

POLSKA AKADEMIA NAUK

OŚRODEK INFORMACJI NAUKOWEJ

PL ISSN 0324-8194

# ZAGADNIENIA INFORMACJI NAUKOWEJ

1989

OSSOLINEUM

Nr 1 (54)

POLSKA AKADEMIA NAUK

---

OŚRODEK INFORMACJI NAUKOWEJ

# ZAGADNIENIA INFORMACJI NAUKOWEJ

1989

Nr 1 (54)

---

ZAKŁAD NARODOWY IMIENIA OSSOLIŃSKICH  
WYDAWNICTWO POLSKIEJ AKADEMII NAUK

Rada Redakcyjna  
Zdzisław Bobiatyński, Bożenna Bojar, Władysław M. Grabski,  
Andrzej Gromek, Małgorzata Klossowska, Barbara Krygier,  
Juliusz L. Kulikowski, Anna Leśniewicz (sekretarz redakcji),  
Bronisław Ługowski (redaktor naczelny), Elżbieta Malinowska,  
Maria Szomańska, Olgierd A. Wojtasiewicz

Do roku 1971 czasopismo ukazywało się pod tytułem\*  
„BIULETYN ODIIN PAN”

Adres Redakcji

Ośrodek Informacji Naukowej PAN  
00-330, ul. Nowy Świat 72 (Pałac Staszica)

Maszynopis niniejszego numeru  
przekazano do Wydawnictwa 16 VI 1989

Wykonano ze składopisu dostarczonego przez OIN PAN

PL ISSN 0324-8194

Zakład Narodowy im. Ossolińskich – Wydawnictwo, Wrocław 1989 r.

Nakład 640 egz. Objętość erk. wyd. 8,70; ark. druk. 12,25; erk. A1 – 16,4. Papier  
offset., kl. III, 70 g, 61 × 86. Składopis przekazano do druku i druk ukończono w  
listopadzie 1989 r. Zam. 694/89. Cena zł 1300.–

Zakład Graficzny Wydawnictw Naukowych w Łodzi, ul. Żwirki 2

ELŻBIETA ARTOWICZ

Katedra Filologii Węgierskiej  
Wydział Neofilologii UW

## INTERPRETACJE TERMINU "RELEWANCJA" W INFORMACJI NAUKOWEJ

Geneza i pierwsze definicje relewancji. Relewan-  
cja jako relacja między elementami systemu in-  
formacyjno-wyszukiawczego oraz jako relacja  
między zbiorem informacyjnym systemu i otocze-  
niem systemu. Relewancja w świetle zasad eks-  
plikacji semantycznej R. Carnapa.

W literaturze poświęconej zagadnieniom sprawności funkcjonowa-  
nia systemów informacyjno-wyszukiawczych (SIW), a w szczególności  
kwantytatywnej ocenie wyników wyszukiwania informacji w systemach  
dokumentacyjnych termin relewancja służy jako podstawowa kategoria,  
do której odwołują się autorzy poszczególnych prac. Pierwsze użycie  
tego terminu w informacji naukowej przypisuje się S.C. Bradfordowi,  
który posłużył się w latach trzydziestych określeniem "relewantny",  
kwalifikując artykuły z czasopism naukowych jako "relewantne"  
bądź "nierlewantne". Miało to miejsce w sytuacji coraz bardziej  
drastycznego kryzysu komunikacyjnego w nauce spowodowanego:

- wzrostem liczby publikacji naukowych związanych z ilościowym roz-  
wojem kadr naukowych,
- brakiem lub niedoskonałością kryteriów selekcji dokumentów w roz-  
rastających się zbiorach bibliotecznych,

---

"Zagadnienia Informacji Naukowej" 1989 nr 1(54)

- równocześnie zachodzącymi procesami dezintegracji (specjalizacji) i integracji w nauce (powstawanie interdyscyplinarnych dziedzin wiedzy).

Kryzys ten rozpoczął się zdaniem niektórych badaczy już w wieku XVII, w momencie powstania pierwszych czasopism naukowych, które atakowały element mechanizmu selekcjonowania i rozpowszechniania informacji w środowiskach specjalistów zajmujących się określonymi problemami naukowymi. Były one równocześnie środkami integrującymi informacyjnie te środowiska /23 , 21/.

Dyskusje nad metodami efektywnej selekcji "relewantnej" informacji rozpoczynają się faktycznie w latach czterdziestych i pięćdziesiątych naszego wieku wraz z powstawaniem SIW jako komplementarnych wobec bibliotek mechanizmów komunikowania wiedzy. J.W. Perry i A. Kent wprowadzają wówczas propozycje jednostek kwantytatywnego pomiaru sprawności wyszukiwania "recall" (kompletność wyszukiwania) oraz "precision" (dokładność, trafność wyszukiwania) /31/. Kryteria te, a także inne, przedstawione w późniejszym okresie tzw. współczynniki efektywności wyszukiwania: szumu, ciszy, efektywności, czułości systemu (sensitivity) oraz precyzji (specifity) definiowano na podstawie rozróżnienia dokument relewantny - nierelewantny oraz wydany - niewydany przez system /25/. Przeprowadzono - zwłaszcza w latach pięćdziesiątych - sześćdziesiątych ogromną liczbę testów i eksperymentów poświęconych pomiarom sprawności wyszukiwania informacji. Doczekały się one obszernej literatury zanalizowanej i omówionej m.in. przez T. Saracevica /25/. Uwagę metodologów i praktyków informacji absorbowały w pierwszej kolejności metody oceny sprawności wyszukiwania, natomiast wyniki - jak twierdzi Saracevic - w stopniu nieco mniejszym. Kryterium relewancji przyjmowano początkowo jako "oczywiste", intuicyjnie zrozumiałe - coś, co ma znaczenie w efektywnej komunikacji między

użytkownikiem a systemem informacyjno-wyszukiwawczym, co jest zre-  
latywizowane do wiedzy, sposobu pytania, autora pytania, jego po-  
trzeb informacyjnych oraz sytuacji, w jakiej odbywa się proces wy-  
szukiwania i jego ocena. Analizowano czynniki wpływające na orze-  
kanie o relewancji, dokonywano jej pomiarów, a równocześnie ignoro-  
wano w zasadzie dorobek innych dziedzin filozofii, logiki w zasad-  
niczej kwestii, jaką była definicja relewancji.

Refleksja nad potrzebą sformułowania "dobrej" i "mocnej"  
definicji relewancji pojawia się w środowiskach informacyjnych na  
przełomie lat pięćdziesiątych i sześćdziesiątych. Podejmowane wów-  
czas próby sformułowania definicji relewancji przyjmują formę hipot-  
tez wyliczających możliwe do uwzględnienia argumenty relacji między  
elementami procesu komunikacyjnego, które można w uproszczeniu o-  
kreślić za pomocą formuły. RELEWANCJA JEST TO

CÓŚ (miara, stopień, wymiar, ocena, relacja) - (A)

CZEGOŚ (odpowiedniości, użyteczności, związku, usatysfakcjonowania,  
dostosowania, zbieżności) - (B)

ISTNIEJĄCE MIĘDZY CZYMŚ (dokumentem, artykułem, formą tekstu, in-  
formacją, faktem) - (C)

oraz CZYMŚ (zapytaniem informacyjnym, instrukcją wyszukiwawczą, wy-  
korzystywaną informacją, potrzebą informacyjną, punktem wi-  
dzenia) - (D)

OKREŚLONE PRZEZ PEWNEGO (osobę, użytkownika, autora sądu, autora  
zapytania, specjalistę w dziedzinie informacji) - (E) /25/. Inter-  
pretacja znaczenia terminu podstawowego w ocenie wyszukiwania infor-  
macji oraz w projektowaniu języków informacyjno-wyszukiwawczych ze  
względu na zakres i rodzaj wiedzy w nich odwzorowywanej obciążona  
była wieloma błędami, które można ogólnie określić jako niespeł-  
nianie warunków poprawnej eksplikacji znaczeń terminów naukowych,  
wyrażające się w:

- posługiwaniu się w definicjach terminami semantycznie złożonymi lub nieostryimi wyrażeniami języka potocznego, jakkolwiek formułowane definicje traktowano aksjomatycznie,

- braku określenia związków semantycznych między terminami występującymi w definicjach, denotującymi różne elementy i układy procesu komunikacji między użytkownikiem i SIW.

Należy także zwrócić uwagę na fakt, iż większość autorów interpretowała relewancję jako relację wieloargumentową, natomiast ocenę dokumentów wyszukanych przez SIW traktowano w kategoriach celu absolutnego, o którego osiągnięciu rozstrzyga się binarnie. Dokument wyszukany powinien być relewantny, niewyszukany - jest nierelewantny.

Przyczyn niekonsekwencji i problemów z definiowaniem relewancji w systemach informacji należałoby upatrywać nie tylko w niedostatkach metodologicznych prac poświęconych temu zagadnieniu. Dyskusje nad kwantytatywną oceną wyników wyszukiwania informacji na podstawie potocznego, "zdroworozsądkowego" kryterium relewancji słabną z początkiem lat siedemdziesiątych bądź też przenoszą się w sferę rozważań teoretycznych. Okres ten stanowi pewną cezurę w rozwoju systemów informacyjnych, oznacza bowiem początek rozwoju ogólnie dostępnych dialogowych systemów informacji (on-line). W systemach tradycyjnych, do których należą zarówno tzw. systemy manualne jak i skomputeryzowane, funkcjonujące w trybie wsadowym, dialog z użytkownikiem powinien być doprowadzić do uzyskania odpowiedzi relewantnej we wszelkich możliwych aspektach. Użytkownik nie musi bowiem możliwości śledzenia poszczególnych etapów wyszukiwania i ingerowania w jego przebieg (zmiany strategii wyszukiwawczej). W systemach dialogowych (interakcyjnych) pojawiają się możliwości ingerencji w proces wyszukiwania czyli zmiany sformułowania instrukcji wyszukiwawczej lub całej strategii, co oznacza możliwość

zmiany kryteriów relewancji określonych w wyjściowej wersji instrukcji wyszukiwawczej. Wprowadza się wówczas inne kryteria oceny informacji wydanej przez system, między innymi pertynencję.

Przyczyn sporów nad znaczeniem terminu relewancja i źródeł trudności z jej zdefiniowaniem należy upatrywać także w złożoności procesów komunikacji i wieloznaczności terminu komunikacja, które można opisywać w aspekcie aplikacyjnym lub teoretycznym. Ujęcie aplikacyjne dominuje w początkowym okresie dyakusji, jakkolwiek należy podkreślić, że wprowadzeniu przez Bradforda terminu relewancja przyświecał cel teoretyczny, a mianowicie formułowanie kwantytatywnych prac komunikacji w nauce związane z powstawaniem bibliometrii i naukometrii.

Przegląd różnych ujęć metodologicznych i definicji relewancji charakterystycznych dla lat pięćdziesiątych, sześćdziesiątych i początku lat siedemdziesiątych przeprowadził T. Saracevic, do którego pracy odwołujemy się na wstępie jako ilustrującej problem w stopniu najbardziej reprezentatywnym. Jako cechę znamioną dla poszczególnych postaw metodologicznych wymienia swoisty izolacjonizm, przejawiający się w ignorowaniu istniejących już interpretacji oraz jednostronność często połączoną z nadmiernym pragmatyzmem. Brak teoretycznego rozwiązania problemu relewancji uznaje za jedną z najpoważniejszych (jeśli nie największą) przeszkodę w rozwoju SIW, stwierdzając: "...Historia nauki dostarcza wielu dowodów na to, iż nie ma nic bardziej praktycznego niż mocna teoria /../. Brak teorii prowadzi najczęściej do zgadywanek z niewielkim prawdopodobieństwem pozytywnego wpływu"<sup>1/</sup>. Autor ten podjął także próbę przedstawienia

---

<sup>1/</sup> ... The history of science provides numerous proofs that there is nothing more practical than a good theory. /../Lack of theory most often leads to guessing games with little probability of positive impact... (s. 337-338).



własnej klasyfikacji, a właściwie typologizacji interpretacji relewancji w informacji naukowej, przyjmując za punkt wyjścia dwa pytania:

- jakie czynniki przyjmowano w danej definicji jako podstawę opisu znaczenia terminu,
- jakie relacje specyfikowane są w definicji terminu?

W wyniku analizy definicji i poglądów poszczególnych autorów na relewancję wymienia następujące elementy systemu komunikacji w nauce uwzględnianie w eksplikacji tego terminu, stanowiące odpowiedź na pytanie pierwsze:

- wiedza przedmiotowa,
- literatura przedmiotu,
- języki reprezentacji wiedzy,
- zbiory informacyjne systemu,
- adresaci (użytkownicy) informacji wraz z ich wiedzą przedmiotową,
- system informacyjno-wyszukiwawczy,
- otoczenie użytkownika (environment) czyli realia wykorzystania informacji,
- systemy wartości stanowiące punkt odniesienia przy orzekaniu o relewancji informacji wydanej przez system.

Sformułowanie pytania drugiego implikuje interpretację relewancji jako pewnej relacji między przyjętymi elementami procesu komunikacji. Jednakże zaproponowana przez Saracevica definicja relewancji wydaje się zbyt ogólna i obciążona przy tym wadami wymienionymi wyżej, a polegającymi na stosowaniu terminów nieostrych, jak sam stwierdza trudno definiowalnych. Według tego autora .."relewancja jest miarą efektywności kontaktu między źródłem informacji a odbiorcą w procesie komunikacji..", jest częścią wiedzy, informacji i komunikacji. Zarzut definicji zbyt szerokiej wynika przede wszystkim z posługiwania się terminami, których znaczenie próbuje u-

ściślić, czyni to jednak niekonsekwentnie. Zaznaczając, że w tezaursie Rogeta pole semantyczne terminów "komunikacja" i "informacja" zawiera ponad sto terminów skojarzonych, ogranicza się do podania "organizacji" jako cechy odróżniającej komunikację od informacji. Organizacja odwzorowuje - jego zdaniem - proces nawarstwiania się wiedzy (sedimentation). Rozróżnienie to traci moc w przypadku samouczących się inteligentnych systemów informacji, a ponadto należy zauważyć, że cechą każdego zbioru informacji jest jakieś zorganizowanie, różniące się stopniem i/lub sposobem. Równie enigmatycznie określone są związki między wiedzą a literaturą przedmiotu ... "są powiązane lecz także dość odmienne /../, różnią się strukturą, dynamiką, łączy je cecha redundancji"... /25, s.332/. Podobnie nieostry terminem w przytoczonej definicji jest "efektywność kontaktu", o której nie wiadomo, jakiej sfery dotyczy: wiedzy, informacji czy też działań i jak jest interpretowana - binarnie czy relatywnie. To zbyt szerokie i nieostre ujęcie relewancji jest ponadto niespójne z uwagami autora na temat innych definicji i podsumowującym stwierdzeniem o istnieniu różnych rodzajów (systemów) relewancji, zorganizowanych w kompleksowe nadsystemy, które przypuszczalnie posiadają pewne własności uniwersalne, różniąc się cechami specyficznymi dla każdego z nich.

Seracevic wyróżnił kilka ujęć relewancji, uwzględniając częściowo chronologię ich powstawania. Są to ujęcia eksponujące aspekty:

- wiedzy przedmiotowej, w którym relewancja jest relacją między wiedzą o przedmiocie lub na temat przedmiotu (on or about the subject) i pytaniem lub problemem dotyczącym przedmiotu (question or topic),
- literatury przedmiotu, w którym relewancja jest relacją między literaturą przedmiotu i pytaniem dotyczącym przedmiotu,

- **logiczny**, w którym relewancja jest relacją między wnioskiem i przesłanką we wnioskowaniu dedukcyjnym lub probabilistycznym,
- **systemowy**, w którym relewancja jest relacją między systemem informacyjnym oraz pytaniem lub literaturą przedmiotu lub użytkownikiem,
- **odbiorcy informacji**, w którym relewancja jest oceną, orzeczeniem o relacji między zapytaniem i dokumentem, sformułowaną przez człowieka,
- **pertynencji**, a więc odbiorcy informacji, w którym relewancja jest relacją między zbiorem wiedzy użytkownika i wiedzą przedmiotową lub literaturą przedmiotu,
- **pragmatyczny**, w którym relewancja jest relacją między problemem rozwiązywanym aktualnie przez użytkownika oraz informacją dostarczaną przez system informacyjny przy uwzględnieniu kryterium użyteczności i preferencji jako podstawy orzekania o relewancji.

Kryteria wyróżnienia powyższych ujęć relewancji budzą pewne wątpliwości związane przede wszystkim z nieostrym określeniem elementów procesu komunikacji i ich powiązań. Brak w nich także wyraźnej interpretacji terminu systemu informacyjno-wyszukiwawczy ukazującej jakie elementy układu komunikacji w nauce traktowane są jako część systemu, jakie zaś należą do jego otoczenia. Cechą dystynktywną SIW jest transformacja (przetwarzanie) informacji: wejściowej w odpowiednio ustrukturalizowany zbiór informacyjny systemu, a następnie tegoż zbioru w informację wyjściową drogą selekcyjonowania informacji na podstawie określonej w instrukcji wyszukiwawczej relacji relewancji technicznej. Relacja ta określa kryterium podobieństwa, jakiegó powinny spełniać elementy zbioru informacyjnego do instrukcji wyszukiwawczej, odwzorowującej zapytanie użytkownika, a więc będącej także transformacją tego zapytania. Stąd też w charakterystyce różnych ujęć relewancji uwzględniających różne elemen-

ty procesu komunikacji między systemem informacyjnym i użytkownikiem należałoby oddzielić te definicje, które umieszczają relewancję w sferze relacji:

- między elementami systemu informacyjnego, a więc jego zbiorami i wyrażeniami języka informacyjno-wyszukiwawczego umożliwiającą poezzczególne rodzaje transformacji informacji,

- między systemem informacyjnym (jego zbiorami i wyrażeniami języka odwzorowującymi zapytania informacyjne) i elementami jego otoczenia (użytkownikiem, jego wiedzą, potrzebami informacyjnymi, sytuacją, w której odbywa się proces wyszukiwania informacji). Należy ponadto stwierdzić, że chronologiczna prezentacja różnych definicji relewancji tylko częściowo odpowiada precyzji kryteriów, jakimi kierowali się poszczególni autorzy, stąd też ma w przyjętym układzie znaczenie drugorzędne.

