawsze zresztą wystarczająco wyraźnie określonych. Komplementarnym, a w niektórych typach języków informacyjno-wyszukiwawczych jedynym sposobem prezentacji ich systemów semantycznych, jest odwoływanie się – explicite lub implicite – do znaczenia skorelowanych z jednostkami leksykalnymi tych języków wyrażen języka naturalnego. W sposób wyraźny ma to zawsze miejsce w wypadku języków informacyjno-wyszukiwawczych o notacji sztucznej, podających, jak w Uniwersalnej Klasyfikacji Dziesiętnej, tak zwane odpowiedniki słowne, mniej lub bardziej wyraźnie dzieje się to w językach o notacji paranaturalnej. W jednym i drugim wypadku przyjmuje się założenie, że użytkownik języka sam domyśli się znaczenia wyrażenia języka informacyjno-wyszukiwawczego na podstawie znajomości znaczenia przyporządkowanego mu lub równokształtnego z nim wyrażenia języka naturalnego, jak w wypadku języka haseł przedmiotowych, języka słów kluczowych czy języka deskryptorowego. Niestety, taki sposób prezentacji systemu semantycznego języków informacyjno-wyszukiwawczych bywa przyczyną wielu nieporozumień i błędów w użyciu wyrażen języków informacyjno-wyszukiwawczych. Rzadko kiedy bowiem przy prezentacji systemu semantycznego języków informacyjno-wyszukiwawczych zwraca się uwagę na dwie podstawowe sprawy:

- że języki informacyjno-wyszukiwawcze i języki naturalne odwzorowują dwie zupełnie różne rzeczywistości. Pierwsze (informacyjne) rzeczywistość dokumentacyjną – wyrażeniom tych języków przyporządkowane są klasy dokumentów zawierające odpowiednie informacje, natomiast drugie (naturalne) rzeczywistość obiektywną (całe uniwersum) – ich wyrażeniom przyporządkowane są klasy obiektów (relacji, sytuacji) realnego świata;

- że rekonstruując znaczenie wyrażen języków informacyjno-wyszukiwawczych na podstawie przyporządkowanych im wyrażen języków naturalnych nie wolno sugerować się znaczeniem i użyciem tych ostatnich, nawet jeśli się pamięta o tym, że należy tu brać pod uwagę ich użycie w funkcji metainformacyjnej, gdyż znaczenie wyrażen języków informacyjno-wyszukiwawczych jest współwyznaczane sztuczną paradygmatyką tych języków, znacznie nieraz odbiegającą od paradygmatyki danego języka naturalnego. Sugerowanie się tą ostatnią może powodować błędy w użyciu wyrażen języków informacyjno-wyszukiwawczych oraz stanowić przyczynę niepowodzeń w wyszukiwaniu informacji.

Wyjątkowo niekonsekwentnie była prezentowana gramatyka języków informacyjno-wyszukiwawczych. Najczęściej podawane są tylko niektóre reguły składniowe. Zakłada się, że użytkownik języka sam będzie umiał je wyabstrahować z podanych przykładów wyrażen złożonych bądź też, że zachowa się „zdrowo-

rozsądkowo” i, na przykład, nie będzie powtarzał tych samych wyrażen elementarnych w wyrażeniu złożonym, choć nigdzie *explicite* nie mówi się, że język takich powtórzeń nie dopuszcza (jak na przykład w wypadku języków o notacji paranaturalnej, deskryptorowych czy też języków słów kluczowych). Prezentując gramatykę języków informacyjno-wyszukiwawczych, wprawdzie używa się terminologii lingwistycznej, ale zazwyczaj niekonsekwentnie, co często świadczy o nierozumieniu istoty struktury języka.

Wypracowane w lingwistyce strukturalnej metody tworzenia modeli lingwistycznych, a więc zasad formalizacji i prezentacji systemów językowych, miały istotne znaczenie także w projektowaniu języków sztucznych. Tylko przy takim opisie projektowanego języka można się ustrzec błędów w projektowaniu i zachować pełną kontrolę nad systemem językowym. Koniecznie trzeba tu wspomnieć o roli tak zwanej semantyki składnikowej, pod wpływem której zrodził się zupełnie nowy typ języków informacyjno-wyszukiwawczych. Były to kody semantyczne, a najbardziej z nich znane to stworzony w Ośrodku Dokumentacji i Komunikacji Uniwersytetu w Cleveland w stanie Ohio USA język dla systemu informacji metalurgicznej, tak zwany kod semantyczny Perry’ego i Kenta (koniec lat pięćdziesiątych XX w.) oraz ukraiński język R_x -kodów, stworzony przez E. F. Skorochod’kę. W Polsce prace nad kodami semantycznymi dla systemu informacji naukowej związane były z tworzeniem systemu SINTO i we współpracy z IINTE prowadzone były w latach siedemdziesiątych w Katedrze Lingwistyki Formalnej Uniwersytetu Warszawskiego przez Bożennę Bojar i Olgierda Adriana Wojtasiewicza. Kody semantyczne to języki, których wyrażenia elementarne odwzorowują w swej formie elementarne informacje, uznane za relewantne dla użytkownika systemu. Dzięki takiej strukturze planu wyrażania języki te umożliwiają jednoznaczne odwzorowanie wszystkich informacji objętych ich polem semantycznym *explicite* w tekście, istnieje w nich bowiem jedno-jednoznaczne przyporządkowanie między elementarnymi informacjami a elementarnymi wyrażeniami tworzącymi komunikat. Języki takie mogłyby również pełnić funkcję metajęzyków w stosunku do innych języków, zarówno naturalnych, jak i informacyjno-wyszukiwawczych, opisując znaczenie ich wyrażen, co rozwiązałoby problem tak zwanego języka pośrednika w procesie przekładu i pozwoliłoby uspołnić systemy informacyjne posługujące się różnymi językami informacyjno-wyszukiwawczymi. I choć prace nad stworzeniem takiego uniwersalnego języka semantycznego dla systemów informacyjnych nie są obecnie kontynuowane, to doświadczenie wtedy zdobyte, a przede wszystkim świadomość semantycznej struktury informacji, na pewno jest i będzie wykorzystywane przy tworzeniu hipertekstowych systemów informacyjnych i coraz sprawniejszych wyszukiwarek internetowych.

Z metod i ustaleń lingwistyki czerpano również w okresie tworzenia dla systemów informacyjno-wyszukiwawczych języków deskryptorowych. Tu niezbędną była nie tylko wiedza o strukturze systemu paradygmatycznego języka, ale także znajomość systemów terminologicznych poszczególnych dyscyplin naukowych i ustaleń terminologii jako nauki o tych systemach. Wiedza lingwistyczna przy projektowaniu systemów i języków informacyjno-wyszukiwawczych niezbędna jest także przy konstruowaniu pola semantycznego języka – tu przydają się informacje nie tylko z klasycznej już semantyki strukturalnej, ale również ze

współcześnie prowadzonych badań w ramach lingwistyki kognitywnej, zajmującej się między innymi relacjami między strukturą języka – przede wszystkim jego systemem semantycznego – a strukturą wiedzy jego użytkowników oraz problemami przetwarzania informacji językowych. To ta wiedza lingwistyczna pozwoliła na tworzenie systemów informacyjnych coraz bardziej przyjaznych (jak to się modnie nazywa) dla użytkownika, przyjaznych nie tylko pod względem technicznych udogodnień we współpracy z systemem, ale przede wszystkim przyjaznych, bo umożliwiających coraz łatwiejsze dotarcie do informacji naprawdę mu potrzebnych.

Bardzo długo użytkownikami języków informacyjno-wyszukiwawczych byli praktycznie tylko bibliotekarze i dokumentaliści, bo to oni przygotowywali opisy dokumentów i oni też byli dysponentami informacji. Czytelnik w bibliotece i interesant ośrodka dokumentacji języków tych znać nie musiał, bo to nie on prowadził wyszukiwania. Sama rzeczywistość informacyjna jawiła się jeszcze wtedy jako w pełni porządkowalny zbiór, a przejawem takiej wiary była chociażby ciągle rozwijana Uniwersalna Klasyfikacja Dziesiątka.

Nowe narzędzie wyszukiwawcze zmieniło też pozycję użytkownika w systemie. Chodziło o to, żeby użytkownik informacji był obsługiwany przez system jak najlepiej, stąd wiele badań poświęcano badaniom potrzeb użytkowników oraz doskonaleniu systemów. Zadowolony użytkownik miał otrzymać wszystkie potrzebne mu dane o dokumentach zawierających interesujące go treści, dane, którymi system dysponował. Wiele starań poświęcano więc odpowiedniemu określeniu zakresu systemów informacyjnych, gromadzeniu w systemie informacji o wszystkich dokumentach należących do wytyczonego zakresu, wymianie informacji między systemami, a przede wszystkim wypracowaniu odpowiednich kryteriów opracowania rzeczowego, tak aby w charakterystykach treściowych (opisie rzeczowym) dokumentów znalazły się wszystkie informacje o ich zawartości treściowej, które mogą być przydatne użytkownikom systemu.

Tak określone priorytety wyznaczyły listę teoretycznych i praktycznych problemów, którymi powinni się zająć teoretycy i praktycy języków informacyjno-wyszukiwawczych. Przede wszystkim uświadomiono sobie, że dotychczas używane języki informacyjno-wyszukiwawcze, które dobrze spełniały swoją rolę w odniesieniu do zbiorów bibliotecznych, nie najlepiej sobie radzą w systemach informacyjno-wyszukiwawczych, gdzie przewiduje się znacznie większy udział w procesie wyszukiwania informacji samego użytkownika. Przede wszystkim dlatego, że tradycyjne języki informacyjno-wyszukiwawcze były w zasadzie kodami jednostronnymi, zarówno bowiem w procesie formułowania w nich komunikatu, jak i w procesie wyszukiwania informacji posługiwali się nimi tylko pracownicy systemu – dla normalnego użytkownika systemu informacyjnego były niedostępne, nie były mu bowiem prezentowane w formie jawnej, a w wypadku języków typu klasyfikacji sztuczna notacja i znacznie rozbudowana struktura hierarchiczna stanowiły bariery nie do pokonania. Nie najlepiej także sprawdzały się w procesie wyszukiwania informacji, a to przede wszystkim ze względu na tzw. cechę prekoordynacji, którą wówczas ściśle wiązano z typem języka. Istotną ich wadą było także to, że nie nadawały się do wszechstronnego odwzorowania treści dokumentu w charakterystyce wyszukiwawczej, a rekompensowanie tej cechy tzw. przydziałem wielokrotnym proces indeksowania dokumentu czyniło znacz-

nie bardziej pracochłonnym. Stąd idea stworzenia języków informacyjno-wyszukiwawczych zupełnie innego typu, wolnych od takich ograniczeń.