#### RELEWANCJA JAKO RELACJA MIĘDZY ELEMENTAMI SIW

Najstarsza w historii rozwoju skomputeryzowanych systemów informacyjnych jest interpretacja binarna, zgodnie z którą każda odpowiedź systemu spełniająca warunek minimalnej zgodności z instrukcją wyszukiwawczą, a więc zawierająca wyrażenia języka informacyjno-wyszukiwawczego (JIW) równokształtne z wyrażeniami instrukcji jest relewantna. W ujęciu tym, obecnie określanym terminami "relewancja techniczne", "relewancja formalna", "relewancja inżynierska", relewancja jest relacją między planami wyrażania tekstów dokumentów SIW (charakterystyk wyszukiwawczych) i instrukcji wyszukiwawczych. Skala oceny relewancji jest binarna.

Kryterium relewancji **Bradforda** stosowane w bibliometrii oraz w językoznawstwie statystycznym zakłada interpretację relewancji jako funkcji przyporządkującej zbiorom informacyjnym systemu nauki

złożonym z tekstów dokumentów pierwotnych (artykułów) cechą "bycia przedmiotem (tematem) treści" na podstawie wyrażen języka naturalnego (terminów). Skala oceny relewancji jest tu wielostopniowa. Pozwela wyznaczyć w zbiorze informacyjnym pewne podzbiory (strefy) odpowiadające przedziałom funkcji relewancji /1/. Analogicznie w badaniach bibliometrycznych przeprowadzonych przez:

- Lotkę nad dystrybucją częstości prac różnych autorów w poszczególnych dyscyplinach /18/,
- Price'a nad peradygmatem (aieciami) cytowań w literaturze naukowej (21),
- Urquarta nad wykorzystaniem czasopism w bibliotekach naukowych /32/,

wykazywano prawidłowości organizacji zbiorów informacji odwzorowującej działanie zasady relewancji w komunikacji w nauce na podstawie badania relacji między tekstami dokumentów i wyrażeniami języka, będącymi wykładnikami cech relewantnych, na przykład elementami opisu bibliograficznego /24, 21/.

Ilościowego potwierdzenia prawomocności zasady relewancji dostarczają także modele statystyczne języka naturalnego opracowane przez Zipfa wraz ze sformułowaną przez niego tezą o ekonomii wysiłku w komunikacji językowej, popartą analizą częstości wyrażen w tekstach oraz w słowniku. Analiza ta wykazała, że istnieje zależność między częstością użycia wyrazów, ich długością oraz liczbą znaczeń, wiekiem, pochodzeniem i rangą /24/.

Próby powiązania prawidłowości ustalonych przez bibliometrię z relacją relewancji podejmowali Kozakchov oraz Saracevic, który wysunął hipotezę o zdeterminowaniu struktury zbiorów informacyjnych nauki w przekroju diachronicznym przez mechanizm relewancji /17/.

Wśród poglądów na relewancję łączących ją ze sferą relacji między zbiorami informacji a wyrażeniami językowymi wyróżnia się

podgrupa poglądów autorów postulujących konieczność wykorzystania w interpretacji tego terminu aparatu logiki i/lub matematyki.

Interesującą analizę problemów związanych z interpretacją relewancji przedstawił J.D. Hillman /11/. Punktem wyjścia jego rozważań przebieg dyskusji na Międzynarodowej Konferencji na temat Informacji Naukowej w Waszyngtonie w 1958 r. W dyskusji tej podkreślano, iż kierunek poszukiwania "dobrej" definicji relewancji jest wyznaczany przez problem powiązania informacji w systemie, a więc jej organizacji (relatedness). Zarysowały się wówczas dwa nurty w próbie definiowania relewancji:

- reprezentowany przez B.C.Vickery'ego, charakteryzował się ujmowaniem relewancji jako cechy systemów, określanej przez projektanta SIW według formuły odwzorowującej związek między klasami dokumentów (charakterystyk wyszukiwawczych dokumentów) w systemie: "A jest związane z B wtedy i tylko wtedy, jeśli A i B występują razem w pewnym C, a ponadto uwzględnia się wszystkie wywodzące się z nich klasy, należące do pewnego łańcucha o określonej długości". Taka formuła relewancji stanowi podstawę klasyfikacji hierarchicznych. Została określona jako interpretacja słabej relewancji.

- reprezentowany przez Bar-Hillela, a podzielany przez Hillmana eksponował problem interpretacji znaczenia dla potrzeb przetwarzania informacji i odwzorowania związków semantycznych między zapytaniem użytkownika i zbiorem dokumentów SIW, co określono terminem "conceptual relatedness".

Zdaniem Hillmana konieczne jest skonstruowanie teorii relewancji opisującej formalnie relację podobieństwa sądów użytkowników wyrażających ocenę relewancji dokumentów wyszukanych przez system w odpowiedzi na zapytania informacyjne. Na tej podstawie należałoby określić uniwersalną zasadę ustalania klas dokumentów "podobnych semantycznie" (conceptually similar) z uwzględnieniem zmienności

i wielowartościowości relewancji. Sformułowanie teorii relewancji wymagałoby, zdaniem tego autora, rozwiązania następujących problemów:

- a) zdefiniowania klas ekwiwalencji znaczeniowej przy ekstensjonalnej interpretacji znaczenia, przyjmując za podstawę podobne sądy użytkowników uporządkowane parami,
- b) zdefiniowania "powiązania znaczeniowego" (relatedness of concepts) przez określenie możliwości uatalenia klas dokumentów ekwiwalentnych znaczeniowo na podstawie szeroko rozumianej relacji podobieństwa.

Do zbudowania teorii relewancji próbował Hillman wykorzystać logiczną metodę podobieństwa klas przedstawioną przez R. Carnapa w dwu wersjach. W wersji pierwszej metoda ta oparta jest na symetrycznej, zwrotnej relacji częściowej identyczności (part-identity) - R, której dziedzinę stanowi wykaz par jednostek elementarnych (charakterystyk wyszukiwawczych dokumentów w systemie informacyjnym). Warunek związku znaczeniowego (treściowego) spełniony byłby wówczas, gdy:

1. Każda badana para klasy znaczeniowej jest częściowo identyczna (part identical).
2. Jako klasy znaczeń należy traktować największe spośród spełniających warunek, co formalnie definiuje się następująco:

$$(x)(y) [(x \in \alpha \cdot y \in \alpha) \rightarrow R_{x,y}] \cdot (x)(\{y\}) [(y \in \alpha \rightarrow R_{x,y}) \supset x \in \alpha].$$

Ograniczenia metody ustalania relewancji na podstawie relacji częściowej identyczności nazwał Hillman "problemem towarzyszenia" (companionship difficulty)\*powstającym w przypadku klas znaczeniowo niezależnych oraz "niedoskonałej wspólnoty" (imperfect community), gdy każda para zbioru charakterystyk wyszukiwawczych posiada wspólne cechy znaczeniowe, lecz nie posiada ich cały zbiór, na przykład:

1.abc 2.bcd 3.cde 4.def

W drugiej wersji metoda Carnapa zakłada konstrukcję klas znaczeniowo powiązanych na podstawie relacji częściowego podobieństwa (part-similarity) - A. Relacja ta zachodzi między każdą z dwu jednostek elementarnych (charakterystyk wyszukiwawczych dokumentu - ChWD), które posiadają wspólną cechę, bądź też własność jednej z nich jest podobna do własności drugiej jednostki. Relacja częściowego podobieństwa ma służyć do budowy tak zwanych kręgów podobieństwa (similarity circles). Klasa własności  $\beta$  (cech) powinna spełniać następujące warunki:

- powinna należeć w całości do każdego kręgu podobieństwa, który zawiera co najmniej połowę  $\beta$  oraz
- dla każdej jednostki elementarnej  $x$  należącej do  $\beta$  istnieje przy najmniej jeden krąg podobieństwa, do którego nie należy  $x$ , lecz który zawiera w całości  $B$ .

Podobieństwo własności (similarity between qualities) -  $A_q$  zdefiniowane zostało formalnie:

$$A_q = Df \bar{\alpha} \bar{\beta} \{ \alpha \beta \in \text{qual.} \alpha \beta \in A_e \}.$$

- relacja  $A_q$  jest klasą wszystkich  $\alpha$  oraz klasą wszystkich  $\beta$ , gdzie  $\alpha$  i  $\beta$  są klasami podobnymi takimi, że każdy element  $\alpha$  jest częściowo podobny (Ae do każdego elementu  $\beta$ ).

W konkluzji Hillman stwierdza, że przy ekstensjonalnej interpretacji znaczenia nie jest możliwe zdefiniowanie relacji podobieństwa (związku znaczeniowego) dokumentów przy wykorzystaniu metody Carnapa także w wersji drugiej z dwu powodów:

- nie istnieje teoria umożliwiająca organizację i strukturalizację zapytań użytkowników dla wszystkich dokumentów relevantnych ze względu na cechę "przedmiot dokumentu",
- matematycznie dowiedziono, że nie jest możliwe utworzenie klas obiektów podobnych przyjmując za podstawę wspólność predykatów



(predicate sharing);... "liczba predykatów wspólnych dla dowolnej pary rozróżnialnych obiektów jest stała. Zatem jeśli podobieństwo ocenia się na podstawie wspólności predykatów, żadne dwa obiekty nie są do siebie bardziej podobne niż jakakolwiek inna para. Liczba predykatów wspólnych dla dowolnego zbioru  $n$  różnych obiektów jest stała, jakkolwiek zbiór by to był /27/.

Wykazawszy nieadekwatność metody Carnapa do ustalania podobieństwa znaczeniowego klas dokumentów w SIW, a tym samym jej nieprzydatność do sformułowania uniwersalnej teorii relewancji przy ekstensjonalnej interpretacji znaczenia, Hillman wyraził przypuszczenie, że innym obiecującym kierunkiem poszukiwań definicji i teorii relewancji może być badanie indywidualnych podstaw sądów użytkowników o relewancji dokumentów wyszukanych przez SIW.

Autorem, którego poglądy wywarły inspirujący wpływ na interpretację terminu relewancja oraz na rozwój kryteriów oceny sprawności systemów informacyjnych jest W.S. Cooper. W artykule "A Definition of Relevance for Information Retrieval" poddał krytyce dość rozpowszechnioną definicję Cuadry i Kattera, która jest prezentowana wśród ujęć wiążących relację relewancji z elementami otoczenia systemu, a której odmówił statusu definicji /5/. W szczególności krytycznie odniósł się do psychologizującego nurtu w interpretacji relewancji, podkreślając konieczność wprowadzenia następujących rozróżnień i założeń:

- informacja jest zawsze reprezentowana w formie językowej,
- relewancja jest relacją definiowaną dla reprezentacji językowych,
- podstawową jednostką reprezentacji informacji jest zdanie,
- dane zawarte w zbiorze informacyjnym systemu można traktować jako zbiór zdań oznajmujących poprawnie zbudowanego języka (well-formed), na przykład języka rachunku predykatów pierwszego rzędu,

- podstawowym konstruktem logicznym niezbędnym do zdefiniowania relewancji w systemie informacyjnym jest relacja wynikania (implikacji),
- wyszukiwanie informacji w SIW ma charakter inferencyjny i polega na dedukowaniu bezpośrednich odpowiedzi (wniosków) na pytania (przesłanki);

Jako źródło nieporozumień w dyskusjach nad relewancją Cooper wskazuje fakt nierozróżniania poszczególnych aspektów wyszukiwania informacji, a mianowicie:

- utożsamianie potrzeby informacyjnej użytkownika związanej ze sferą stanów psychicznych, niedostępnych bezpośrednio obserwacji z zapytaniem informacyjnym, które jest wstępnym, przybliżonym odwzorowaniem tej potrzeby,.
- utożsamianie zapytania informacyjnego z instrukcją wyszukiwawczą, które tylko niekiedy są identyczne,
- utożsamianie logicznej relewancji informacji z użytecznością informacji. Użyteczność informacji (utility) jest, jak twierdzi, zdefiniowana przez różnorodne czynniki pozasystemowe, spośród których wymienia łatwość orzekania o relewancji (humanly discernible relevance), wartość przypisaną wyszukanej informacji oraz stopień jej wiarygodności (credibility). Użytkownik z zasady nie dowierza precyzji systemu. Podkreśla przy tym, że kryterium użyteczności nie jest nacechowane emocjonalnie, ma charakter neutralny. Celem systemu jest dostarczenie informacji użytecznej, nie zaś zaledwie relewantnej.

Logiczną relewancję informacji określa jako dwuargumentową relację między zdaniami zbioru informacyjnego i zdaniem (zdaniami) odwzorowującymi potrzebę informacyjną użytkownika. ... "Zdanie ze zbioru informacji jest logicznie relewantne dla odwzorowanej potrzeby informacyjnej wtedy i tylko wtedy, gdy jest elementem minimal-

negu zbioru przesłanek pewnego zdania składowego w odwzorowaniu potrzeby /.../. Minimalny zbiór przesłanek jest rozumiany jako najmniejszy z możliwych, to znaczy taki, że usunięcie któregośkolwiek z jego elementów spowoduje, że pozostałe zdania przestaną stanowić logiczną konsekwencję zbioru przesłanek"... /5, s.20/.

Powyższe definicje opatrzone zostały zastrzeżeniami i komentarzem, których zreferowanie wydaje się istotne dla dalszych rozważań nad relewancją:

1. Definicja relewancji sformułowana w języku naturalnym jest jedynie definicją "w zasadzie, ponieważ sprowadza się do zastąpienia terminu relewancja innym, równie nieostrym terminem "implikacja" (wynikanie), niezdefiniowanym dla języka naturalnego, konieczna jest sformalizowana definicja relewancji.

2. Istnieje różnica między wyszukiwaniem informacji w systemach faktograficznych i dokumentacyjnych z punktu widzenia reguł informacji, na jakich opiera się selekcja informacji. Wyszukiwanie informacji faktograficznej ma charakter wnioskowania dedukcyjnego na podstawie zbioru informacji i reguł wnioskowania. W systemie dokumentacyjnym proces ten jest logicznie podobny do wnioskowania probabilistycznego. Zaangażowane są w nim dwa zbiory informacji (systemu i wiedza użytkownika) oraz program wyszukiwawczy. Na początku procesu wszystkie zdania zbioru informacji są - z punktu widzenia systemu - elementami zbioru przesłanek, posiadają jednakowe prawdopodobieństwo bycia relewantnymi. Zmienia się ono w miarę wyszukiwania na mocy kryteriów arbitralnie przyjętych w systemie, bowiem dla każdego zdania składowego instrukcji wyszukiwawczej musi być znaleziony minimalny zbiór przesłanek. Probabilistyczny proces wyszukiwania i proces dedukcji (realizowany w mózgu użytkownika) są, zdaniem Coopera, nierozróżnialne w systemie dokumentacyjnym, co sprawia, że prawdopodobieństwo zbudowania doskonałego SIW

jest prawie żadne.

3. Logiczna definicja relewancji dla potrzeb wyszukiwania informacji ma równocześnie wymiar matematyczny i filozoficzny. W wersji podstawowej jest adekwatna jedynie dla systemów faktograficznych udzielających odpowiedzi binarnych na pytania o rozstrzygnięcie. Przy uwzględnieniu różnic w logicznym charakterze procesów wyszukiwania, zaznacza jej autor konieczność odwzorowania relacji między zdaniami składowymi instrukcji wyszukiwawczej, które tworzą pewne drzewo zależności (component statement tree), co formułuje następująco:

..."Zdanie zbioru informacyjnego (w pamięci systemu lub użytkownika) jest logicznie relewantne dla odwzorowania potrzeby informacyjnej wtedy i tylko wtedy, gdy:

- jest elementem minimalnego zbioru przesłanek dla zdania składowego w drzewie zależności odwzorowującym potrzebę informacyjną oraz

- dla wszystkich zdań składowych w tym drzewie (z wyjątkiem wierzchołka) będących poprzednikami tego zdania istnieją zbiory przesłanek /5, s.37/".

Znaczenie poglądów S.C. Coopera na temat relewancji jest wieloaspektowe:

- sformułowana definicja, jakkolwiek poddawana później krytyce, wyraźnie określa konieczność oparcia teoretycznych podstaw wyszukiwania informacji na aparacie logiki formalnej,
- wyraźnie określona została dziedzina relacji relewancji - wyrażenia (zdania) języka reprezentacji informacji,
- podjęto próbę scharakteryzowania procesu wyszukiwania informacji (jej transformacji) w systemach faktograficznych i dokumentacyjnych w kategoriach odmiennych reguł wnioskowania (dedukcyjnego i probabilistycznego, wywodzącego się z logiki indukcji).

**Probabilistyczne** ujmowanie relewancji reprezentuje jedna z pierwszych prac w informacji naukowej Marona i Kuhna, w której podjęto także próbę parametryzacji relewancji /19/. Zdefiniowali oni relewancję jako wieloargumentową relację między: zapytaniem użytkownika, zakresem tematycznym dokumentu, instrukcją wyszukiwawczą oraz dokumentem dostarczonym przez system. Podstawę procesu wyszukiwania miały stanowić ustalone arbitralnie wzorce odpowiedzi prawdopodobnie relewantnych. Zakładali oni przypisywanie ChWD wartości  $p$  reprezentującej systemową ocenę prawdopodobieństwa relewancji, nazywaną wskaźnikiem relewancji lub statusem wyszukiwawczym (relevance number, retrieval status value), posiadającą ustaloną wartość progową ( $p_0$ ). System wyszukiwał dokumenty, których wskaźnik relewancji przekraczał wartość progową. Kryterium relewancji interpretowane było binarnie, co oznaczało równoczesne zakładanie zawodności systemu.