Takimi językami wydawały się języki generowane z języków naturalnych, a więc łatwe dla użytkownika, który powinien znać podstawową terminologię interesującej go dziedziny wiedzy. Mógłby wtedy sam posługiwać się nimi w procesie wyszukiwania informacji, a więc samodzielnie formułować instrukcję wyszukiwawczą dla systemu i samodzielnie kontrolować poszczególne etapy wyszukiwania. Musiałyby to być jednak języki o możliwie najprostszej składni, którego to warunku nie spełniały wywodzące się z leksyki języka naturalnego, mające dobrze ugruntowaną w bibliotekarstwie tradycję języki haseł przedmiotowych. Idealnymi językami wydawały się zrazu języki słów kluczowych, bo w wypadku tzw. swobodnych słów kluczowych nawet nie trzeba było ich tworzyć, a współrządne indeksowanie, w którym były używane, zapewniało swobodne i łatwe wyszukiwanie. Dodatkowym argumentem za używaniem w systemach informacyjno-wyszukiwawczych tych właśnie języków była perspektywa, jaką stwarzały „elektroniczne maszyny liczące”, automatyzacji w systemach informacyjno-wyszukiwawczych wszystkich operacji tekstowych, a więc nie tylko wyszukiwania, ale także tworzenia charakterystyk wyszukiwawczych. Już wtedy myślano także nie tylko o tworzeniu skomputeryzowanych systemów informacji dokumentacyjnej, ale także o systemach hybrydowych, przetwarzających tekst dokumentów i łączących informacje o charakterze dokumentacyjnym z informacjami faktograficznymi, a więc właściwie o systemach informacyjnych przetwarzających wiedzę. Do urzeczywistnienia tych planów miały się przyczynić przede wszystkim wyniki badań prowadzonych przez naukowców pręźnie się wówczas rozwijającej w obrębie językoznawstwa nowej dyscypliny, tzw. przekładu maszynowego, z której powstała lingwistyka komputerowa. To właśnie lingwistyka komputerowa miała pozwolić na uporanie się w systemach informacyjnych z zasadniczą wadą języków słów kluczowych, którą było związanie ich z konkretnymi językami naturalnymi, co ograniczało ich używanie w uniwersalnych systemach informacyjno-wyszukiwawczych i powodowało brak kompatybilności systemów różnych krajów.

W tworzonych wtedy systemach informacyjno-wyszukiwawczych ogromną wagę przywiązywano do sprawności wyszukiwawczej, stąd wiele prac poświęcano w informacji naukowej poprawianiu takich parametrów systemu, jak kompletność, relewancja i pertynencja. Chodziło więc o zapobieganie zarówno tzw. ciszy informacyjnej, jak i szumowi informacyjnemu, a wszystko to w trosce o zadowolenie użytkownika systemu. Proste języki słów kluczowych nie najlepiej się tu sprawdzały, lepsze miały być języki deskryptorowe. Zachowywały one wszystkie zalety języków o notacji naturalnej, ale ich rozbudowana paradygmatyka, prezentowana w tezaurusach, używanych nie tylko przy indeksowaniu dokumentów, ale także w procesie wyszukiwawczym, poprzez wyszukiwanie hierarchiczne i relacje asocjacyjne zapewniała większą kompletność wyszukiwanej informacji, a poprzez wprowadzenie w planie syntagmatycznym tzw. wykładników związku zgody oraz specyfikację relacji skojarzeniowych znacznie zmniejszała szum informacyjny, a więc poprawiała współczynnik relewancji. Temu ostatniemu miało też służyć tzw. ważenie deskryptorów. Odpowiednią relewancję informacji w procesie indeksowania mogły zapewnić tylko dziedzinowe języki deskryptorowe,

wiele wysiłku i czasu poświęcono więc w informacji naukowej na tworzenie takich właśnie, coraz bardziej szczegółowych tezaurusów specjalistycznych.

Troska o użytkownika i jego potrzeby oraz rosnąca świadomość roli informacji naukowej w rozwoju gospodarki spowodowały masowe powstawanie różnego rodzaju specjalistycznych ośrodków informacyjnych, przede wszystkim związanych z różnymi dziedzinowymi instytutami badawczymi i dużymi przedsiębiorstwami. Dla stworzonych w nich systemów informacji dziedzinowej i specjalistycznej gromadzących literaturę fachową budowano języki o odpowiednio ograniczonym polu tematycznym, które łączyć miały tzw. klasyfikacje szczytowe, takie jak powstała w CINTE Polska Klasyfikacja Tematyczna. Prace nad tworzeniem takich języków informacyjnych, umożliwiających łączenie języków dziedzinowych i współpracę różnych systemów informacyjnych prowadzone były w wielu krajach i były przejawem wiary w realną możliwość stworzenia kontrolowanego hipersystemu informacyjnego.

Ciągłe dążenie do zapewnienia systemom informacyjnym optymalnych parametrów wyszukiwawczych zmuszało teoretyków informacji naukowej do stworzenia takich języków informacyjno-wyszukiwawczych, które zapewniałyby wyszukiwanie odpowiednio relewantnej informacji w optymalnym czasie, a tu jednym z najważniejszych warunków było zapewnienie odpowiednio relewantnego odwzorowania treści dokumentów w charakterystykach wyszukiwawczych. Można to było osiągnąć tylko poprzez odpowiednie języki informacyjno-wyszukiwawcze, przygotowane do odpowiednio szczegółowego indeksowania – klasyczne języki informacyjne, tworzone dla zbiorów uniwersalnych, zakładały bowiem zawsze stosunkowo daleko idące uogólnienia. Teoria języków informacyjno-wyszukiwawczych, czerpiąc z teoretycznych osiągnięć semantyki języka naturalnego, problem ten próbowała rozwiązać poprzez budowanie języków zupełnie nowego typu, wspomnianych już kodów semantycznych, a więc języków kodujących bezpośrednio w planie wyrażania informacji dystynktywnych dla użytkowników.

Języki informacyjno-wyszukiwawcze dziś ... czy jutro?