Wiele analogii z ujęciem Marona i Kuhna wykazuje model wyszukiwania Swetsa, w którym wskaźnik relewancji - nazywany wskaźnikiem pertynencji (pertinence number), ustalano na podstawie funkcji odwzorowującej dystrybucję dokumentów uznanych za relewantne (lub nie-relewantne) /29/.

Z ujęciami Marona i Kuhna oraz Swetsa polemizował Bookstein, zwracając uwagę, że w modelu Marona i Kuhna binarna interpretacja relewancji jest stosowana przy semantycznym kryterium selekcji informacji "bycia przedmiotem/tematem" (topicality). Bookstein wykazał matematyczną nieadekwatność obu modeli i zaproponował tak zwaną operacyjną interpretację relewancji (operational interpretation of relevance) oraz rozróżnienie relewancji systemowej i relewancji dla użytkownika, a ponadto wprowadzenie stopni relewancji ustalane w pewnym jej continuum zależnie od tego, kto podejmuje decyzje - system lub użytkownik. Koncepcja ta nie wnosi w zasadzie

pomysłów interpretacyjnych poza interpretację relewancji w sferze systemu - otoczenie (użytkownik), co jest ekwiwalentne rozróżnieniu relewancji i pertynencji, omawianemu w dalszej części artykułu /2/.

Interesującą koncepcję relewancji wywodzącą się z nurtu logiczno-probabilistycznego przedstawili Goffman i Nevill, autorzy tak zwanej epidemicznej teorii komunikacji w nauce, nawiązującej nazwą do zasady rozprzestrzeniania się chorób przez efektywny kontakt /10/. W teorii tej relewancja jest interpretowana jako relacja ekwiwalencji umożliwiająca podział zbioru informacyjnego systemu na klasy ekwiwalencji. Posługując się aksjomatami teorii miar, autorzy ci wykazali, że z natury probabilistyczny pomiar relewancji powinien uwzględniać także wewnętrzną dynamikę zbioru informacyjnego i wzajemne związki między jego elementami (charakterystykami wyszukiwawczymi dokumentów). Uzasadniali to dynamizmem procesu komunikacyjnego, jakim jest wyszukiwanie informacji, powodującym, że relewancja nie jest stałą cechą jednego podzbioru informacyjnego w SIW. Procedury interrogacyjne mogą bowiem wpływać na zmianę oceny relewancji poszczególnych klas ChWD. Na podstawie twierdzeń prawdopodobieństwa statystycznego, określających zmiany procesu w czasie dla danej populacji wskazano różne warunki, w których zbiór informacyjny ujawnia różne rodzaje powiązań między jego elementami. Zaproponowali termin "warunkowe prawdopodobieństwo relewancji" zdeterminowane przez aktualne powiązania w zbiorze informacji i procedurę interrogacyjną, stwierdzając: "...jakakolwiek miara efektywności dostarczanej informacji musi zależeć od tego, co dotychczas wiadomo"... /10, s.202/. Ujęcie warunkowe relewancji jako relacji prawdopodobnej, ocenianej wielowartościowo i stopniowalnej, mierzonej przy uwzględnieniu relacji między elementami zbioru informacyjnego w danym czasie było pierwszą propozycją teoretyczną łączącą aspekt interakcji w procesie komunikacji ze statycznym

ujęciem relewancji w dystrybucji zbiorów informacji w nauce przedstawionym przez Bradforda.

## RELEWANCJA JAKO RELACJA MIĘDZY ZBIOREM INFORMACYJNYM SYSTEMU I OTOCZENIEM SYSTEMU

Spośród interpretacji relewancji jako relacji między informacją (zbiorem systemu) i elementami jego otoczenia bardzo rozpowszechnione w literaturze lat sześćdziesiątych było sformułowanie Cuadry i Kattera /6/. ... "Relewancja jest zgodnością w kontekście między zapytaniem i artykułem, to jest zakresem, w jakim artykuł obejmuje materiał odpowiedni dla zapytania informacyjnego" ... /.../ Relewancja jest terminem niewyjaśnionym" ... Określenie to, nie spełniające warunków poprawnej definicji jest symptomatyczne dla wspomnianego na wstępie pragmatycznego psychologizującego nurtu w dyskusjach nad relewancją, w którym utożsamiano ocenę relewancji z jej definicją. Spory na temat czynników wpływających na orzekanie o relewancji informacji wydanej przez system były szczególnie intensywne po publikacji raportu z tzw. eksperymentu cranfieldskiego /4/. Niektórzy z autorów kwestionowali w ogóle przydatność kryterium relewancji jako subiektywnego i efemerycznego /20/. Taube twierdził wręcz, że dokładność i kompletność wyszukiwania informacji są kryteriami niedefiniowalnymi, nieprzydatnymi w ocenie SIW /30/. Cuadra i Katter wymieniali następujące czynniki wpływające na orzeczenie o relewancji:

1. cechy badanego dokumentu: tematyka, stopień trudności, związłości, styl,
2. warunki formułowania orzeczenia: czas, organizacja zbioru dokumentów, jego wielkość, sposób sformułowania zadania dla oceniającego,

3. wyrażenia określające sposób orzekania o relewancji,
4. cechy orzekającego: doświadczenie, wykształcenie, postawa. /za:  
23/

Salton zauważa, że różnica zdań wśród orzekających o relewancji powstawała w tzw. przypadkach granicznych, którym przypisywano niski współczynnik relewancji (ibid.).

Próby ustalenia aplikacyjnych kryteriów relewancji były uwikłane w problem zdefiniowania innego, równie nieprecyzyjnego jak zdroworoządkowa relewancja terminu "potrzeba informacyjna" (information need). Praktyka serwisów informacyjnych zawsze wykazywała kłopoty z artikulacją potrzeb informacyjnych przez użytkowników. Włączenie aspektu psychologicznego i socjologicznego do orzekania o relewancji prowadziło do jej interpretacji jako relacji między informacją wydaną przez system a domniemanym stanem umysłu użytkownika uświadamiającego sobie potrzebę informacyjną. Do wprowadzenia pewnego ładu interpretacyjnego i równoczesnego uznania możliwości istnienia różnych odmian relewancji przyczyniła się omówiona praca Coopera oraz inne, m.in. O'Connora, wskazujące konieczność rozróżnienia potrzeby informacyjnej jako stanu umysłu od językowego odwzorowania tego stanu /5/. Później walor porządkujący dyskusje miało ponadto wprowadzenie już w latach sześćdziesiątych kryterium pertynencji jako relacji wiążącej wydaną przez system informację z potrzebą informacyjną odróżnianej od relewancji jako relacji językowej między tekstami języka. Nie wyjaśniało to jeszcze kategorii użyteczności informacji istotnej w ustalaniu pertynencji. /8, 12/

W propozycjach autorów próbujących uściślić kryteria relewancji i pertynencji niewyraźny termin "potrzeba informacyjna" jest zastępowany terminem "zbiór wiedzy", reprezentujący zbiór informacji użytkownika stanowiący punkt odniesienia w ocenie informacji wydanej przez system. Użycie terminu wiedza zamiast informacja nie



jest przypadkowe, wyznacza bowiem inny wymiar opisu relacji między zbiorami (podzbiorami) systemu i elementami rzeczywistości pozajęzykowej. Foskett wprowadza opozycję "public knowledge - private knowledge" /8/. Rozszerza w ten sposób sferę rozważań o sumaryczne zbiory informacji w danej dziedzinie nauki oraz algorytmy jej conceptualnego przetwarzania, będące wynikiem pewnego consensusu specjalistów (np. paradygmat Kuhna), które tworzą wiedzę społecznie akceptowaną - "public knowledge". Wiedzę prywatną, zindywidualizowaną (private knowledge) tworzy pewien podzbiór wiedzy społecznej, różniący się nie tylko wielkością z powodu ograniczeń pojemności pamięci człowieka, lecz przede wszystkim sposobem jej indywidualnego przetwarzania. ... "Relewancja jakiegokolwiek elementu wiedzy oznacza, że można go traktować jako spójny z ogólnym paradygmatem w szerszej dziedzinie, jako uznany i uznawalny element consensusu między specjalistami. Pertynencja oznacza, że jest on zgodny ze szczególnym paradygmatem, jaki jednostka próbuje konstruować we własnym umyśle. /../ Relewancja jest tym, co determinuje zawartość wszelkich serwisów informacyjnych, na przykład służby informacyjne selekcionują materiały według pewnego consensusu, nie zaś według kryteriów jednostkowych. /.../ Pertynencja jest natomiast tym, co determinuje dobór materiałów przez bibliotekarza lub pracownika służby informacyjnej dla konkretnego użytkownika. Dobór ten jest bardziej subiektywny, mniej zgodny z consensusem. Wkład jednostki do rozwoju wiedzy można oceniać w zakresie takim, w jakim udaje się jej transformować to, co dla niej pertynentne w to, co jest relewantne względem istniejącego consensusu".. /ibid./

Ujęcie Fosketta rozwija Kemp, wskazując pary terminów w innych dziedzinach związane z rozróżnieniem "bycie wspólnym, społecznym" (public-ness) i "bycie indywidualnym" (private-ness) analogicznie jak relewancja i pertynencja, np. denotacja - konotacja

w językoznawstwie i psychologii, semantyka - pragmatyka w językoznawstwie, komunikacja formalna - komunikacja nieformalna w socjologii /12/.

W interpretacji relewancji jako relacji między odwzorowaniem w systemie informacyjnym zbioru wiedzy społecznie akceptowanej i między zapytaniem informacyjnym za cechę semantycznie prymarną przyjmuje się "bycie przedmiotem treści" oraz ewentualnie relację między przedmiotami treści, co nazywane jest **relewancją przedmiotową** (topical relevance) /28/. Z kolei w interpretacji pertynencji jako relacji między zbiorem wiedzy społecznie akceptowanej odwzorowanej w systemie informacyjnym i między zbiorem wiedzy jednostki (powstałej w wyniku selekcji i nawarstwiania się wiedzy społecznie akceptowanej) istotne znaczenie ma ocena (wartościowanie - judgement), co określane jest terminem **relewancja interpretacyjna** (interpretational relevance) /ibid./.

Stosunkowo słabe rozpoznanie psychologicznych podstaw wydawania sądów przez człowieka stanowi o nieadekwatności ujęć relewancji w kategoriach relacji między zbiorami wiedzy i jej oceną psychologiczną. Podkreśla to P. Wilson, autor koncepcji tzw. **relewancji sytuacyjnej**, w ramach której do określenia relacji między odpowiedzią systemu informacyjnego i zbiorem wiedzy użytkownika wykorzystane jest pozasystemowe kryterium "użyteczności informacji", wyróżnione przez W. S. Coopera /33/. Punktem wyjścia koncepcji relewancji sytuacyjnej jest definicja relewancji logicznej Coopera, wobec której Wilson wysunął zarzut definicji zbyt wąskiej, uwzględniającej tylko jeden z rodzajów relewancji logicznej związanej z systemami dedukcyjnymi. Zarzut ten wydaje się niezupełnie słuszny, zważywszy przytoczone zastrzeżenia Coopera, który podkreślał nieadekwatność obecnego stanu logiki indukcji wobec potrzeby mocnej definicji relewancji, twierdząc równocześnie, że wniosko-

wania probabilistyczne można interpretować dedukcyjnie, opisując relację implikacji jako stopień prawdopodobieństwa.

Poglądy P. Wilsona związane z relewancją sytuacyjną dotyczą dwu aspektów: jej podstaw logicznych oraz użyteczności informacji. Polemizując z Cooperem w kwestii reguł wnioskowania, na których opiera się wyszukiwanie informacji, twierdzi, że procesy inferencyjne człowieka mają charakter niededukcyjny (non-deductive, non-demonstrative). Jego zdaniem, nawet przy braku mocnej, ogólnej teorii indukcji najmocniejsze uzasadnienie relewancji stanowi logika konfirmacji, której przedmiotem są relacje między przesłankami i wnioskiem w sytuacji, gdy przesłanka podtrzymuje wniosek, lecz nie zachodzi między nimi relacja "dedukowalności" (entailment). Tę odmianę relewancji nazywa Wilson relewancją ewidencyjną (evidential relevance). Można ją określić na podstawie prawdopodobieństwa empirycznego potwierdzenia wniosku na podstawie przesłanek, na potwierdzenie czego przytacza definicję relewancji przyjętą w teorii prawdopodobieństwa Keynesa: "... element informacji I jest relewantny dla wniosku (h) ze względu na przesłanki (e), jeśli stopień konfirmacji (prawdopodobieństwa) (h) przy świadectwie (e) oraz przy I jest większy lub mniejszy niż stopień konfirmacji przy samym (e)"... /13/. Za podstawę w teorii systemów informacyjno-wyszukiwawczych uznaje Wilson relację podtrzymywania (support).

Definiując relewancję sytuacyjną i jej odmiany, Wilson wyróżnia te elementy systemu informacyjnego i jego otoczenia, które, jego zdaniem, są niezbędne do sprecyzowania kryterium użyteczności informacji i równoczesnego wyeliminowania nieprecyzyjnych kryteriów psychologicznych:

- **zainteresowania** użytkownika (concern) wyznaczone przez zbiór atrybutów sytuacji, w jakiej się znajduje. Zbiór ten jest określany

subiektywnie, w postaci takiej, jaka jest identyfikowana przez użytkownika,

- **preferencje** dotyczące alternatyw związanych z zainteresowaniami, a więc ze zbiorem atrybutów sytuacji. Nie zakłada się hierarchizacji preferencji, jak to ma miejsce w teoriach ekonomicznych, a jedynie przyjmuje, że mają charakter holistyczny, nie zaś atomistyczny;

- **zbiór wiedzy użytkownika**, traktowanej jako synonim informacji, a więc bez uwzględniania psychologicznej kategorii przekonań (belief); W zbiorze tym wyróżnia podzbiory "wiedzy ogólnej" i "wiedzy sytuacyjnej". Wiedzę sytuacyjną wyznaczają atrybuty sytuacji w danym czasie;

- **zapytanie informacyjne** odwzorowujące atrybuty sytuacji, a więc zainteresowania użytkownika (concern question);

- **zbiór zdań** stanowiących **możliwe odpowiedzi** na zapytanie (concern set statements);

Ponadto przyjmuje Wilson następujące założenia:

a) istnieje wiele rodzajów relewancji (relevance is not a single notion but many),

b) należy odróżniać relewancję psychologiczną, określającą efekty wykorzystania informacji od relewancji logicznej (evidential), będącej relacją między elementem informacji (item of information), indywidualnym odwzorowaniem rzeczywistości czyli zbiorem wiedzy użytkownika oraz sytuacją użytkownika w tej rzeczywistości;

c) zbiór wiedzy użytkownika pozostaje w pewnym związku ze zbiorem potencjalnych odpowiedzi;

d) zdania tworzące zbiór potencjalnych odpowiedzi powinny spełniać warunek rozłączności i adekwatności tak, aby dokładnie jeden element zbioru odpowiedzi był elementem opisu sytuacji. Spełnienie

tego warunku powinien zapewniać poprawnie skonstruowany język (well formed). (Por. Cooper);

e) warunkiem relewancji sytuacyjnej jest relewancja logiczna; Oznaczając  $I$  - zbiór wiedzy użytkownika w danym czasie, Wilson stwierdza: "...element  $I_j$  jest relewantny sytuacyjnie, jeśli łącznie z innymi elementami zbioru  $I$  jest relewantny logicznie dla pewnego zapytania, to jest, jeśli jest elementem minimalnego zbioru przesłanek w  $I$ , z którego wynika logicznie jedna z możliwych odpowiedzi na to zapytanie"... /33, s.465/

..."element  $I_j$  jest relewantny ewidencyjnie dla pytania, jeśli istnieje możliwa odpowiedź, której wiarygodność (credibility) zostanie zmniejszona lub zwiększona po jego rozważeniu, to znaczy, jeśli istnieje podzbiór elementów  $I_a$  nie zawierający  $I_j$ , którego użycie z  $I_j$  spowoduje, że wiarygodność odpowiedzi będzie mniejsza lub większa niż przy wykorzystaniu samego  $I_a$ . Jeśli potencjalne odpowiedzi na zapytanie tworzą rodzinę wyczerpujących i rozłącznych hipotez, to nowa informacja będzie relewantna ewidencyjnie wówczas, gdy spowoduje zmiany w rozkładzie prawdopodobieństwa wśród elementów tej rodziny, przy czym nie można zakładać z góry warunku rozłączności i adekwatności. /ibid./

Wyróżnia kilka odmian relewancji sytuacyjnej. Jest ona **bezpośrednia** (direct) wówczas, gdy element zbioru wiedzy użytkownika ( $I$ ) jest elementem zbioru atrybutów wyznaczających zainteresowania, natomiast **pośrednia**, gdy element zbioru ( $I$ ) nie jest elementem zbioru zainteresowań. Jako warunek relewancji sytuacyjnej wymieniana jest psychologiczna akceptacja relewancji logicznej, zdeterminowana przez zbiór preferencji. Wyjaśnia to, zdaniem tego autora, fakt, że informacja relewantna logicznie jest kwalifikowana w danej sytuacji jako "interesująca - nieinteresująca, nudna". Preferencje decydują o motywacji, chęci dowiedzenia się czegoś.