Praktyka języków informacyjno-wyszukiwawczych wybrała jednak inne rozwiązanie – szczegółową rozbudowę leksyki paranaturalnych języków informacyjnych. Tak uczyniono na przykład przy tworzeniu języka KABA, rozbudowując leksykę języka haseł przedmiotowych do rozmiarów znacznie przekraczających zasób czynnej leksyki przeciętnego użytkownika języka polskiego, często przy tym zmieniając denotację i konotację wyrażań, co w rezultacie zamiast do poprawy relewancji prowadzi do zwiększenia szumu informacyjnego, i jeszcze groźniejszego w informacji zjawiska tzw. ciszy informacyjnej.

Efektom tworzenia tak rozbudowanych języków informacyjno-wyszukiwawczych jest również niewspółmierność kosztów ich budowy i użycia w procesie indeksowania do efektów wyszukiwania, ogromna część takiej leksyki nigdy bowiem nie jest używana w procesie wyszukiwania przez przeciętnych użytkowników, których wiedza o przedmiocie i dziedzinie w momencie podjęcia wyszukiwania jest zazwyczaj zbyt uboga na to, by mogli się posłużyć jednostką leksykalną

o odpowiednio wąskiej denotacji, a więc o odpowiednio dużej rozdzielczości wyszukiwawczej. Śledząc rozwój tego języka, można odnieść wrażenie, że twórcy jego zapomnieli o wypracowanych przez semiotykę kryteriach, jakie powinny spełniać kody optymalne (o których pisał Tadeusz Wójcik w swej „Prakseosemiiotyce”, a o czym nie raz pisała także zajmująca się projektowaniem języków informacyjno-wyszukiwawczych Jadwiga Woźniak). Dobry język to taki, który jest dostosowany do: nadawcy komunikatu, odbiorcy komunikatu (adresata) i rzeczy komunikowanej. Dostosowany do jego użytkowników, a więc nie tylko umożliwiający przekazywanie im informacji relewantnej, ale przede wszystkim umożliwiający im sprawne posługiwanie się językiem. Projektując taki język, trzeba przede wszystkim pamiętać o tym, że w wypadku języka informacyjno-wyszukiwawczego o notacji paranaturalnej zawsze zagraża interferencja systemu języka naturalnego. Należy to uwzględnić, projektując strukturę semantyczną i paradygmatyczną języka. Z dużą dozą ostrożności należy też określić liczebność systemu leksykalnego takiego słownika, pamiętając o tym, że zbyt mała liczba jednostek daje małą rozdzielczość informacji, ale zbyt duża uniemożliwia sprawne posługiwanie się językiem, który staje się wręcz niemożliwy do opanowania zarówno dla osób sporządzających opisy dokumentów, jak i dla wyszukujących. Często niemożliwe jest wtedy odnalezienie w systemie języka adekwatnego wyrażenia, jeśli system nie przewiduje rozbudowanej kategoryzacji semantycznej na poziomie metajęzykowym. Wszystko to powoduje, że wysiłek włożony w tworzenie takiego języka staje się absolutnie nieopłacalny, gdyż i tak wyrażenia zbyt szczegółowe nie będą używane w procesie indeksowania, a jeśli nawet będą, to informacja taka będzie stracona dla użytkownika, ponieważ w większości wypadków, nie posiadając odpowiedniej wiedzy o treści danego dokumentu, nie dotrze do odpowiednich wyrażen przy opracowywaniu strategii wyszukiwawczej. I tu właśnie należy korzystać z ustaleń lingwistyki dotyczących zasobu czynnego i zasobu biernego systemu leksykalnego używanego przez użytkowników języka, znajomości konstytuujących go opozycji semantycznych, znajomości znaczenia wyraźnego i operacyjnego wyrażen oraz z ustaleń dotyczących struktury pamięci lingwistycznej użytkowników. Problemami tymi zajmuje się przede wszystkim psycholingwistyka i lingwistyka kognitywna.

Wiedza o strukturze systemu semantycznego w kompetencji użytkownika języka jest absolutnie niezbędna przy tworzeniu współczesnych wielkich systemów informacyjnych, używających do charakteryzowania dokumentów słów kluczowych, odnoszących się w wyszukiwaniu informacji nie tylko do kompetencji ich użytkowników w zakresie formy wyrażen, ale również ich znaczenia i znajomości struktury paradygmatycznej. Przy optymalizacji działania tych systemów mogą być przydatne badania nad strukturą definicji w tradycyjnych słownikach jednojęzycznych. Chodzi tu o tak zwaną metodę ciągów definicyjnych, odwzorowujących, jak stwierdzono, paradygmatyczną strukturę kompetencji językowej użytkowników języka oraz o metodę tzw. gniazd leksykalnych, odwzorowujących relacje asocjacyjne. Metoda ta może być wykorzystana zarówno do generowania wyjściowej struktury systemu leksykalnego, jak i do polepszenia wyników wyszukiwania informacji w systemach stosujących języki słów kluczowych. Należy także wykorzystywać słowniki frekwencyjne, co pozwoli na ustalenie optymalnego zasobu leksykalnego tych języków.

Skokowy rozwój elektroniki i informatyki w końcu ubiegłego wieku sprawił, że komputery „nauczyły się” przetwarzać teksty języka naturalnego, można więc było użyć ich nie tylko do wyszukiwania informacji w specjalnie spreparowanych zbiorach informacji, ale do tworzenia takich zbiorów i przeszukiwania po prostu tekstów dokumentów.

Opracowując algorytmy analizy tekstu, konieczne należy skorzystać z osiągnięć lingwistyki tekstu, zajmującej się strukturą tekstu, w tym strukturą tematyczno-rematyczną wypowiedzenia, strukturą akapitu, a także informacyjną strukturą tekstu spójnego oraz zależnościami pomiędzy miejscem wypowiedzi w tekście a ich wartością informacyjną. Stworzenie algorytmu dokładnie analizującego tekst wymagałoby skorzystania z wypracowanych w lingwistyce algorytmów analizy składniowej, identyfikującej wszelkie kategoriale wykładniki foryczne (anafora i katafora), a także ustaleń dotyczących wykorzystania w strukturze tekstu spójnego wyrażen pozostających w stosunku do siebie w relacji hierarchii generycznej. Dobrej automatycznej analizy tekstu nie dałoby się również dokonać bez znajomości statystycznej struktury tekstu, a więc ustaleń lingwistyki statystycznej.