Kolejną odmianą relewancji sytuacyjnej, jedną z wielu możliwych, jest relewancja praktyczna, wyznaczana przez cele i plany działania (concern, goals). Ten rodzaj relewancji określany jest w socjologicznej teorii wiedzy "relewancją motywacyjną" (motivational relevance) /28/. W ujęciu socjologicznym wyróżnia się "motywy-czymś spowodowane" (because-of-motives), kiedy zidentyfikowane atrybuty sytuacji powodują zmianę organizacji zbioru informacji jednostki i motywują do dowiedzenia się czegoś, zdobycia nowej informacji, a ponadto "motywy określone przez cel" (in-order-to), kiedy zinterpretowane atrybuty informacji sytuacyjnie relewantnej i interesującej motywują do podjęcia działań, niezależnie od tego czy są prawdziwe, czy fałszywe. Prowadzi to do stwierdzenia, że relewancja informacji jest niezależna od logicznych kryteriów fałszu i prawdy.

Próbę wyjaśnienia zasady działania kryterium użyteczności informacji podjął M. Kochen, który przedstawił sformalizowany opis tzw. funkcji użyteczności informacji (utility function) /14/. Rozróżnia on, za Cooperem, relewancję informacji i użyteczność, umieszczając relewancję w sferze relacji językowych, użyteczność natomiast w sferze relacji pozajęzykowych zachodzących między informacją i systemem preferencji, których zmiana powoduje zmiany w funkcji użyteczności. Definicja funkcji użyteczności ( $u$ ) została sformułowana na podstawie czterech aksjomatów teorii użyteczności odnoszących się do użytkownika  $U$ , danego zapytania  $Q$  oraz każdego dokumentu w zbiorze informacyjnym systemu.<sup>2/</sup>

---

<sup>2/</sup> Skwantyfikowaną definicję użyteczności wprowadzili do teorii gier J. Neumann i O. Morgenstern. Jest obliczana na podstawie współczynników prawdopodobieństwa, gdy każdy wybór dokonywany przez jedną osobę lub zespół w warunkach niepewności posiada określone prawdopodobieństwo, inne niż określone prawdopodobieństwa innych wyborów. Za: T. Pszczołowski: Mała encyklopedia prakseologii i teorii organizacji. Wrocław 1978 s. 265.

- A1. Dla każdego z dwu dokumentów  $(d)$  i  $(d')$  użytkownik  $U$  powinien być zdolny stwierdzić, który z nich preferuje bądź zajmuje wobec nich postawę indyferentną.
- A2. Jeśli użytkownik  $U$  preferuje  $(d)$  zamiast  $(d')$  oraz  $(d'')$ , to preferuje  $(d)$  zamiast  $(d'')$  dla każdego  $(d, d', d'')$ ,
- A3. Jeśli użytkownik  $U$  preferuje  $(d)$  zamiast  $(d')$  oraz  $(d')$  zamiast  $(d'')$ , to potrafi określić prawdopodobieństwo  $P$  takie, że dla wszystkich  $(d, d', d'')$  zajmuje postawę indyferentną między opcją pewnością dla  $d'$  i opcją polegającą na zastosowaniu metody, za pomocą której wyszuka  $(d)$  z prawdopodobieństwem  $P$  oraz  $(d')$  z prawdopodobieństwem  $1 - P$ .
- A4. Jeśli  $U$  zajmuje postawę indyferentną wobec  $(d)$  i  $(d^x)$ , zaś  $(d^x)$  jest zastąpione przez  $(d)$  w pewnej strategii wyszukiwawczej  $(d)$  z prawdopodobieństwem  $P$  lub  $(d'')$  z prawdopodobieństwem  $1 - P$  tak, jak w A3, wówczas  $U$  zajmuje postawę indyferentną wobec tej strategii. Funkcja użyteczności  $u(d; U, Q)$  wyznacza dla każdego dokumentu  $(d)$  wartość, przy danym  $U$  i  $Q$  tak, że jeśli  $U$  preferuje  $(d)$  zamiast  $(d')$ , to  $u(d) > u(d')$ , a spodziewana użyteczność strategii jest spodziewaną użytecznością dokumentów w tej strategii.

Zasadniczą słabością definicji funkcji użyteczności jest odwoływanie się do niejasnego kryterium "jest zdolny", "potrafi", co przyznaje pośrednio sam autor, zaznaczając, że zdobywanie wiedzy przez użytkownika może zmienić porządek preferencji, stąd też wartości funkcji są pośrednio zdeterminowane przez zmiany w organizacji zbioru informacji (wiedzy) użytkownika.

## RELEWANCJA W ŚWIETLE ZASAD EKSPLIKACJI SEMANTYCZNEJ R. CARNAPA

Analiza przedstawionych interpretacji terminu relewancja w teorii systemów informacyjno-wyszukiwawczych prowadzi do konstatacji anelagicznych do sformułowanych w pracach z zakresu filozofii języka podejmujących problem uniwersalnej, zdroworozsądkowej wizji świata odwzorowanej w języku potocznym /16/. Intuicyjne rozumienie zasady relewancji jako "odpowiedniości, zgodności" z najczęściej nieprecyzyjnie wyartykułowanymi potrzebami, oczekiwaniami użytkownika okszuje się równie zwodnicze jak założenie filozofów na temat istnienia "pierwszych zasad danych przez naturę", pozwalających zrekonstruować "jądro" zbioru cech prymarnych i konstytutywnych dla opisu rzeczywistości. Trwające od wieków spory filozofów w sprawie zdrowego rozsądku doprowadziły do osiągnięcia consensusu, wyrażającego się w stwierdzeniu pluralizmu zdrowych rozsądków uwarunkowanych behawioralnie i kulturowo /16/. Dyskusje nad uniwersalną, "mocną" definicją relewancji informacji mające nieporównywalnie krótszą historię, mierzoną zaledwie w dziesięcioleciach nie tylko nie spełniły zamierzonej roli, lecz zostały doprowadzone do podobnego jak w filozofii punktu, tj. stwierdzenia o istnieniu różnych relewancji.

Autorzy przytoczonych prac zgodni byli w jednej kwestii: relewancja jeat relacją dwu- lub wieloargumentową. Jako jej dziedzinę przyjmowano różne elementy procesu komunikacyjnego należące do SIW lub do jego otoczenia. Wart podkreślenia wydsze się fakt, iż niezależnie od tego, jak szeroko próbowano ustalić dziedzinę relacji w sferze pozasystemowej, w większości przypadków dochodzi się do wniosku o konieczności interpretacji relewancji w sferze relacji odwzorowanych w języku informacyjno-wyszukiwawczym w powiązaniu z zasadami inferencji w procesie wyszukiwania informacji. Dznacza to,



podkreślaną przez wielu autorów (Hillman, Cooper, Wilson, Saracevic i in.), konieczność oparcia informacyjnej teorii relewancji na aparacie logiki, przy ewidentnych jego ograniczeniach, sygnalizowanych w prezentowanych koncepcjach, zwłaszcza zaś aparatu logiki indukcji, której zasadom bliskie są mechanizmy (procedury) selekcjonowania informacji w systemach dokumentacyjnych.

Do metod logiki indukcji nawiązywała scharakteryzowana wyżej próba sformułowania definicji relewancji przez D.J. Hillmana przy założeniu o konieczności ekstensjonalnej interpretacji znaczenia. Próba ta doprowadziła do wyników negatywnych. Należy przy tym zauważyć, że autor ten wykorzystał tylko w niewielkiej części propozycje metodologiczne R. Carnapa zawarte w jego teorii indukcji i prawdopodobieństwa logicznego /3/. Według Carnapa prawdopodobieństwo logiczne jest relacją między zdaniami, określającą status hipotez naukowych poprzez tzw. funkcję konfirmacji (confirmation function, c-function), za pomocą której ustala się stopień potwierdzenia pewnych zdań przez inne zdania. Jej logiczny charakter można wyjaśnić przez analogię z relacją implikacji w logice dedukcji, relacją zachodzącą między hipotezą (h) i świadectwem (przesłanką) (e), jeśli traktuje się implikację indukcyjną jako konwers implikacji dedukcyjnej, określanej za pomocą terminów "entailment", "implikacja konieczna". Carnap twierdził, że rozwiązania problemów implikacji w logice dedukcji i indukcji należy poszukiwać w sferze semantyki języka (tu semantyki logicznej). Prawdopodobieństwo zdania zależy bowiem od języka, w którym zostało sformułowane i wyraża się w stosunku do wszelkich możliwych zdań tego języka, odwzorowujących możliwe stany i własności obiektów, a więc ich cechy dystynktywne, które nazwał możliwymi światami (possible worlds), a które z kolei tworzą pewne klasy. Związek wynikania logicznego (implikacji) między zdaniami odwzorowuje właśnie funkcja konfirmacji, której wartości mogą

ulegać zmianie pod wpływem wprowadzenia do zbioru zdań nowej informacji (zdania) - (i). Stopień potwierdzenia hipotezy (h) przy korpusie dowodowym (e) uzupełnionym o (i) może się zmniejszyć, zwiększyć lub pozostać niezmienny, stąd nowa informacja (i) może być uznana za **relewantną pozytywnie, negatywnie lub irrelewantną**. W takim ujęciu relewancję możemy interpretować jako funkcję definiowaną na podstawie m- wartości określających stopień naailenia związku logiczno-semantycznego między wyrażeniami pewnego języka. Formułując zasady prawdopodobieństwa logicznego niezbędego do ustalenia funkcji konfirmacji zdań typu: .."eksperyment ponownie potwierdza teorię T", "eksperyment dostarcza nowych argumentów za..", "teoria jest potwierdzona w stopniu znacznie wyższym przez dane z eksperymentu znane obecnie niż przez dostępne przed dwudziestu laty", Carnap poświęcił wiele uwagi dyscyplinie metodologicznej swojej teorii eksponując rolę semantyki języka nauki, a więc reguł jego interpretacji warunkującej poprawność wszelkich procesów inferencyjnych. Pomiar relewancji, którym nie zajmujemy się w niniejszej pracy, zależy, w świetle wywodów Carnapa, od interpretacji semantycznej terminów wyjaśnianych w nauce, a zaczerpniętych z języka potocznego (tzw. pre-scientific terms), niekiedy wprowadzanych niezależnie do teorii naukowej. Wprowadzanie nowych pojęć winno odbywać się w dwu etapach:

- konstrukcji systemu aksjomatycznego,
- interpretacji tego systemu polegającej na ustaleniu znaczeń jego wyrażeń elementarnych (primitive axiomatic terms).

Interpretacja semantyczna wyrażeń języka nauki prowadzi do stworzenia tzw. definicji korelacyjnych lub korelacji epistemicznych pozwalających ustalić związki między tymi wyrażeniami oraz dokonywać ich transformacji. Budowa definicji korelacyjnych języka to inaczej proces eksplikacji, polegający na dostarczeniu definicji terminu

poprzez skorelowanie jego znaczenia ze znaczeniami innych terminów. (Termin, którego znaczenie jest wyjaśniane nazywany jest explicandum, zaś służący do wyjaśnienia - explicatum). Eksplikacja semantyczna terminów elementarnych języka powinna opierać się na następujących czterech zasadach:

- **podobieństwa** do explicatum do explicandum (similarity to the explicandum) - np. potoczne "ryba - zwierzę żyjące w wodzie" i zoologiczne "piscis",
- **dokładności** (exactness) - reguły użycia terminu powinny umożliwić wprowadzenie go do spójnego systemu terminów naukowych,
- **produktywności** (fruitfulness) - explicatum powinno być semantycznie produktywne. Pojęcie jest tym bardziej produktywne, w im więcej wchodzi związków z innymi pojęciami na podstawie obserwowanych faktów, innymi słowy, im większa jest jego użyteczność w formułowaniu praw,
- **proatoty** (simplicity) - explicatum powinno być proste na ile to możliwe, to znaczy w stopniu takim, na jaki pozwalają zasady 1 - 3.

Śród wymienionych zasad próbowano wykorzystać jedynie zasadę podobieństwa klas (por. Hillman) w celu sformułowania definicji relewancji. Wydaje się, że za najważniejszą należałoby uznać zasadę produktywności jako wyjaśniającą problem związków znaczeniowych między wyrażeniami języka informacyjno-wyszukiwawczego, decydujących o stopniu podobieństwa zbiorów informacji w SIW do instrukcji wyszukiwawczej, którą cytowani wyżej autorzy określali terminem "relatedness". Tak więc w świetle powyższych wywodów relewancję należałoby interpretować jako sieć relacji odwzorowanych w polu semantycznym języka informacyjno-wyszukiwawczego (relacji paradygmatycznych) oraz relacji syntagmatycznych w tekstach charakterystyk wyszukiwawczych dokumentów i instrukcji wyszukiwawczych. Relacje te odwzorowują związki między obiektami rzeczywistości dokumenta-

cyjnej i/lub pozadokumentacyjnej oraz cechy tych obiektów semantycznie dystynktywne. Ich ustalenie przy ekstensjonalnej interpretacji znaczenia jest niemożliwe, jak to udowodniono matematycznie. Sieć relacji semantycznych wyznaczających kryteria relewancji dla danego SIW powinna opierać się na intensjonalnej interpretacji znaczenia. Intensjonalna interpretacja znaczenia uzasadnia z kolei wielość relewancji odpowiadających złożoności procesowi komunikacji między użytkownikiem i systemem informacyjnym.

### Literatura

1. BRADFORD S.C. The documentation chaos. W: Documentation, London, Lockwood 1932 s.106-121.
2. BOOKSTEIN A. Relevance. Journal of American Society for Information Science 1979 vol. 29 s. 269-272.
3. CARNAP R. Logical foundations of probability. The University of Chicago Press 1962 613 s.
4. CLEVERDON, C. W. MILIS J.; KEEN E.M. Factors determining the performance of indexing system. Test results. Vol. 2, Aslib-Cranfield Research Project, Cranfield, England 1966.
5. CODPER W.S. A definition of relevance for information retrieval. Information Storage and Retrieval 1971 nr 7 s.19-37.
6. CUADRA C.A. KATTER R.V. Experimental studies of relevance judgments. Final report. Report TM-3520 Vol. 1 Project Summary, Vol.2 Description of Individual Studies, Systems Development Corp. Santa Monica June 1967.
7. DOYLE L.B. Is relevance an adequate criterion in retrieval system evaluation? Proceedings of the American Documentation Institute 1966 Part 2 Washington s.199-200.
8. FOSKETT D.J. A note on the concept of "relevance". Information Storage and Retrieval 1972 vol.8 nr 2 s.457-471.

9. GOFFMAN W. A general theory of communication. W: SARACEVIC T. Introduction to information science. New York Bowker 1970 s.726-747.
10. GOFFMAN W., NEVILL V.A. Communication and epidemic process. Proceedings of the Royal Society 1969 nr 1454 s.316-334.
11. HILLMAN D.J. The notion of relevance. American Documentation. 1964 vol. 15 nr 1 s.26-34.
12. KEMP D.A. Relevance, pertinence and information system development. Information Storage and Retrieval 1974 vol. 10 nr 2 s.37-47.
13. KEYNES J.M. AT treatise on probability. London 1921.
14. KOCHEN M. Principles of information retrieval. Los Angeles Mellville 1974.
15. KOPALIŃSKI W. Słownik wyrazów obcych i zwrotów obcojęzycznych. Warszawa 1978.
16. HOŁÓWKA T. Myślenie potoczne. Warszawa 1986.
17. KOZAKCHOV L.S. Relevancija v informatike i naukovedenii. Naučno-Tekhnickája Informacija Ser.2 1969 nr 8 s.3-11.
18. LOTKA A.J. The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences 1926 vol.16 nr 12 s.283-350 za: /11/.
19. MARON M.E., KUHN J.L. On relevance, probabilistic indexing and information retrieval. Journal of the Association for Computing Machinery 1960 vol.7 nr 3 s.216-244.
20. O'CONNOR J. Relevance disagreements and unclear request forms. American Documentation 1967 vol.18 nr 3 s.165-177.
21. PRICE D.J. de S. Networks of scientific papers. Science 1965 vol.149 nr 3683 s.510-515.
22. REES A.M., SARACEVIC T. The measurability of relevance. Proceedings of the American Documentation Institute Washington 1966 nr 3 s.225-234.
23. SALTON G. SMART. Automatyczny system wyszukiwania informacji. Warszawa 1975 585 s.

24. SAMBOR J. Słowa i liczby. Zagadnienia językoznawstwa statystycznego. Wrocław 1972.
25. SARACEVIC T. Relevance: a review of and a framework for the thinking on the notion in information Science. Journal of American Society for Information Science 1975 vol.26 nr 6 s.321-343.
26. SARACEVIC T., REES A.M. Conceptual analysis of questions in information retrieval systems. Proceedings of the American Documentation Institute Washington 1963 nr 2 s.175-177.
27. SATOSI WATANABE Theorem of ugly duckling. A mathematical explication of inductive inference. Proceedings of International Colloquium. Tihany, Hungary, za /11/.
28. SCHUTZ A. Reflections on the problem of relevance. New Haven Yale University Press 1970 186 s.
29. SWETS J.A. Information retrieval systems. Science 1963 vol.141 nr 3577 s.245-250.
30. TAUBE M.A. A Note on the pseudomathematics of relevance. American Documentation 1965 vol. 16 nr 2 s.69-72.
31. UNGURIAN O. Elementy teorii języków informacyjnych. Warszawa 1976 253 s.
32. URQUART D.J. Use of scientific periodicals. Proceedings of the International Conference on Scientific Information. Washington National Academy of Sciences 1959 s.277-290 za: /25/.
33. WILSON P. Situational Relevance - Information Storage and Retrieval. 1973 vol.9 nr 8 s.457-471.

18.04.1989 r.

## INTERPRETATIONS OF THE TERM "RELEVANCE" IN INFORMATION SCIENCE

### Summary

The article is devoted to the review and the systematization of different interpretations of the term "relevance" which have been formulated in information science. The origin and the early definitions of relevance are presented. The proposal to divide different definitions into two groups is put forward, namely: this in which relevance is interpreted as the relationship between elements of information system and this in which relevance is interpreted as the relationship between file of the system's environment. The interpretation of relevance in the light of the semantic explication of R. Carnap is discussed.

### ИНТЕРПРЕТАЦИИ ТЕРМИНА "РЕЛЕВАНТНОСТЬ" В НАУЧНОЙ ИНФОРМАЦИИ

#### Резюме

Статья посвящена обзору и систематизации различных интерпретации термина "релевантность" сформулированных в научной информации. Представлено происхождение первых определений релеванции. Предлагается разделить разнообразные определения на две группы, а именно: эта, в которой релевантность интерпретирована как отношение между элементами информационной системы и вторая, в которой релевантность считается как отношение между информационным массивом системы и её средой. Обсуждается интерпретация релеванности в свете семантического истолкования Р.Карнапа.

JADWIGA WOZNIAK

Instytut Bibliotekoznawstwa  
i Informacji Naukowej UW

### O TAK ZWANYCH POTRZEBACH INFORMACYJNYCH

Uwagi o określeniach treści terminu potrzeba informacyjna; krytyka ich teoretycznej poprawności i operacyjnej przydatności. Charakterystyka najczęściej stosowanych schematów badań użytkowników informacji. Proponycje ich modyfikacji.

Rozczarowanie co do praktycznych walorów dużej części badań użytkowników informacji oraz trudności w porozumiewaniu się różnych specjalistów z dziedziny informacji wydają się mieć jedną z przyczyn w nieładzie terminologicznym, jaki panuje w niektórych obszarach tej dyscypliny. Jednym z takich pól jest problematyka potrzeb informacyjnych widziana w kontekście sytuacji informacyjnej użytkowników. Tutaj krytycznymi terminami wydają się być m.in. następujące: potrzeba informacyjna (=potrzeba użytkownika), pytanie informacyjne (=zapytanie informacyjne, =kwerenda), pertynencja i relewancja. Aby zapobiec nieporozumieniom, często sprowadzającym dyskusję naukową z poziomu merytorycznego na poziom spierania się o formę językową, potrzeba prac nad systemem leksykalnym języka wykorzystywanego tu do przekazywania informacji. Dlatego w artykule chcemy poddać analizie pierwszy z wymienionych terminów, to jest

---

"Zagadnienia Informacji Naukowej" 1989 nr 1(54)



termin potrzeba informacyjna. Zagadnieniom pytań, pertynencji i re-  
lewancji będą poświęcone odrębne teksty.

Nie wdając się w rozważanie trudnego i kontrowersyjnego pro-  
blemu uniwersalnej definicji i typologii potrzeb, chcę zwrócić  
uwagę na pewne, istotne dla analizy terminu potrzeby informacyjne,  
funkcjonujące w psychologii uogólnienia i obserwacje, zwłaszcza  
dotyczące tzw. potrzeb poznawczych. (Piszę "tzw. potrzeb poznaw-  
czych", żeby zwrócić uwagę, iż użyta nazwa jest tylko jedną ze  
stosowanych i została wybrana arbitralnie spośród innych równo-  
uprawnionych). Przede wszystkim spróbujmy określić, czym jest (jak  
będziemy pojmować termin) potrzeba.

"Jeżeli mówimy, że X ma potrzebę p, znaczy to tyle co, jeżeli  
p jest niedostępne dla X, to następuje zachwianie homeostazy (równo-  
wagi układu), która jest podstawą prawidłowego funkcjonowania X"<sup>1/</sup>.  
Inaczej mówiąc: podmiot działania ma potrzebę p, to znaczy, że  
dla normalnego jego funkcjonowania niezbędne są określone warunki,  
stwarzane przez różnorodne bodźce. Niezaspokojenie potrzeby nie  
oznacza odzwierciedlenia w świadomości jakiegoś faktycznego braku,  
czyli nie "pragnienie", "pożądanie" itp., ale obiektywną sytuację,  
w której organizm jest pozbawiony czegoś istotnego do funkcjonowa-  
nia. Sytuacja ta może (ale nie musi) znaleźć odbicie w jego świadomości.  
Każdej potrzebie daje się więc przyporządkować jakościowo  
i ilościowo pewne bodźce, które ją zaspokajają, jak również wymie-  
nić bodźce, które ją wywołują lub wzmagają. Istnieją dość znaczne  
różnice zdań co do typów i hierarchii potrzeb. Niemniej panuje  
powszechna zgoda, że potrzeby wyższego rzędu wywodzą się z potrzeb  
podstawowych oraz że potrzeby ludzkie nie są czymś stałym, lecz

---

<sup>1/</sup>K. Obuchowski: Psychologia dążeń ludzkich. Warszawa 1972 s. 105

rozwijają się i zmieniają. Wśród potrzeb wyższego rzędu lokują się potrzeby, które nazwaliśmy poznawczymi, choć w pracach różnych autorów są one różnie nazywane, określane i traktowane.

Na przykład Murray wymienia potrzeby wiedzy i potrzeby wyjaśniania oraz aktywność poznawczą w kontekście potrzeb czynnościowych (obejmujących dyspozycje do zajmowania się czynnością ze względu na nią samą), potrzeb psychicznych i twórczych. W przekształconej wersji teoria Murraya wyróżnia 12 wektorów (tendencje do działania) oraz 14 wartości. Wśród wartości wymieniona jest wiedza i w odniesieniu do niej wyróżnić można potrzeby zdobywania, budowania i wyrażania wiedzy. Teoria ta została oparta na obszernym materiale empirycznym, natomiast jej mankamentem jest to, że opisuje zbiór potrzeb jako zestawienie nie powiązanych wzajemnie dyspozycji motywacyjnych, bez wyjaśniania ich genezy i funkcji. Zarzuty takie nie mogą być postawione najpopularniejszej chyba w psychologii teorii potrzeb Masłowa, który przyznaje potrzebom poznawczym wyjątkową rolę, uznając że tworzą one niemal równoległą i komplementarną hierarchię w stosunku do podstawowych potrzeb. Zauważmy, że w obu wymienionych wyżej ujęciach termin "potrzeba" używany był w znaczeniu zastępującym takie terminy, jak popęd, motyw lub dążenie. W psychologii podejmuje się próby, aby terminowi potrzeba nadać (lub przywrócić) znaczenie obiektywne. Jedną z bardziej udanych jest koncepcja Obuchowskiego, w której potrzeba poznawcza przedstawiona jest jako jedna z siedmiu innych. Jej definicja brzmi: "Potrzeba poznawcza jest to właściwość człowieka, która powoduje, że bez otrzymania określonej liczby informacji w każdej sytuacji i bez możliwości przeprowadzenia czynności poznawczych za pomocą pojęć w sytuacjach częściowo nowych człowiek nie może normalnie funkcjonować"<sup>2/</sup>. Problematyki potrzeb poznawczych doty-

<sup>2/</sup> K. Obuchowski: op. cit. s. 166.

czy także koncepcja systemowa Kocowskiego<sup>3/</sup>. Potrzeby są w niej rozumiane jako warunki obiektywne; "potrzeba to tyle co pewien warunek określonego celu człowieka dający się określić obiektywnie"<sup>4/</sup>. Wśród 42 kategorii potrzeb, zgrupowanych w 5 zasadniczych klasach znaleźć można co najmniej 8 ściśle dotyczących aktywności poznawczej. W kategorii potrzeb prokreacji i rozwoju wymienione są potrzeby dydaktyczne, dotyczące zachowań ukierunkowanych na przekazywanie potomstwu informacji oraz umiejętności niezbędnych do samodzielnego życia, oraz potrzeby samokształcenia odnoszące się do nabywania wiedzy przez dorosłych. Również potrzeby samorealizacji, czyli "zbiór koniecznych lub pożądaných zgodności pomiędzy uzdolnieniami, zamiłowaniem, wykształceniem i pełnionymi rolami społecznymi", zdają się zawierać warunek ukierunkowanej i długofalowej aktywności poznawczej. W kategorii potrzeb funkcjonalnych, decydujących o skuteczności działania, wymienia Kocowski potrzeby dotyczące "uzyskiwania, przechowywania i wykorzystywania informacji niezbędnych do realizacji każdego celu, każdej potrzeby". W tej samej grupie umieszczone są także potrzeby sprawnościowe, czyli prakseologiczne, obejmujące zbiór umiejętności, technik i metod działania, dotyczące ogólnych zasad i szczegółowych procedur, algorytmów i heurystyk. Wymienione rodzaje potrzeb, jak wszelkie inne, mają swe odwzorowania subiektywne i z uwagi na to można je wszystkie odnaleźć powtórnie w kategorii potrzeb psychicznych. Kocowski nie podaje jednak ich uzupełniających charakterystyk. Realizacja każdej niemal czynności wymaga w jakimś sensie aktywności poznawczej, a zatem jest uwarunkowana potrzebą poznawczą, chociaż oczywiście nie zawsze

---

<sup>3/</sup>T. Kocowski: Potrzeby człowieka. Koncepcja systemowa. Wrocław 1982.

<sup>4/</sup>op. cit. s. 69.

stanowi ona podstawowy warunek spełnienia celu tej czynności. Systemowy charakter propozycji Kocowskiego uwzględni taki właśnie uniwersalny walor potrzeb poznawczych, chociaż wydaje się, że nie docenia się tu ich autonomicznej roli. Rola ta nie da się sprowadzić do "potrzeb dydaktycznych" i "samokształcenia", a w grupie potrzeb samorealizacji omówiona jest zdawkowo.

Analiza poglądów psychologów na temat potrzeb poznawczych zdaje się prowadzić do dwóch istotnych wniosków:

- właściwie wszyscy autorzy dostrzegają szczególną wagę i znaczenie potrzeb związanych z aktywnością poznawczą;

- zauważalna jest podwójna rola tych potrzeb; z jednej strony "obsługują" one niejako wszystkie inne potrzeby z drugiej zaś pełnią funkcje autonomiczne.

Drugi z wniosków można wyrazić inaczej, stwierdzając, iż **zdobywanie informacji jest warunkiem (bezpośrednim lub pośrednim) zaspokojenia wielu potrzeb, nie tylko poznawczych.**

Weźmy teraz pod uwagę rozpowszechnioną definicję potrzeb informacyjnych, zamieszczoną w "Słowniku terminologicznym informacji naukowej", która mówi, że są to "ujawnione lub nieujawnione (potencjalne) potrzeby użytkowników informacji"<sup>5/</sup>, czyli "osoby lub zespołu osób korzystających z informacji"<sup>6/</sup>. W języku polskim wyraz korzystać jest rozumiany jako "odnosić, osiągać korzyści, mieć pożytek, zysk z czegoś, użytkować, wyzyskiwać coś, używać czegoś"<sup>7/</sup>. Czyli potrzeby informacyjne są to potrzeby osób, które mają pożytek

---

<sup>5/</sup> Słownik terminologiczny informacji naukowej. Warszawa 1979 s.91.

<sup>6/</sup> op. cit. s. 119.

<sup>7/</sup> Słownik języka polskiego. T. 1. A-K. Warszawa 1988 s. 1020.

z informacji, używają jej, wykorzystują. Zdobyć informacji nie jest więc tu celem samym w sobie, nie może też kończyć aktywności skierowanej na zaspokojenie potrzeby, lecz jest niezbędnym warunkiem osiągnięcia tego celu. Wobec tego możemy stwierdzić jedynie tyle, że w definicji chodzi o takie potrzeby, których zaspokojenie jest związane ze zdobyciem informacji. Z wniosków dotyczących poglądów psychologów na temat potrzeb ludzkich wynika zaś, że cecha ta może przysługiwać najróżniejszym (niemal wszystkim) klasom potrzeb. Przytoczona definicja (pomijamy tu, jako w takiej sytuacji nieistotną, kwestię rozumienia sformułowania "nieujawnione (potencjalne) potrzeby") nie przyporządkowuje więc w sposób ostry definiowanemu terminowi zakresu obiektów rzeczywistości pozajęzykowej. W tym momencie warto się zastanowić nad sensownością działań zmierzających do zbudowania poprawnej definicji terminu potrzeby informacyjne. Na obecnym etapie wiedzy bibliotekoznawczej chyba nie potrafimy wskazać istotnych cech konotacyjnych, które pozwoliłyby na wyodrębnienie zakresu "potrzeb informacyjnych" (jeśli taka kategoria potrzeb w ogóle autonomicznie istnieje, bo tylko wtedy uzasadnione byłoby wzbogacenie języka o ten termin) spośród denotacji nazwy potrzeby ludzkie. Czy nie lepiej byłoby potraktować tę nazwę jako skrótowo-zastępcze wyrażenie sensu frazy "informacja dla zaspokojenia potrzeby  $p_i$ " i następnie skoncentrować uwagę na eksplorowaniu z jednej strony różnych cech przysługujących poszukiwanej informacji, z drugiej zaś zachowań osób, które szukają informacji - środka niezbędnego do osiągnięcia pewnego celu.

W dalszej części artykułu zostanie podjęta próba wykazania, że tzw. badania potrzeb użytkowników w istocie nie były i nie są badaniami potrzeb, ale właśnie zachowań czy wręcz samych systemów informacyjno-wyszukiwawczych. W tym celu najpierw krótko zilustrujemy zależności, jakie zachodzą między dowolnymi potrzebami a wywo-

zającymi się z nich zachowaniami, dzięki czemu zyskamy pewne narzędzia terminologiczne, którymi następnie będziemy się posługiwać przy charakteryzowaniu schematów badań użytkowników informacji.

#### POTRZEBA - MOTYW - ZACHOWANIE

Różnego rodzaju potrzeby i chęć ich zaspokojenia powodują, że ludzie podejmują (większy lub mniejszy) trud zdobywania informacji. Nie znaczy to jednak, że zaistnienie potrzeby zawsze inicjuje działanie poszukiwania informacji.

Pierwszym warunkiem, koniecznym choć niewystarczającym, jest uświadomienie faktu istnienia potrzeby. Następnym - sformułowanie odpowiednio silnego, trwałego i zyskującego aprobatę podmiotu motywu działania. Motyw spełnia funkcję sterującą i kontrolującą, musi więc być zrozumiały dla danej osoby i przekonujący. Motyw, będąc werbalizacją celu działania i sposobów jego osiągnięcia, jest bezpośrednim czynnikiem uaktywniającym zachowania. Wśród czynników kształtujących zachowania użytkowników informacji najważniejszymi wydają się być: cechy osobnicze użytkownika, doświadczenie społeczne, wymagania, jakie stawia mu otoczenie, role, jakie pełni w grupach społecznych, wreszcie postawy wobec przedmiotu działania, w tym waga, jaką użytkownik nadaje zaspokojeniu potrzeby oraz przewidywane konsekwencje działania w warunkach niepełnej informacji itp. Czynniki te wyznaczają ważność ewentualnego celu i sposoby jego osiągnięcia. Rezultaty działania w dużym stopniu zależą od wiedzy użytkownika o przedmiocie potrzeby i od możliwości jego uzyskania. Pierwsza część ostatniego stwierdzenia jest sprzeczna z rozpowszechnionym w nauce o informacji i w bibliotekoznawstwie aksjomatem o niechęci użytkowników do ujawniania potrzeb informacyjnych, do wypowiedzania się na temat, jakich informacji i w jakiej

postaci potrzebują. (Zakłada się bowiem, że użytkownicy byłiby w stanie dokładnie określić swoje potrzeby, gdyby tylko chcieli współpracować z systemami). Tymczasem jeśli zgodzić się, że zachowania informacyjne są inicjowane przez użytkowników w sytuacjach wymagających wsparcia informacją, aby móc osiągnąć wyznaczone cele, to można się spodziewać, że użytkownik będzie mógł, w rozsądnych granicach, scharakteryzować te cele. Nierozsądnym byłoby jednak oczekiwać, że będzie on w stanie określić, jak zasób jego informacji powinien się zmienić, o jakie informacje wzbogacić, aby stać się pewnym, czyli wystarczającym do osiągnięcia wyznaczonych celów. Gdy człowiek styka się z czymś, czego nie może w żadnym aspekcie przyporządkować czemuś, co zna, nie jest też w stanie mówić o tym. Istotna jest więc umiejętność odróżniania sytuacji niepewności od sytuacji luki informacyjnej, która jest niepewnością na tyle zdefiniowaną, że można oczekiwać aktywności użytkownika ukierunkowanej na jej przedstawienie innej osobie, na przykład bibliotekarzowi pośredniczącemu między nim a zbiorami informacyjnymi systemu. Aktywność użytkownika, który jednak podjął trud szukania informacji, może się realizować poprzez następujące typy zachowań:

- zachowania użytkownika, który nie nawiązuje kontaktu z systemem informacyjno-wyszukiwawczym (SIW). Przyczyną może tu być albo ignorancja użytkownika, albo przewidywana przez niego wyższa sprawność i skuteczność działania nieformalnych kanałów zdobywania informacji (np. krótszy czas czekania na odpowiedź, która z kolei może zawierać informacje aktualniejsze od przechowywanych w SIW)<sup>8/</sup>;

---

<sup>8/</sup> Zagadnienia nieformalnych kanałów komunikacji było przedmiotem między innymi rozprawy doktorskiej Danuty Koniecznej ("Nieformalne procesy komunikacji naukowej. Łódź 1980, 241 ss. Uniwersytet Łódzki Wydział Filologiczny) oraz publikacji tej samej autorki pt. "Rola nieformalnych procesów w systemie komunikacji naukowej". Warszawa 1982. Prace IINTE nr 44.