Znajomość struktury informacyjnej tekstu spójnego niezbędna jest także przy automatycznym abstrahowaniu informacji dla systemów informacji faktograficznej. Dla użytkownika takich systemów ważnym wskaźnikiem wyznaczającym zaufanie do systemu jest ocena wartości informacji udostępnianej przez system. Do automatycznego wyznaczania wskaźników takiej wartości niezbędna jest znajomość gramatycznych i leksykalnych wykładników oceny wartości informacji obecnych w systemie morfologicznym i leksykalnym języka naturalnego, a więc jego wykładników metainformacyjnych.

Wymienione tu główne nurty w rozwoju języków informacyjno-wyszukiwawczych były pochodną wiary w możliwość stworzenia kontrolowanego i sterowanego systemu informacyjnego, co znacznie ułatwić miały komputery. Wydaje się, że zapomniano jednak o tym, że tak zaprojektowane systemy mogą się rozwijać i działać zgodnie z przewidywaniami tylko do osiągnięcia pewnej wielkości krytycznej. Wielkie systemy działają inaczej i po przekroczeniu pewnego progu stają się niesterowalne, przekształcają się w systemy samoorganizujące się. Sądzę, że w takim stadium przejściowym znajdują się obecnie systemy wyszukiwania informacji.

Wynalezienie i wdrożenie Internetu pozwoliło na stworzenie uniwersalnego systemu informacyjnego o zasięgu światowym, łączącego w swych funkcjach tradycyjny system informacji dokumentacyjnej z nowoczesnym systemem wyszukiwania pełnotekstowego i systemem informacji faktograficznej. To Internet pozwolił na urzeczywistnienie marzeń wszystkich teoretyków, praktyków i użytkowników tradycyjnych systemów informacyjno-wyszukiwawczych i utworzenie multimedialnego systemu hipertekstowego. Tyle tylko, że system ten daleko odbiega od klasycznych systemów informacyjnych, różni się od nich nie tylko wielkością, ale i jakością. To jest właśnie ten moment, kiedy „ilość przemienia się w jakość” (choć wielu użytkowników wyszukiwarek internetowych, szczególnie tych przyzwyczajonych do systemów tradycyjnych, powiedziałoby, że jest to raczej bylejakość). Tradycyjne systemy informacyjno-wyszukiwawcze ściśle odgraniczały rolę nadawcy od odbiorcy. Nadawcami komunikatów byli pracow-

nicy dokonujący kwalifikacji dokumentów włączanych do systemu i przygotowujący ich charakterystyki wyszukiwawcze, oni też wyznaczali możliwe do realizacji ścieżki wyszukiwania poprzez odpowiednią strukturę języków informacyjno-wyszukiwawczych. Odbiorcami systemu byli jego użytkownicy, którym twórcy i administratorzy systemów wyznaczali rolę w zasadzie bieżącej. Długo zresztą trwały kiedyś dyskusje, czy należą oni do systemu informacyjnego, czy też do jego otoczenia. W nowoczesnych systemach podział ten uległ zatarciu – użytkownicy są jednocześnie twórcami zbiorów wyszukiwawczych systemu, sami mogą włączać do niego informacje, sami też wytyczają drogi poszukiwań. A tradycyjne charakterystyki wyszukiwawcze przestały już być potrzebne, bo wyszukiwarki operujące na tekście dokumentu same dokonują identyfikacji słów kluczowych. Nie są więc już potrzebne tradycyjne języki informacyjno-wyszukiwawcze. Nie ma też sensu podejmowanie działań poprawiających tradycyjne współczynniki sprawności systemu, a więc poprawiających kompletność i trafność wyszukiwania. Ta pierwsza wobec ogromnej, ciągle się zmieniającej liczby dokumentów na każdy temat znajdujących się w systemie jest nie do wyliczenia – kompletność nie jest zresztą ważna, bo i tak żaden użytkownik nie byłby w stanie zapoznać się z ich treścią. Ta druga byłaby dla użytkownika istotna, bardziej niż w wypadku systemów o małych zbiorach, ale i tamte w zasadzie zapewniały tylko tzw. relewancję techniczną, a więc nie gwarantowały, że wyszukane dokumenty będą zawierały potrzebne użytkownikowi informacje, nie oceniały więc tzw. pertynencji, czyli odpowiedniości treści wyszukanych dokumentów do faktycznych potrzeb użytkownika. Tym bardziej, że systemy tradycyjne miały w zasadzie sztywne strategie wyszukiwawcze, z góry wyznaczane przede wszystkim przez strukturę języka i nie modyfikowały jej pod wpływem wyników odbytych już poszukiwań.