- zachowania użytkownika, który nawiązuje kontakt z SIW i sam stara się odnaleźć potrzebną informację;
- zachowania użytkownika, który nawiązuje kontakt z osobą (np. bibliotekarzem), która będzie pośredniczyć między nim i systemem informacyjno-wyszukiawczym.

Teraz możemy już zająć się scharakteryzowaniem funkcjonujących modeli badań użytkowników informacji. Szczególną uwagę zwrócimy na określenie bezpośredniego przedmiotu badania, chcemy bowiem wykazać, iż nie są nim potrzeby lecz zachowania informacyjne oraz że badania te nie upoważniają do wyrokowania o **potrzebach** użytkowników.

#### SCHEMATY BADAŃ UŻYTKOWNIKÓW INFORMACJI

Dla najstarszego i najbardziej prymitywnego z funkcjonujących schematów badań użytkowników charakterystyczne jest założenie istnienia potrzeb informacyjnych jako podkategorii potrzeb ludzkich i relewantności wskazanych przez system dokumentów. Przy takim założeniu każdy system informacyjno-wyszukiawczy jest systemem efektywnym, w pełni zaspokajającym potrzeby. Badania potrzeb informacyjnych użytkowników sprowadzają się tu do dość szczegółowych opisów korzystania z systemów (głównie bibliotecznych) i z różnych typów dokumentów. Nie bierze się pod uwagę ani zagadnień motywacji dynamizujących procesy szukania informacji, czyli tego, co w istocie jest wyznacznikiem potrzeb użytkowników, ani problemów jej wykorzystania. Charakterystyczne jest, że nawet w przypadkach badań przeprowadzanych w organizacjach przemysłowych, gdzie systemy informacyjne powstawały z myślą o osiągnięciu takich celów jak promocja innowacji czy redukcja zbytecznych działań, zapotrzebowanie na informację rzadko było kojarzone z czynnikami zewnętrznymi w stosunku do systemów informacyjno-wyszukiawczych, takimi jak motywy czy



wykorzystanie informacji. Pewne zmiany schematu dały się zaobserwować w połowie lat sześćdziesiątych. Polegały one najpierw na zaakceptowaniu faktu, że użytkownicy danego systemu nie stanowią nieodróżnianej grupy osób jednakowo zachowujących się w trakcie szukania informacji. Następnie wyraziły się w postrzeganiu użytkownika jako elementu różnych systemów społecznych. Nadal jednak za miary zaspokojenia potrzeb przyjmowano oceny np.: wiedzy użytkowników o systemach i usługach, z jakich mogliby korzystać, wielkości podzbioru populacji użytkowników pozytywnie oceniających dany system lub jego elementy, zakresu, w jakim użytkownicy korzystają z różnych typów usług informacyjnych, źródeł informacji czy zbiorów dokumentów. Badania te nie były w istocie badaniami potrzeb użytkowników, ale badaniami systemów informacyjno-wyszukiwawczych prowadzonymi nie wprost lecz za pośrednictwem użytkowników. Orientacja "na system" nadal się cieszy dużą popularnością, chociaż nie uwzględnia się w niej między innymi kwestii, że usługi informacyjne, nawet o szerokim zakresie i niekiedy bardzo wyrafinowane, nie muszą być tymi, których oczekuje użytkownik. Jednym z osiągnięć kierunku postrzegającego użytkownika w przestrzeni systemów społecznych, których jest on elementem (kierunku, który można by nazwać socjologicznym) było wprowadzenie terminu zachowanie informacyjne<sup>9/</sup>. Oznacza on celowe, czyli prowadzące do zamierzonego celu, działanie użytkownika, który znalazł się w takiej sytuacji, że posiadane przez niego informacje nie są wystarczające do osiągnięcia tego celu.

---

<sup>9/</sup> Termin ten został wprowadzony do literatury przez D.W. Kinga i V. Palmoura, którzy wzbogacili komunikacyjny model systemu informacyjno-wyszukiwawczego, stworzony przez Lina i Garvey'a (Lin N., Garvey W.D. Information needs and uses. Annual Review of Information Science and Technology 1972 Vol. 7. s.5-37), o nowy element, który nazwali "user behaviour".

Inaczej mówiąc, zachowanie informacyjne to zachowanie użytkownika, który znalazł się w sytuacji wymagającej zasilenia informacją, aby móc się zbliżyć do stanu dla niego korzystnego czy pożądanego, czyli do celu. Powstrzymanie się od działania jest również działaniem i wyraża się określonym zachowaniem. Zapotrzebowanie na informację i jej wykorzystanie są więc elementami wzajemnie się warunkującymi, wyrażającymi się w postaci różnych zachowań informacyjnych. Pierwszym ważniejszym znakiem nowej orientacji w badaniach użytkowników były publikacje Paisleya<sup>10/</sup> i Allena. Kluczowym założeniem było przekonanie, że zachowanie człowieka pozostaje pod wpływem pewnej liczby systemów społecznych. Inaczej mówiąc, człowiek jest elementem różnych systemów i jego zachowanie w dowolnym z tych systemów nie może być odizolowane od wpływu innych systemów. Także i zachowania informacyjne są warunkowane nie tylko przez cechy systemu informacyjno-wyszukiwawczego, z którym użytkownik się kontaktuje, ale także przez systemy społeczne (prawny, ekonomiczny, polityczny, kulturowy itd.), do których należy, i oczywiście, przez jego cechy indywidualne. Aby zrozumieć zachowanie użytkownika informacji trzeba je więc przeanalizować w kontekście wszystkich tych systemów. Socjologicznie zorientowane badania użytkowników wykazały, że:

- istnieją poza systemami informacyjno-wyszukiwawczymi i poza samym zapotrzebowaniem na informację czynniki, które mają niekwestionowany wpływ na zachowania informacyjne użytkowników;
- czynniki te zależą m.in. od miejsca użytkownika w kompleksie systemów społecznych, kategorii problemów, z jakimi się styka i celów, dla osiągnięcia których potrzebuje informacji.

---

<sup>10/</sup> W.J. Paisley: Information needs and uses. Annual Review of Information Science and Technology 1968 Vol. 3. s. 1-30.

Mankamentem socjologicznej koncepcji badań było zbyt mocne przekonanie, że dla sporządzenia charakterystyki istotnych cech użytkownika determinujących jego zachowanie całkowicie wystarczą dane o jego wieku, płci, kulturze, w jakiej został wychowany, pochodzeniu społecznym, profilu, poziomie wykształcenia itd. Pogląd ten jest bowiem na tyle słuszny, na ile słuszne jest zdanie, że przy takich samych potrzebach takie same bodźce muszą doprowadzić do utrwalenia się takich samych zachowań. Dziś jest już widoczne, że badania użytkowników w tych wczesnych okresach były w większości nieskuteczne, jeśli chodzi o formułowanie przesłanek do wykorzystania w projektowaniu jakichkolwiek systemów informacyjnych. W literaturze zwykle wskazuje się na dwie tego przyczyny - metodologiczną niewiarygodność badań, i co za tym idzie ich wyników, oraz tradycyjne pojmowanie potrzeb informacyjnych jako niezróżnicowanej podkategorii potrzeb powszechnych. Do lat siedemdziesiątych prawie w ogóle nie było badań, których przedmiotem byłyby zagadnienia motywacji dynamizujących zachowania informacyjne i niewielka liczba badań, w których usiłowano powiązać zapotrzebowanie na informację z jej wykorzystaniem. Nie było więc możliwości skonstruowania żadnej ramy metodologicznej, w obrębie której można by interpretować różne zachowania informacyjne ujawniające się w odmiennych warunkach, wyjaśniać je w kategoriach ogólniejszych niż odnoszące się do SIW, z którym użytkownik się kontaktował.

W tym samym czasie co podejście socjologiczne pojawił się w badaniach użytkowników inny trend, wyrażający się w uwzględnianiu czynnika indywidualizacji zwanego kognitywnym. Za prekursora tego podejścia można uznać Taylora.

Koncepcja Taylora była próbą określenia czterech poziomów potrzeb i odpowiadających im pytań. Najniższy poziom potrzeb nazwał Taylor potrzebą wewnętrzną (visceral need). Temu poziomowi nie od-

powiada żadna konstrukcja o funkcji pytania, ponieważ potrzeba ta nie jest wyrażana w żadnym języku. Następny poziom tworzy potrzeba uświadomiona (conscious need). Odpowiada jej niewypowiedziane, ale określające obszar niepewności, pytanie, które użytkownik sam sobie zadaje i na które sam usiłuje odpowiedzieć. Trzeci poziom (pojawiający się w wypadku niepowodzenia na poziomie drugim) to potrzeba zwerbalizowana (formalized need), wyrażająca się w postaci pytania czy pytań kierowanych do innych osób w nadziei, że pomogą oni albo lepiej zrozumieć problem albo go rozwiązać. Jeśli jednak rozmowy te nie wyeliminują niepewności, nie rozwiążą problemu, użytkownik może się zwrócić do systemu informacyjno-wyszukiwawczego z zapytaniem informacyjnym (odpowiada mu czwarty poziom potrzeb - potrzeba wynegocjowana (compromised need). Może tym się ona różni od pytania z poziomu trzeciego, że opis problemu zwykle jest tu precyzyjniejszy, zaś samo pytanie może być sformułowane z uwzględnieniem znanych użytkownikowi ograniczeń SIW. Reprezentanci orientacji kognitywnej nawiązują do sytuacji, w której użytkownik styka się z problemem i rozpatrują ją z punktu widzenia związku z wewnętrznymi strukturami wiedzy, postaw, osobowości, przekonań, otoczenia, sytuacji, problemu itd. Sytuacja problemowa w tym świetle to taka sytuacja, w której podmiot działania uznaje te wewnętrzne struktury za niewystarczające do rozwiązania problemu, i co za tym idzie, do osiągnięcia celu. Obraz sytuacji stworzony przez użytkownika, który stanął przed problemem do rozwiązania jest taki, że wymagane są zasilenia informacją pochodzącą ze źródeł zewnętrznych, aby osiągnąć poziom pewności niezbędny do podjęcia skutecznego działania. Sytuacja problemowa, jak każda sytuacja, obejmuje zarówno czynniki mniej lub więcej subiektywne i reakcje zachowań mogą być rozpatrywane tylko w związkach z całym kontekstem, tj. sytuacją taką, jaka istnieje sprawdzalnie oraz jaką wydawała się istnieć

zainteresowanej osobie. Tak więc można przyjąć, że sytuacja problemowa, mająca wiele wspólnego z "visceral need" Taylora, jest przyczyną zachowań informacyjnych, w których wiedza i cele użytkownika przenikają się z efektami i wykorzystaniem informacji. Głównym osiągnięciem orientacji kognitywnej wydaje się być szukanie źródeł zachowań informacyjnych w innych obszarach niż tzw. potrzeby informacyjne, wyrażające się między innymi w analizowaniu problemów, z jakimi stykają się użytkownicy informacji. Problemy te mają wpływ na ich stany psychiczne, tak samo jak wewnętrzne stany użytkowników mogą być interpretowane jako problemy same w sobie.

Jednym z bardziej znaczących przedstawicieli orientacji kognitywnej w badaniach użytkowników są Wersig<sup>11/</sup> i Belkin<sup>12/</sup>. Orientacja ta wyraźnie zdaje się tworzyć alternatywną postawę badawczą wobec tego, co do tej pory było robione w obszarze badań użytkowników.

Podsumujmy to, co do tej pory zostało napisane.

1. Zakładamy, że nie istnieje taka podkategoria potrzeb jak potrzeby informacyjne, ale różne potrzeby mogą być zaspokajane (bezpośrednio lub pośrednio) przez informację.

---

<sup>11/</sup>G. Wersig: The problematic situation as a basic concept of information science in the framework of social sciences: a reply to N.J.Belkin: Theoretical problems of informatics: new trends in informatics and its terminology. Moscow 1979. FID 56B,

<sup>12/</sup>N.J. Belkin: Cognitive models and information transfer. Social Science Information Studies 1984 nr 4 s.111-129; N.J. Belkin, R.N. Oddy, H.M. Brooks: ASK for information retrieval: Part I. Background and theory. Journal of Documentation 1982 nr 2 s.61-71; N.J. Belkin, R.N. Oddy, H.M. Brooks: ASK for information retrieval: Part II. Results of a design study. Journal of Documentation 1982 nr 3 s. 145-164; N.J. Belkin, T. Seeger, G. Wersig: Distributed expert problem treatment as a model for information system analysis and design. Journal of Information Science 1983 vol.5. nr 4 s.153-168.

2. Potrzeby mogą dynamizować zachowania, ale reakcje na bodźce otoczenia oraz na bodźce wytwarzane przez samego człowieka mają miejsce tylko wtedy, gdy bodźce te działają na tle odpowiednich potrzeb.

3. Reakcją człowieka (mającego określone potrzeby) na bodźce jest stan zakłóconej homeostazy (sytuacja problemowa), której odzwierciedleniem w świadomości jest motywacja do działania.

4. Sformułowanie motywu, spełniającego określone warunki, umożliwia rozpoczęcie działania, poprzez wyznaczenie celu (kierunku działania) i programu działania. Przyczyną działania może być kilka motywów, nawet sprzecznych ze sobą, co pozwala wyjaśnić pewne specyficzne sposoby działania, jego niekonsekwencje lub powstrzymanie się od działania. Natomiast wybór motywu jest uwarunkowany przez osobnicze i społeczne doświadczenie jednostki.

5. Działania, o których mowa w punkcie 4., objawiają się w postaci różnych zachowań.

Dążąc do zaproponowania kompleksowego modelu użytkowników informacji za punkt wyjścia przyjmujemy sytuację problemową oraz niektóre kształtujące ją czynniki, wynikające z widzenia użytkownika w otoczeniu oddziałujących na niego systemów społecznych. Uwagę skoncentrujemy tylko na tych sytuacjach i tylko na tych zachowaniach, których celem jest pozyskanie informacji potrzebnej do rozwiązania problemu poprzez komunikowanie się użytkownika z systemem informacyjno-wyszukiwawczym. Poza zakresem analizy pozostaną więc zagadnienia związane z powstrzymywaniem się od działania<sup>13/</sup> oraz problemy

---

<sup>13/</sup> Zagadnienie to coraz częściej jest podnoszone w literaturze jako istotny element, który powinien stać się przedmiotem badań, jeśli projektowany system czy język informacyjno-wyszukiwawczy ma być optymalny, tzn. zdolny do zaspokojenia potrzeb wszystkich potencjalnych użytkowników. W literaturze anglojęzycznej ten typ badań nosi nazwę nonuser studies.

zachowań informacyjnych skierowanych na nieformalne kanały komunikowania się.

Tylko niewielka ilość informacji pochodzi z bezpośredniego doświadczenia. Pozostała uzyskiwana jest za pośrednictwem innych ludzi. W ten sposób stają się oni nie tylko elementami, ale i konstruktorami obrazu świata innych ludzi. Tak więc czynnikiem kształtującym informację pozyskiwaną przez użytkownika jest jego otoczenie. Strukturę otoczenia wyznaczają systemy i grupy społeczne, których elementem jest dany użytkownik. Najczęściej w literaturze wskazuje się na cztery systemy społeczne, jako na czynniki kształtujące wzorce zachowań informacyjnych. Są to: system kulturowy, polityczny, ekonomiczny i prawny. W określonych przez nie warunkach użytkownicy informacji wypełniają różne role społeczne, których wpływ na kształt zachowań informacyjnych wydaje się być bardziej znaczący (w sensie bardziej różnicujący zachowania poszczególnych osób) niż tych wielkich systemów społecznych.

Analiza ról społecznych w świetle literatury przedmiotu jest jednym z najwartościowszych źródeł informacji o wzorcach zachowań informacyjnych. Spośród ról pełnionych przez użytkowników informacji jedną z ważniejszych wydaje się być rola zawodowa<sup>14/</sup> (szeroko określana, tak że w jej zakresie mieszczą się i role stricte zawodowe, i role wynikające z przynależności do innych grup stawiających członkom pewne formalne zadania, np. rola bycia studentem,

---

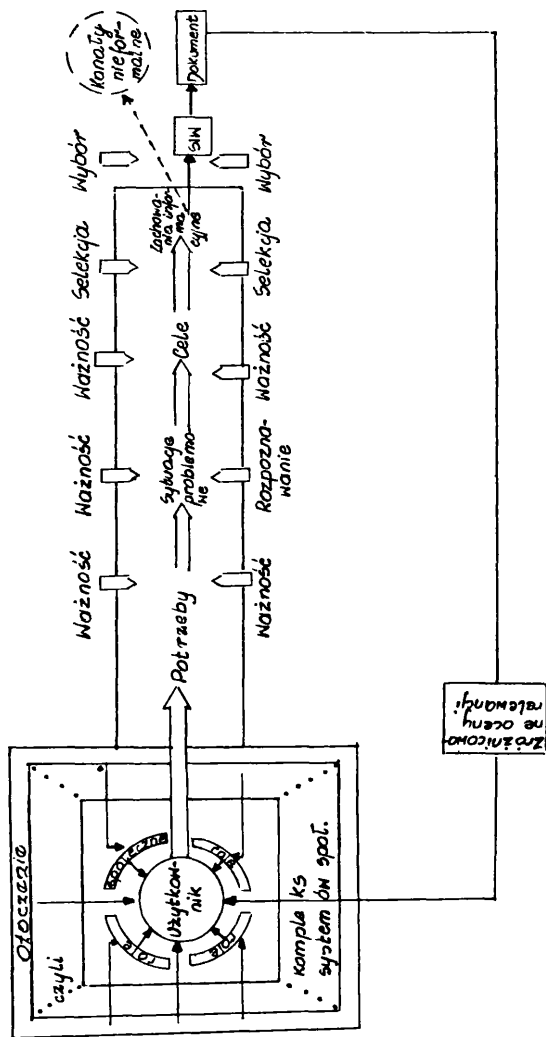
<sup>14/</sup> Wskazanie na rolę zawodową jako jedną z ważniejszych ról społecznych wpływających na projektowanie systemu informacyjno-wynikawczego jest konsekwencją ograniczenia rozważań do dokumentacyjnych SIW. Jeśli określono by SIW na przykład tak, że w ich zakresie znalazłyby się faktograficzne systemy, np. "udzielające" porad prawnych, to na równi z rolą zawodową należałoby analizować inne role, jak np.: rolę bycia obywatelem, rolę bycia rodzicem lub opiekunem, rolę bycia właścicielem itd.

rola bycia uczniem itp.), realizowana w ramach określonego szczebla hierarchii pewnej formalnej struktury organizacyjnej, charakteryzującej się właściwymi jej cechami, klimatem, ukierunkowanej na osiągnięcie określonych celów. Analiza roli zawodowej, skojarzona z cechami osobniczymi użytkownika, determinuje m.in. pojawianie się pewnych potrzeb (nie tylko poznawczych, również innych, np. emocjonalnych, takich jak potrzeba samorealizacji czy uznania), celów, problemów i co za tym idzie schematów zachowań informacyjnych. Tak więc warunkiem koniecznym, choć niewystarczającym, powodzenia badań jest rzetelna analiza miejsca użytkownika w wybranym ze względu na cele badania podsystemie grup społecznych i wynikających z nich ról społecznych. Niezwykle istotne jest, żeby badania realizowane według tego nowego schematu nie atępiały sobie celów uniwersalnych i nie stawały się pretekstem do pochopnych uogólnień. Wyznaczony obszar badawczy muszą konstruować względnie jednorodne, porównywalne elementy (tak sami użytkownicy, jak i ich cele, problemy, warunki działania itp.), w liczbie umożliwiającej metodycznie poprawną analizę. Do tej pory<sup>15/</sup> nie został wymieniony jeszcze jeden system, który w największym stopniu ma wpływ na kształt zachowań informacyjnych. Jest nim sam użytkownik - jego cechy indywidualne. Struktury osobowościowe, procesy i kategorie poznawcze, wartości, procesy pamięci, wychowanie, wiek, płeć itd. wre-

---

15/ Wyróżnione poziomy oddziaływania na zachowania informacyjne - społeczny i grupowy - nie reprezentują jednolitej płaszczyzny teoretycznej ani nie ukazują "przejścia" prawidłowości jednego poziomu w prawidłowość funkcjonowania drugiego poziomu. Procesy grupowe są immanentną właściwością grup, podczas gdy procesy zachodzące na poziomie społecznym są efektem świadomej i zamierzonej działalności zorganizowanej społeczności. Z tego powodu stanowią dwa różne (co nie znaczy, że niezależne) stany otaczające użytkownika rzeczywistości.





Schemat zależności elementów w modelu badania użytkowników informacji

sztanie język są najistotniejszymi indywidualnymi modulatorami zachowań, także informacyjnych. Jak pokazują między innymi porównawcze badania kultur, rodzaje działań podejmowanych w toku rozwoju danej społeczności w znacznej mierze kształtują sposób widzenia rzeczywistości, co znajduje bardzo silne odzwierciedlenie w języku tej społeczności. Poznając język jednostka uzyskuje dostęp do kulturowego zbioru kategorii poznawczych, których może używać do kodowania i dekodowania informacji, ale które jednocześnie w pewnym stopniu narzucają jej sposób widzenia rzeczywistości. Język pełni więc podwójną rolę - jednego z czynników kształtujących informację, i przez to wpływających na kształt zachowań informacyjnych, oraz jedynego środka, za pomocą którego możliwy jest wgląd w poszczególne elementy modelu i w ogóle jego skonstruowanie. Zwróćmy uwagę, że z tego, co zostało tutaj napisane o użytkownikach informacji, wynika m.in. wniosek, iż na obecnym etapie wiedzy efektywnymi mogą być tylko takie badania użytkowników, których zakres jest na tyle wąski, by móc dostrzec wszystkie istotne (wskazane wcześniej) czynniki i je zanalizować. Praktyka potwierdza małą użyteczność badań "potrzeb informacyjnych" różnych kategorii użytkowników (chemików, pracowników nauki, kierowników itd.), jeśli są to osoby realizujące nieporównywalne cele, rozwiązujące nieporównywalne problemy w różnych warunkach organizacyjnych i informacyjnych. Wydaje się, że zostało wykazane, iż tzw. potrzeby informacyjne są zmienną dynamiczną, zależną od:

- celów i zadań wynikających z aktywności użytkownika;
- warunków, które umożliwiają realizowanie tych zadań i osiągnięcie celów;
- cech indywidualnych użytkownika;
- badanego obszaru wyznaczonego w zakresie określonym przez czynniki wymienione w poprzednich punktach.

Prawidłowo przeprowadzone badania użytkowników systemów informacyjno-wyszukiwawczych są w stanie dostarczyć konkretnych informacji wspomagających procesy projektowania systemów informacyjno-wyszukiwawczych. Niezbędne jest jednak uwzględnienie wymienionych wyżej założeń i przesłanek oraz zastosowanie rzetelnego aparatu metodycznego<sup>16/</sup>.

8.05.1989 r.

#### ON SO CALLED INFORMATION NEEDS

##### Summary

The author formulates some remarks on the meaning of the term "information needs" which is widespread in information science. This term is criticized view of its theoretical incorrectness and operational unusefulness. The outline of the most frequently used patterns of the investigation of information needs is presented. Then, there is formulated the proposal aiming at the building of a pattern in which the enigmatic concept of information need could be replaced

---

<sup>16/</sup> W artykule ze względu na ograniczenia co do objętości nie było mowy o metodyce realizowania badań użytkowników w nowym schemacie. Wartościowe informacje na ten temat można znaleźć między innymi w pracach T.O. Wilsona, na przykład takich jak: Office automation and information services: final report on a study of current developments. Sheffield [b.r.]; T.O. Wilson, D.R. Streatfield: You can observe a lot ... . A study of information use in Local Authority Social Services Departments. Sheffield 1980. Occasional Publications Serie. No 12.

with the analysis of goals to be achieved by means of information and with the analysis of the users' behaviour when storing the information.

## О ТАК НАЗЫВАЕМЫХ ИНФОРМАЦИОННЫХ ПОТРЕБНОСТЯХ

### Резюме

Автор формулирует некоторые замечания на значение термина "информационные потребности", который является широко распространенным в научной информации. Этот термин подвергает критике из-за его теоретической неправильности и операционной неполезности. Представляет характеристику наиболее часто используемых образцов исследования информационных потребностей. Затем формулирует предложение намеренное на создание образца, в котором нечеткое понятие информационной потребности можно бы заместить анализом целей достигаемых с помощью информации, а также анализом поведения потребителя при сборе информации.



ARKADIJ W. SOKOŁOW

Leningradzki Instytut Kultury  
im. N.K. Krupskiej

### POTRZEBY INFORMACYJNE\*

Różnorodność i sprzeczność definicji potrzeb informacyjnych. Pojęcie potrzeb informacyjnych w naukach społecznych i w informacji naukowej.

Przypomnijmy: "pojęcie potrzeb informacyjnych stało się jednym z centralnych pojęć w teorii informacji naukowej i działalności biblioteczno-bibliograficznej. Zaspokajanie potrzeb informacyjnych traktowane jest obecnie jako podstawowy cel działania placówek informacji naukowo-technicznej i bibliotek, jako sposób realizowania przez te instytucje ich funkcji społecznych" /4, s.1/. Wiele już napisano o naturze i istocie potrzeb informacyjnych; od dawna trwają prace nad sformułowaniem teorii potrzeb informacyjnych - jednakże, jak przekonująco dowiódł D.I. Blumenau, problem potrzeb informacyjnych pozostaje wciąż nierozwiązany, mimo "boom'u" publikacji na ten temat /2/. Dlaczego? Przyczyny są co najmniej dwie.

---

\*tytuł oryginału: Информационные потребности и информационный подход к потребностям.

---

"Zagadnienia Informacji Naukowej" 1989 nr 1(54)

Pierwsza przyczyna to dyletantyzm. Większość autorów piszących o potrzebach informacyjnych nie wglębia się w problematykę potrzeb i zainteresowań, dyskutowaną w filozofii, socjologii, ekonomii i psychologii /9/. W rezultacie wyjściowe definicje w rodzaju: "potrzeby informacyjne - to zapotrzebowanie na informację, konieczność posiadania, deficyt informacji" są zupełnie nieprzydatne, same potrzeby informacyjne wynikają bowiem z całego systemu specyficznych potrzeb współczesnego człowieka (grupy, społeczeństwa). Naiwne próby różnych fachowców modelowania potrzeb informacyjnych a priori zaciemniają jedynie nadzwyczajną złożoność i zagmatwanie problemu ludzkich potrzeb, chociaż przyświeca tym próbom wzniosły ideał zaopatrzenia w informację działań produkcyjnych w sposób zróżnicowany i wyprzedzający zapotrzebowanie. Zjawisko "niezgłaszania potrzeb informacyjnych", barwnie opisane przez O.E. Bułyj-Smarjana /3/ to nie przykry przypadek, lecz następstwo niezrozumienia prawdziwych duchowych potrzeb uczonych i specjalistów. Na tle dyletanckiej beztroski, z jaką traktowane jest to zagadnienie, prace S.D. Kogotkova wydają się być pocieszającym i obiecującym wyjątkiem. Tworzą one podstawę dla rozwoju kierunku psychologicznego w teorii potrzeb informacyjnych.

Druga przyczyna. Nie mając jasnego wyobrażenia o tym, czym jest informacja, nie można wyjaśnić czym jest zapotrzebowanie na informację. W literaturze naukowej, potrzeby informacyjne określają przynajmniej trzy definicje:

a. Organiczne (biologiczne) potrzeby dopływu do żywych organizmów informacji sensorycznych, to znaczy odbieranych zmysłami, których uzyskanie jest równie niezbędne dla normalnych procesów życiowych jak przemiana materii i wymiana energii ze środowiskiem zewnętrznym ("metabolizm informacyjny").

b. Duchowe potrzeby człowieka (grupy, społeczeństwa) otrzymywania informacji społecznej (potrzeby społeczno-informacyjne).

c. Zawodowe potrzeby w zakresie informacji naukowo-technicznej uczonych i specjalistów z różnych dziedzin gospodarki narodowej<sup>1/</sup>. W świetle tego ujęcia, szeroko rozpowszechnionego w krajowym systemie informacji naukowo-technicznej, okazuje się, że uczniowie, robotnicy, kolchoźnicy, emeryci właściwie nie posiadają potrzeb informacyjnych.

Rozbieżność i niespójność istniejących poglądów na temat informacji zmusza do wyjaśnienia treści pojęcia "potrzeb informacyjnych", przy jednoczesnym pominięciu terminu "informacja". Istnieją dwie definicje tego typu, nie obejmujące jednak biologicznego aspektu problemu.

1. Potrzeby informacyjne - to potrzeby pozyskania wiedzy. Można je przedstawić jako różnicę pomiędzy wiedzą potrzebną a wiedzą posiadaną. Jeśli człowiek wie, jak rozwiązać określone zadanie - nie posiada potrzeb informacyjnych; jeśli jednak tego nie wie - wówczas informacja jest mu niezbędna. Wadą tej definicji jest fakt, że potrzeby informacyjne - chcąc nie chcąc - utożsamiane są z potrzebami poznawczymi.

2. Potrzeby informacyjne - to potrzeby kontaktów międzyludzkich. Niewątpliwie, potrzeby tego typu są podstawowymi potrzebami duchowymi ludzkości, zwanymi tradycyjnie potrzebami komunikacyjnymi. Wydaje się, że w tym przypadku potrzeby informacyjne są tożsame z potrzebami komunikacyjnymi.

---

<sup>1/</sup> Dla przykładu: "...jako potrzeby informacyjne rozumiana jest konieczność uzyskania informacji zgodnej z charakterem realizowanych działań lub prac". (Terminologičeskij slovar po informativkie. Red. . zespół: S.T. Arseniev i in. Moskva 1975 s. 159)



Gdyby zgodzić się na przedstawione definicje i wziąć pod uwagę istnienie mocnej wzajemnej więzi i pewnych wspólnych zakresów pojęciowych poznania i komunikacji, to należałoby uznać, że potrzeby informacyjne przekształcają się w potrzeby komunikacyjno-poznawcze (poznawczo-komunikacyjne). To jednak nie wyczerpuje zagadnienia.

Podobnie jak społeczeństwo (grupa, jednostka) nie może funkcjonować bez uczestniczenia w procesach poznawczych i komunikacyjnych, tak samo nie może ono istnieć bez pamięci społecznej (grupowej, jednostkowej). Na równi z potrzebami komunikacyjno-poznawczymi ludzkość posiada właściwe sobie potrzeby gromadzenia (magazyinowania, akumulowania) i wyszukiwania wytworów kulturowych (to znaczy faktów, idei, wiedzy społecznej, spuścizny kulturalnej w ogóle). Nie ma żadnych podstaw, aby te potrzeby kumulacyjno-wyszukiawcze nie mogły być traktowane jako potrzeby informacyjne.

Ostatecznie, współcześni specjaliści i ludzie w ogóle potrzebują nie tyle relewantnej informacji, wedle zasady: "im więcej, tym lepiej", lecz raczej chcą się uchronić przed informacyjnym nadmiarem (przeciążeniem). Działalność placówek informacji naukowo-technicznej w zakresie analizy i syntezy informacji stymulowana jest rzeczywistą potrzebą selekcji wartościującej (kontroli, regulowania, kierowania) efekt procesów komunikacyjno-poznawczych. Potrzebę wartościowania możemy również z pełnym uzasadnieniem traktować jako potrzebę informacyjną.

W rezultacie potrzeby informacyjne mogą przejawiać się na różne sposoby i w rozmaitych formach: komunikacyjno-poznawczych, kumulacyjno-wyszukiawczych, wartościująco-orientujących - chociaż jako takie nie posiadają samodzielnej formy. Dlatego z logicznego punktu widzenia błędne jest twierdzenie, że specjalista posiada,

na przykład potrzeby informacyjne, komunikacyjne, poznawcze i selekcji wartościującej. Wymienione typy potrzeb nie są równorzędne, ostatnie z nich istnieją obiektywnie, poddają się obserwacji i pomiarom, natomiast potrzeby informacyjne..?

Zatem nasuwa się wniosek, że najistotniejszy z praktycznego i teoretycznego punktu widzenia problem potrzeb informacyjnych znajduje się w metodologicznym impasie. Programy działalności placówek informacji naukowo-technicznej i bibliotek skierowane na zaspokajanie potrzeb zawodowych i pozazawodowych dezorientują ze względu na nieostre i powierzchowne definiowanie tych potrzeb. Są nawet powody, aby wątpić w obiektywne ich istnienie. W tej sytuacji zapotrzebowanie na nowe idee staje się problemem bądźco ważnym.

Zanim zajmiemy się potrzebami informacyjnymi, należy uściślić rozumienie terminu "potrzeby", mającego status interdyscyplinarny. Odłóżmy na bok często spotykane, tautologiczne definicje, określające potrzeby jako "stan zapotrzebowania", "nieodzowność w danych warunkach", "sprzeczność pomiędzy posiadaniem a niezbędnym", "deficyt czegoś ważnego", itp. Znacznie bardziej efektywne wydaje się być traktowanie potrzeb jako funkcjonalnych cech systemów, wynikających z oczywistej przesłanki, iż potrzeby - to cecha wyróżniająca istoty żyjące, to atrybut życia.

Istoty żywe i ich społeczności stanowią dynamiczne, samoregułujące się i samorozwijające się systemy. Każdy typ żywego systemu posiada charakterystyczne dlań, specyficzne mechanizmy formułowania, doznawania i zaspokajania potrzeb; jednakże funkcje owych potrzeb są w każdym przypadku takie same: sygnalizująca (obzwierciedlająca) i pobudzająca. Sygnalizująca funkcja potrzeb polega na wytwarzaniu sygnału o pojawieniu się niezgodności (rozdźwięku) zarówno zewnętrznych - pomiędzy istniejącymi warunkami środowiska

zewnątrznego i warunkami niezbędnymi dla normalnej egzystencji danego systemu - jak i sprzeczności wewnętrznych, czyli naruszenia wewnętrznej stabilności procesów życiowych (brak pożywienia, destabilizacja sfery psychicznej, itp.).

Sygnał o pojawieniu się niezgodności może zostać odebrany podświadomie przez system nerwowy, albo też przez świadome stosowanie przez człowieka lub społeczeństwo procesów poznawczych. W każdym jednak przypadku nie jest on biernym odbiciem złożonej sytuacji, lecz posiada aktywną, pobudzającą siłę. Pobudzająca funkcja potrzeb przejawia się w aktywizacji żywego systemu w kierunku określonych działań, kompensujących istniejące niezgodności. Sygnał wskazujący na pojawienie się niezgodności ujawnia się jako doznanie fizjologiczne: (głodu, pragnienia, chłodu itp.) oraz jako psychiczne stany niezadowolenia, niepokoju, rozdrażnienia. Kiedy sprzeczność znika, potrzeby dezaktywizują się i ich pobudzający wpływ wygasa, aby pojawić się znów podczas kolejnego cyklu ujawnienia się niezgodności.