W porównaniu z tradycyjnymi, uporządkowanymi, ale statycznymi systemami wyszukiwania informacji, dynamiczne systemy internetowe z pozoru jawić się mogą jako chaos, ale jak wynika z teorii chaosu, w skomplikowanych układach wszelkiego typu z czasem pojawia się porządek wyższego rzędu, bo lokalne procesy zachodzące w układzie wzajemnie się dostosowują (tu można przywołać zamieszczony w 1 numerze dodawanego do Polityki „Niezbędnika inteligenta” ciekawy wywiad Sławomira Mizerskiego z prof. Michałem Tempczykiem, który też kiedyś zajmował się informacją naukową, zatytułowany *Chaotyczna harmonia świata*). Czynnikiem porządkującym chaos internetowych systemów informacyjnych jest tzw. mądrość tłumu, o której pisał James Surowiecki w książce *Mądrość tłumów (The Wisdom of Crowds)*. „Mądrość tłumów doskonale funkcjonuje, dowodzi Surowiecki, nie tylko w świecie finansów, lecz też w sferach nowoczesnej nauki i technologii. Zaskakująca sprawność wyszukiwarki Google na przykład opiera się na rozległym badaniu opinii publicznej. Używając jej rzadko zastanawiamy się, w jaki sposób wśród setek tysięcy stron znalezionych w odpowiedzi na naszą kwerendę kilka pierwszych zawiera zwykle najbardziej przydatne informacje. Otóż o tym, jakie strony wypływają na wierzch, decydują sami niezliczeni użytkownicy wyszukiwarki dokonując subiektywnych ocen, która ze znalezionych stron zasługuje na to, by zrobić sobie na niej zakładkę.” (Krzysztof Szymborski: *Jarmark idei*, Polityka, „Niezbędnik inteligenta” nr 1).

W rzeczywistości taki system wyszukiwawczy jest o wiele bardziej wyrafinowany. Na przykład pozycja strony w wyniku wyszukiwania w systemie Google nie

zależy, tak jak to się dzieje w wielu innych wyszukiwarkach, od liczby wystąpień szukanych wyrażień na danej stronie, ale od oceny, jaka została przyznana stronie przez w pełni automatyczny algorytm PageRank (oczywiście strony tych wyrażień nie zawierające nie są tu brane pod uwagę). Algorytm PageRank przyznaje stronom oceny na podstawie liczby linków prowadzących do nich od innych stron. Wskaźnik ten wydaje się wartościowy i obiektywny, gdyż strony wartościowe i popularne prawdopodobnie zdobędą uznanie autorów innych stron internetowych o podobnej tematyce, a ci umieszczają do nich odsyłacze w swoich serwisach. Dodatkową obiektywność zapewnia to, że odsyłacze prowadzące do ocenianej strony są poza kontrolą autora tej strony, nie może więc on oszukać systemu. System bierze pod uwagę nie tylko liczbę odsyłaczy prowadzących z innych stron do ocenianej strony, ale także uwzględnia oceny wcześniej im przyznane, w wyniku czego linki prowadzące od dobrze ocenionych stron mają większe znaczenie przy wyznaczaniu oceny strony. Dodatkowo na wagę odsyłacza prowadzącego do ocenianej strony wpływa liczba innych odsyłaczy znajdujących się na stronie będącej źródłem rozważanego odsyłacza. Im więcej jest takich odsyłaczy, tym mniej wnoszą one do oceny innych stron. Ma to na celu wyeliminowanie wpływu stron zawierających sztucznie stworzone długie listy odsyłaczy. Na ocenę stron mają wpływ odsyłacze prowadzące od powszechnie znanych i cieszących się zaufaniem użytkowników katalogów internetowych, takich jak katalog stron Yahoo, które swoją wysoką oceną „obdarowują” wiele innych stron. Jednak najbardziej liczącymi się w wycenianiu stron są linki prowadzące od dobrze ocenionych stron zawierających niewiele linków.

Wpływ na ocenę serwisu internetowego ma także jego wielkość, tzn. im więcej różnych stron składa się na cały serwis, tym wyższą będzie on miał ocenę. Mechanizm oddziaływania wielkości serwisu na ocenę jest taki sam jak przy odsyłaczach pochodzących z innych serwisów, każda strona serwisu oceniana jest bowiem oddzielnie. Prawdopodobnie (bo Google nie ujawnia takich szczegółów) każda strona jest traktowana zupełnie oddzielnie, a więc wewnętrzne odsyłacze łączące różne strony serwisu także przyczyniają się do wzrostu oceny. Ma to sens, gdyż istnieje duże prawdopodobieństwo, że strona należąca do rozbudowanego serwisu rzeczywiście ma większą wartość niż strona małego serwisu. Jednak dla naprawdę popularnych serwisów ten wewnętrzny czynnik ma mniejsze znaczenie, gdyż liczba odsyłaczy prowadzących do nich z zewnątrz na pewno przewyższa liczbę odsyłaczy wewnętrznych, ma więc także większy wpływ na ocenę strony.

Oceny stron w systemie Google można obejrzeć, korzystając z dodatku Google Toolbar do przeglądarek internetowych (Internet Explorer i Firefox). Na tym dodatkowym pasku narzędziowym wyświetlana jest ocena aktualnie oglądanej strony w skali od 1 do 10 (0 dla nie ocenionych stron).

Mechanizmowi działania tej internetowej wyszukiwarki celowo poświęcałam więcej uwagi, chciałam bowiem pokazać, że samoorganizowanie się takich dużych systemów informacyjnych podporządkowane jest przede wszystkim potrzebom informacyjnym i możliwościom percepcyjnym użytkownika informacji, zapewniając równocześnie obiektywność i optymalność – co do wagi, relewancji i odpowiednio wyważonej kompletności – wyszukiwanej informacji.

Wyszukiwanie informacji w internetowych systemach może być dodatkowo optymalizowane na podstawie wyników monitorowania przez system zachowań użytkowników, co w tradycyjnych, nawet stosunkowo niewielkich systemach informacyjno-wyszukiwawczych było trudne. Do wypracowania odpowiednich narzędzi przyczynił się przede wszystkim rozwój handlu elektronicznego w ostatnich latach. Duże popularne serwisy internetowe są w stanie wyprodukować dziennie nawet kilka gigabajtów danych o zachowaniu się odwiedzających je internautów. Najlepszym sposobem przetworzenia i zanalizowania tych danych jest zastosowanie systemów OLAP (OnLine Analytical Processing). Analizowane dane mogą zawierać informacje o lokalizacji komputera internauty, porze odwiedzin, ścieżce, jaką poruszał się internauta, przeglądając serwis, o obiektach, które oglądał, o tym, czy dokonał transakcji (plus szczegóły transakcji), jakich słów kluczowych użył do wyszukiwania, i wiele innych. Aby dodatkowo poprawić przydatność takich analiz, dane te powinny być połączone z również gromadzonymi przez system danymi o poprzednich zachowaniach informacyjnych użytkownika (historia wyszukiwań) oraz o innych relewantnych dla działania samego systemu atrybutach.

Tak więc i internetowy system informacyjny i jego użytkownik jest „sam sobie sterem, żeglarzem, okrętem” i żadne specjalnie tworzone języki informacyjno-wyszukiwawcze, ani też opieka ze strony informacji naukowej nie jest tu już potrzebna.

Nowoczesne systemy informacyjno-wyszukiwawcze są systemami hipertekstowymi, umożliwiającymi użytkownikowi wyszukiwanie informacji nie tylko w zbiorze informacji dokumentacyjnych, ale również w zbiorze tekstów (dokumentów źródłowych), do których metainformacje systemu dokumentacyjnego odsyłają. Umożliwiają użytkownikowi dotarcie do interesujących go informacji zawartych w tekstach nie tylko drogą pośrednią przez dokument pochodny, ale bezpośrednią, dzięki sieci łączącej teksty lub ich fragmenty zawierające informacje powiązane ze sobą określonymi relacjami semantycznymi lub asocjacyjnymi. Tworzenie systemów hipertekstowych można zautomatyzować, korzystając z relacji intertekstualnych bezpośrednio wyrażonych w tekstach lub ich aparatach metatekstowych, takich jak przypisy, bibliografia, innego rodzaju odesłania, a także znajomości struktury tekstu spójnego, znajomości struktury pola semantycznego dziedziny wiedzy oraz struktury paradygmatycznej języka. I tu niezbędne jest więc korzystanie z wyników badań językoznawstwa, przede wszystkim semantyki i lingwistyki tekstu.

Tej samej wiedzy wymaga konstruowanie dobrych wyszukiwarek informacji w rozległych sieciach informacyjnych, podpowiadających użytkownikowi optymalną strategię wyszukiwawczą, uwzględniającą jego potrzeby informacyjne i możliwości w akwizycji informacji.

Nowoczesne systemy informacyjne mają być przyjazne dla użytkownika. Przyjazne, to znaczy także łatwo dostępne i znoszące wszelkiego rodzaju bariery informacyjne utrudniające użytkownikowi dostęp do zbioru wyszukiwawczego i zbioru informacyjnego systemu. Kluczem otwierającym system jest zawsze język: sztuczny, informacyjno-wyszukiwawczy, w wypadku zbioru dokumentacyjnego, naturalny, w wypadku zbioru informacyjnego. W systemach hipertekstowych integrującym oba te zbiory językiem dostępu, umożliwiającym korzystanie

z zawartych w nim informacji każdemu użytkownikowi, powinien być jego język naturalny. Konieczne jest więc skonstruowanie odpowiednich algorytmów tłumaczących – w najprostszym przypadku słowa kluczowe używane w wyszukiwarkach, w przyszłości być może całe frazy i teksty języka naturalnego. Algorytmów, których tworzenie musi poprzedzać formalizacja systemów językowych oraz pragmatycznej ekwiwalencji tekstów różnych języków. A to już pole badań językoznawców, z których rezultatów informacja naukowa może i powinna korzystać.

W informacji na naszych oczach dokonuje się rewolucja: skokowe przejście od coraz bardziej wyspecjalizowanych systemów informacyjnych do uniwersalnego systemu informacyjnego, jakim jest Internet, a od coraz bardziej „wymyślnych” i rozbudowanych języków informacyjno-wyszukiwawczych powrót do używania języka naturalnego w funkcji metainformacyjnej, czego przejawem jest używanie w wyszukiwarkach słów kluczowych.

Warto więc zastanowić się nad tym, czy we współcześnie rozwijających się systemach informacyjnych języki informacyjno-wyszukiwawcze będą jeszcze potrzebne, czy trzeba je jeszcze tworzyć, doskonalić i rozwijać. A może to już koniec języków informacyjno-wyszukiwawczych?

Na to pytanie moja odpowiedź jest pozytywna, choć może nie jest pozytywna dla tych, którzy językami tymi się zajmują.

Tak, to już koniec języków informacyjno-wyszukiwawczych rozumianych tak, jak to pokazałam w pierwszej części mojego wystąpienia. Nie uratują ich również tak modne obecnie w informacji naukowej tzw. ontologie, które mają usprawnić wyszukiwanie informacji i poprawić wszystkie jego parametry. Sądzę, że ontologie – wbrew temu, co głoszą ich pomysłodawcy i twórcy - nie są w językach informacyjno-wyszukiwawczych żadną nową jakością. Pomysł polega po prostu na wpisaniu w paradygmatykę języka szeroko rozumianych relacji hierarchicznych, mereologicznych i asocjacyjnych (w tym intratekstualnych i intertekstualnych), co było już odwzorowywane w tezaurusach bardziej rozbudowanych języków deskryptorowych, a i wcześniej funkcjonowało chociażby w katalogach w postaci rozbudowanej sieci odsyłaczy.

Wszystko to oczywiście nie znaczy, że istniejące języki informacyjno-wyszukiwawcze nagle przestaną być używane. W bibliotekach ostaną się języki haseł przedmiotowych, które językami haseł przedmiotowych są już tylko z nazwy, bowiem komputerowe przetwarzanie haseł przedmiotowych i możliwość wyszukiwania według dowolnego elementu charakterystyki wyszukiwawczej dawno już zmieniło ich prekoordinacyjny charakter. Ostanie się zapewne także Klasyfikacja Kongresu, choćby ze względu na tradycję i chęć zachowania jednorodności w opracowywaniu zbiorów. Również z Uniwersalnej Klasyfikacji Dziesiętnej nie zrezygnują tak łatwo powołane do sprawowania nad nią pieczy gremia (to prawo natury ogólnej, głoszące, że instytucje powołane do robienia czegoś łatwo nie rezygnują i same generują swoje zadania nawet wtedy, gdy produkty ich działania przestają być użyteczne). Ale na pewno nie warto już wkładać wysiłku ani w tworzenie nowych języków informacyjno-wyszukiwawczych, ani w dalszą rozbudowę już istniejących, bo:

– i tak ich słowniki rozrosły się do rozmiarów, nad którymi już nie panują ich użytkownicy – używanie w procesie indeksowania jednostek leksykalnych o bardzo wąskiej denotacji (słownictwo specjalistyczne), tak jak to się dzieje

w systemie KABA, nie będzie zdyskontowane w procesie wyszukiwania, bo poszukujący informacji, żeby ich użyć, musieliby już w momencie rozpoczęcia wyszukiwania dysponować tą wiedzą, której dopiero poszukują;

– zdyskontowanie rozbudowy języków informacyjno-wyszukiwawczych, a właściwie wprowadzenia do nich jakichkolwiek zmian wymaga natychmiastowej reklasyfikacji całych zbiorów, co nigdy nie jest czynione, bez tego bowiem wcześniej wprowadzone do zbioru wyszukiwawczego charakterystyki wyszukiwawcze dokumentów generują i szum, i ciszę informacyjną w systemie, czego użytkownik nie jest świadomy – wyszukiwanie bowiem odbywa się zazwyczaj według aktualnej wersji języka na tekstach realizujących jej wcześniejsze stadia (taka reklasyfikacja jest oczywiście nieopłacalna);

– wiedza rozwija się tak szybko, że języki informacyjno-wyszukiwawcze za nią nie nadążają, a przecież użytkownicy informacji w przeważającej części poszukują informacji aktualnych. Dla nich liczy się przede wszystkim aktualność i czas dostępu do informacji, mniej ważna jest pertynencja, a najmniej ważna kompletność wyszukanej informacji. Nikt już nie będzie cierpliwie czekał, aż specjaliści wprowadzą zmiany do języka informacyjno-wyszukiwawczego i przygotują informację pochodną po uprzedniej dokładnej analizie treści dokumentu. Wszystkiego na zadany temat i tak nikt nie przeczyta, a pozycjonowanie wyszukanej informacji zapewniające relewancję i pertynencję system musi zapewniać automatycznie. Tak więc w nowoczesnych systemach informacyjnych wyszukiwanie informacji musi się odbywać w procesie przetwarzania informacji utrwalonych w języku naturalnym, co już się dzieje w Internecie (dokładniej omawiałam to w poprzednim moim wystąpieniu, pokazując, jak działa wyszukiwarka Google), a będzie to możliwe dzięki programom komputerowym wykorzystującym efekty formalizacji języka naturalnego i dokonującym takich formalizacji, także w zakresie semantyki i relacji paradygmatycznych umożliwiających wyszukiwanie informacji na pożądanym poziomie szczegółowości bądź ogólności oraz relacji tekstowych zapewniających relewancję i kompletność informacji. Oczywiście funkcja metainformacyjna tych systemów nie zniknie, tyle tylko że metainformacje zapisane będą w języku naturalnym, a nie w specjalnym sztucznie stworzonym języku, na podstawie relacji między wyrażeniami języka naturalnego oraz reprezentowanymi przez nie dokumentami będzie też prowadzone wyszukiwanie. Funkcję pełnią dotychczas w systemach informacyjnych przez języki informacyjno-wyszukiwawcze przejmie więc język naturalny - funkcja metainformacyjna i metatekstowa jest zresztą jedną z funkcji języka naturalnego w komunikacji odbywającej się w tym kodzie, bo człowiek zawsze udzielał nie tylko informacji, ale i informacji o informacjach, czyli metainformacji – ale, zgodnie z definicją dotychczas obowiązującą w informacji naukowej, nie będzie to już sztuczny język informacyjno-wyszukiwawczy.

Wykorzystano teksty:

Bojar, Bożenna. *Związki teorii języków informacyjno-wyszukiwawczych z językoznawstwem* (referat wygłoszony przed wieloma laty w Bibliotece Narodowej, tekst niepublikowany).

Bojar, Bożenna (2005). *Czy to już koniec informacji naukowej?* (referat wygłoszony na VIII Forum Informacji Naukowej i Technicznej, Zakopane 2005).

Bojar, Bożenna i inni (2002). *Słownik encyklopedyczny informacji, języków i systemów informacyjno-wyszukiwawczych*. Warszawa: Wydawnictwo SBP.

Dembowska, Maria, red. (1979). *Słownik terminologiczny informacji naukowej*. Wrocław: Ossolineum.

Żmigrodzki, Zbigniew; Babik, Wiesław; Pietruch-Reizes, Diana (2006). *Informacja naukowa: rozwój, metody, organizacja*. Warszawa: Wydawnictwo SBP.

Summary

Information-retrieval languages (IRL) has been a topic of information science research for many years, because of the latter concentration (either in theoretical or practical aspects) on information-retrieval systems (IRS). The article answers the questions concerning the future of information-retrieval languages. In particular it focuses on: (a) defining IRL, (b) IRL functioning in different types of IRS, (c) IRL dependence on users' information needs, as well as possibilities of technical devices being used in information processing, (d) development of IRL theory.