Aby przytoczone rozważania nie zasłużyły sobie na miano domorosłego teoretyzowania - powołajmy się na autorytety. Psycholog K.K. Płatonov twierdzi, iż "potrzeby - to zjawisko psychiczne, będące reakcją na pojawiające się braki powodujące naruszenie równowagi pomiędzy organizmem, jednostką i grupą z jednej a środowiskiem (wewnętrznym albo zewnętrznym) lub społeczeństwem z drugiej strony. Tego rodzaju reakcja składa się zawsze z dwóch elementów - aktywności ukierunkowanej na zaspokojenie wynikłej potrzeby, w równoczesnym związku z drugim niezbędnym elementem - doznawaniem potrzeb" /10, s.106/. Filozofowie piszą, że "...potrzeba jest wynikiem wzajemnej reakcji obiektywnych cech środowiska, w którym uformułował się i funkcjonuje dany obiekt oraz obiektywnych cech samego funkcjonującego obiektu. Mechanizm "pojawiania się" tych po-

trzeb polega na odwzorowaniu przez podmiot różnicy pomiędzy stanem istniejącym a pożądanym, ujawniające się jako instykt, doznanie, rozumowanie bądź idealizacja. Tak więc potrzeby to uznanie czegoś brakującego za coś niezbędnego" /11, s.10/. Na koniec, jeszcze jeden pogląd: "...potrzeby - to podstawowa właściwość jednostki, polegająca na dążeniu do posiadania elastycznego systemu sygnałów, stale regulujących jej cykliczne współdziałanie ze środowiskiem" /1, s.94/.

Reasumując, co zostało powiedziane, otrzymujemy następującą definicję: **potrzeby - to podstawowa zdolność żywych systemów do aktywnej reakcji na niezgodność pomiędzy istniejącymi a wymaganymi wewnętrznymi i zewnętrznymi warunkami ich funkcjonowania.** Innymi słowami, potrzeby to szczególne zdolności aktywizowania innych, określonych zdolności żywych systemów przy pojawianiu się niezgodności naruszających niezmienność procesów życiowych.

Potrzeby ludzkie dzielone są powszechnie na dwa rodzaje: biologiczne (biogenne, witalne, organiczne) i społeczno-kulturowe (socio-genne, społeczne). Te ostatnie dzielą się dalej na potrzeby materialne i duchowe, wynikające z wymogów i uwarunkowań społecznych. Nie ma potrzeby podejmować dyskusji na temat wzajemnych relacji różnego rodzaju potrzeb i ich struktur na poziomie jednostkowym i społecznym. Interesują nas jedynie te potrzeby, które możemy nazwać "informacyjnymi". Tego typu potrzeby pojawiają się tylko w tym przypadku, gdy dokonuje się **informacyjnego podejścia** do problematyki potrzeb, to znaczy kiedy potrzeby rozpatrywane są przez pryzmat kategorii informacji.

Podejście informacyjne zbliżone jest do takich ogólnonaukowych ujęć metodologicznych, jak podejście systemowe, funkcjonalne bądź operacyjne. Jest ono metodą poznawczą, mającą ambicje opisanie i wyjaśnienia obiektywnej rzeczywistości za pomocą systemu po-

jęciowego, opartego na bazie pojęcia informacji (systemu, funkcji, działalności operacyjnej). Równocześnie - i to jest zaletą podejścia informacyjnego - podmiot poznający może zaproponować własną interpretację pojęcia informacji, ponieważ nie istnieje ogólnie przyjęta jej definicja. Matematycy i fizycy, cybernetycy, fizjologowie i biolodzy, pedagodzy, psychologowie, socjologowie, językoznawcy i łącznościowcy przyjmują ujęcia, które są wygodne z ich punktu widzenia, co powoduje istnienie w nauce różnorodnych i rozbieżnych definicji informacji, a także brak powszechnie uznanego, uogólnionego pojęcia informacji.

Pomiędzy podejściem informacyjnym a informacją istnieje zależność - informacja pojawia się tam, gdzie stosowane jest podejście informacyjne do poznawania rzeczywistości - i na odwrót: podejście informacyjne nie jest możliwe bez stosowania pojęcia informacji /12/. W relacji tej czynnikiem pierwszoplanowym jest podejście informacyjne, zaś stosowanie pojęcia informacja - konsekwencją tego podejścia. Informacyjne podejście stosowane jest nie tylko we wszystkich niemal współczesnych dyscyplinach nauki (stąd też wynika status informacji jako kategorii ogólnonaukowej), lecz także stanowi bazę pojęciową dla szeregu dziedzin nauki związanych z informacją, bazujących na pojęciu informacji, a w konsekwencji - także na podejściu informacyjnym.

Zasady stosowania podejścia informacyjnego są bardzo proste. Obiekty systemowe, zwłaszcza zaś żywe systemy, charakteryzują procesy odbicia, gdyż odwzorowanie jest atrybutem materii. Te procesy odwzorowania, których typowymi przedstawicielami w naukach społecznych są procesy poznania, zarządzania i komunikacji, noszą miano procesów informacyjnych. W systemach tych bez trudu można wskazać nadawców (przekazicieli) i odbiorców (użytkowników), kanały, komunikaty, kody, zakłócenia, szumy i inne standardowe elementy syste-

ów informacyjnych. W przypadku, gdy nadawcy i użytkownicy informacji są systemami żywymi, wówczas stanowią podmioty działalności informacyjnej i można im przypisać potrzeby, zainteresowania, motywacje. Przejście do informacyjnego sposobu wyrażania procesów odwzorowania, a także - co należy podkreślić - procesów organizacyjnych, to znaczy zmniejszających entropię tworzy przesłanki dla wykorzystywania matematycznego aparatu teorii Shannona lub innych teorii matematycznych, co stosowane jest często przez fizyków, cybernetyków i łącznościowców. Zresztą odwoływanie się do matematyki nie jest konieczne. S.D. Kogotkov wyjaśniając mechanizm powstawania potrzeb informacyjnych /5; 4/ z powodzeniem wykorzystywał podejście informacyjne, obywatelając się bez wzorów matematycznych.

Większa otwartość podejścia informacyjnego nie może jeszcze wyjaśniać jego popularności w nauce. Znacznie ważniejszą przyczyną jest możliwość uogólnienia, właściwa tej metodzie poznawczej. Podejście informacyjne jest to podejście uogólniające, pozwalające na dostrzeganie tego, co wspólne w różnych jakościowo zjawiskach. Wyrazistym tego przykładem jest w tym względzie pojęcie informacji społecznej. Pojęcie to stosowane bywa do opisu wymiany wiedzy w procesach komunikacji społecznej lub opisu jej generowania w procesie poznania społecznego. Wykorzystywane bywa ono także do określania stopnia emocjonalnego oddziaływania wytworów sztuki (informacja estetyczna jest odmianą informacji społecznej), a także służy do przedstawiania oddziaływań w zarządzaniu społecznym. Okazuje się, że **informacja społeczna** to kategoria, która w ramach podejścia informacyjnego służy do opisywania i wyjaśniania sposobów uzyskiwania i przekazywania wiedzy, wpływów emocjonalnych i stymulatorów w procesach zarządzania w środkowisku społecznym.

Pojęcie potrzeb w informacji społecznej, to znaczy potrzeb społeczno-informacyjnych, które tak często pojawiało się w różnych

kontekstach naszego artykułu - jest także pojęciem uogólniającym, lub - jeśli ktoś chce inaczej - pojęciem "zbierającym" (syntezującym). Oznacza ono cechę funkcjonalną ludzi, sygnalizującą niedostateczny poziom wiedzy, niezrównoważenie sfery emocjonalnej albo brak orientacji w danej sytuacji, a także pobudzającą do wykorzystywania zasobów i środków poznawczych i komunikacyjnych do kompensowania powstałych rozdzźwięków. Niemniej jednak, same potrzeby informacyjne nie są cechą funkcjonalną ludzi. Ze względu na swój abstrakcyjny charakter, potrzeby informacyjne (podobnie jak sama informacja) występują nie jako zjawisko realne, lecz jako pojęcie, służące do poznawania istoty danej rzeczywistości. Obiektywnie (to znaczy niezależnie od podmiotu poznającego) wyróżniamy potrzeby komunikacyjne, poznawcze, mnemoniczne i selekcji wartościującej. Można badać te potrzeby i opracowywać praktyczne sposoby ich zaspokajania nie interpretując ich w kategoriach informacyjnych. Niemniej nie można stosować podejścia informacyjnego bez określenia pojęcia potrzeb informacyjnych, podobnie jak nie można już obejść się bez pojęcia informacji, działalności informacyjnej, środowiska informacyjnego, itp. Tak więc w naukach społecznych i w ramach podejścia informacyjnego - **potrzeby informacyjne są pojęciem uogólniającym, oznaczającym całokształt potrzeb duchowych w sferze poznania, komunikacji, zapamiętywania i orientacji, właściwych poszczególnym ludziom, grupom społecznym i społeczeństwom.**

Łatwo wykazać, że organiczne (biogenne) potrzeby informacyjne nie są niczym innym, jak wynikiem podejścia informacyjnego do procesów fizjologicznych i psychologicznych, związanych z obcowaniem żywych organizmów ze środowiskiem zewnętrznym i pomiędzy sobą. Podobne ujęcie uprawnione jest także w naukach biologicznych.

Nas jednak najbardziej interesują profesjonalne potrzeby informacyjne, które - w świetle tego, co dotąd powiedziano - stanowią

całokształt potrzeb komunikacyjnych, poznawczych, mnemoniczych i wartościujących - cechujących nie wszystkich ludzi w ogóle, lecz tylko poszczególne grupy zawodowe i specjalistów. Chodzi tu już nie o wiedzę społeczną i wartości duchowe w ogóle, lecz o specjalistyczną wiedzę i umiejętności, których potrzeba wynika z różnicy między potrzebami a poziomem kwalifikacji (kompetencji) kadr specjalistów.

Prześledźmy konsekwencje praktyczne, wynikające z powyższych rozważań teoretycznych. Placówki informacji naukowo-technicznej, stawiające sobie za główny cel zaspokajanie potrzeb informacyjnych uczonych i specjalistów, nie mogą zrealizować tego celu inaczej, niż tylko drogą zaspokajania abstrakcyjnie wyobrażalnych potrzeb informacyjnych, to znaczy potrzeb: komunikacyjno-poznawczych, kumulacyjno-wyszukiwawczych i duchowych, wartościujących. Wynikają stąd trzy istotne (immanentne), społeczne funkcje placówek informacji naukowo-technicznej, odpowiadające trzem wspomnianym kategoriom potrzeb. Konkretne źródła informacji i usługi informacyjna pełnią zazwyczaj jednocześnie kilka funkcji, tym niemniej można je pogrupować według funkcji dominującej:

- funkcja komunikacyjno-poznawcza dominuje w przypadku przeglądów dokumentacyjnych, abstraktów, informacji sygnalnej, selektywnej dystrybucji informacji, dialogowej obsługi informacyjnej, itp.

- funkcja kumulacyjno-wyszukiwawcza przeważa przy gromadzeniu i organizacji zbiorów informacyjnych, w systemach informacyjno-wyszukiwawczych, zbiorach bibliotecznych i archiwalnych, bankach danych systemów informacji naukowo-technicznej w państwowych zasobach archiwalnych, bibliotekach prywatnych, kartotekach i in.



- funkcja selekcji wartościującej (zalecająca, oceniająca) wychodzi na pierwszy plan przy prowadzeniu prac naukowo-badawczych i konstrukcyjno-doświadczalnych, w systemach obsługi kadr kierowniczych (informacja tematyczna oraz analityczno-syntetyczna), w propagandzie nauki i techniki, w przeglądach analitycznych, bibliografiach zalecających, ekspertyzach, itp.

W praktyce, prawdziwe funkcje są zawoalowane i przybierają postać funkcji stosowanych (celowościowych), np.: wspieranie postępu technicznego, kształcenie ustawiczne, wdrażanie optymalnych doświadczeń, itp. Realizacja "funkcji celowościowych" możliwa jest jedynie poprzez wypełnianie konkretnych funkcji, odgrywających rolę "funkcji-środków". Dlatego też, właśnie te prawdziwe zadania mają kluczowe znaczenie dla realizacji przez placówki informacji naukowo-technicznej (biblioteki, archiwa, służby bibliograficzne i inne) ich podstawowej funkcji społecznej, jaką jest zaspokajanie potrzeb informacyjnych.

Reasumując, z punktu widzenia podejścia informacyjnego uprawnione są wszystkie trzy interpretacje pojęcia potrzeb informacyjnych, występujące w literaturze naukowej, ponieważ wiążą się one z odmiennym rozumieniem pojęcia informacja. Niemniej, określanie potrzeb informacyjnych przy zastosowaniu pojęcia informacji jest w każdym z tych przypadków bezpłodne. W celu określenia potrzeb informacyjnych należy odnaleźć te realnie istniejące potrzeby, które posiada przedmiot, do których można zastosować podejście informacyjne. Właśnie owe rzeczywiste, a nie abstrakcyjne potrzeby informacyjne należy zaspokajać, właściwie ukierunkowując działalność służb informacyjnych.

(Artykuł poświęcony pamięci S.D. Kogotkova)

## Literatura

1. BELIK A.P. Social'naja forma dviženija. Moskva 1982 270 s.
2. BLJUMENAU D.I. K utočneniju ischodnych ponjatij teorii informacionnyh potrebnostej. Naučno-techničeskaja Informacija Ser. 2 1986 nr 2 s.7-12.
3. BURYJ-ŠMAR'JAN O.E. Problema "nepotreblenija" informacii: kak ee rešit' ? Naučno-techničeskaja Informacija Ser. 1 1984 nr 10 s. 13-21.
4. KOGOTKOV S.D. Formirovanie informacionnyh potrebnostej. Naučno-techničeskaja Informacija Ser. 2 1986 nr 2 s.1-7.
5. KOGOTKOV S.D. Nekotorye teoretičeskie problemy razrabotki psihologičeskoj modeli povedenija čitatelej. Tez. dokl. i soobšč. konf. po itogam naučno-issledovatel'skoj raboty Gos. biblioteki SSSR im. V.I. Lenina za 1982 god. Moskva 1983 s. 44-46.
6. KOGOTKOV S.D. Nekotorye voprosy teorii informacionnyh potrebnostej. Naucno-techničeskaja Informacija Ser.1. 1979 nr 2 s.1-3.
7. KOGOTKOV S.D. Professional'nye informacionnye potrebnosti i ich rol' v bibliotečno-bibliografičeskoj dejatel'nosti. [Autoreferat do pracy doktorskiej]. Moskva 1979 16 s.
8. KOGOTKOV S.D. Sravnitel'nyj analiz putej izučenija informacionnyh potrebnostej. Teoretiko-metodologičeskie problemy sovremennogo sovetskogo bibliografovedenija. Mežvuznyj Sbornik Naučnyh Trudov Moskva 1981 s.85-98.
9. MARGULIS A.V., Stepanov E.I. Potrebnosti čeloveka: metodologičeskie problemy kompleksnogo issledovanija (Analitičeskij obzor literatury). Voprosy Filosofii 1986 nr 4 ž.54-68.
10. PLATONOV K.K. Problemy sposobnostej. Moskva 1972 312 s.
11. PROBLEMA potrebnostej v etike i estetike. Leningrad 1976 175 s.
12. SOKOLOV A.W. Informacija a podejście informacyjne. Zagadnienia Informacji Naukowej 1987 nr 1(50) s. 3-22.

02.04.1989 r.

## INFORMATION NEEDS

### Summary

Multiplicity and incompatibility of the existing in science formulations of the term information needs is explained as the result of informational approaches to different realities. The information needs interpreted in the realm of social sciences represent a generalizing concept which refers to the complexity of the inner needs within the processes of cognition, communication, memorizing and orientation. Information need is not the real property of living systems but it is an abstract concept used within the informational approach for the description and explanation of the properties of needs.

## ИНФОРМАЦИОННЫЕ ПОТРЕБНОСТИ

### Резюме

Разнообразие и несовместность бытующих в науке трактовок информационной потребности объясняется тем, что они являются результатом информационного подхода к различным реалиям. В сфере общественных наук информационная потребность – это обобщающее понятие, обозначающее совокупность духовных потребностей в познании, коммуникации, запоминании, ориентации, свойственных отдельным людям, коллективам, обществам. Информационная потребность не реальное свойство живых систем, а абстрактное понятие, служащее в рамках информационного подхода для описания и объяснения потребностных свойств.

LEONID V. SACHARNYJ

Państwowa Biblioteka Publiczna  
w Leningradzie

**STRUKTURY KOJARZENIOWE W JĘZYKU NATURALNYM**  
**I JĘZYKU INFORMACYJNO-WYSZUKIWAWCZYM**  
**(na przykładzie języka haseł przedmiotowych)**

Rozumienie relacji kojarzeniowych w języku informacyjno-wyszukiawczym (jiw). Struktury kojarzeniowe w mowie ludzkiej. Pole kojarzeniowe w języku naturalnym. Relacje między jednostkami leksykalnymi jiw. Dialog użytkownika z systemem jako współdziałanie struktur kojarzeniowych. Typy kompensacyjnych struktur kojarzeniowych w jiw.

W artykule podjęto próbę przedstawienia wszystkich typów relacji kojarzeniowych w języku informacyjno-wyszukiawczym (na przykładzie języka haseł przedmiotowych) z punktu widzenia uniwersalnych mechanizmów kojarzenia charakterystycznych dla języka naturalnego.

**RELACJE KOJARZENIOWE W JĘZYKU INFORMACYJNO-WYSZUKIWAWCZYM**

We współczesnej działalności informacyjnej przyjęło się rozumienie relacji kojarzeniowych jako wszystkich tych, które w ramach

---

\*tytuł oryginału: Asocjativnye struktury v estestvennom jazyke i predmetizacjonnom ipja (k postanovke problemy), W: Predmetnyj polak v tradicionnych i netradicionnych informacionno-poiskowych sistemach Leningrad 1988 s. 5-35.

---

"Zagadnienia Informacji Naukowej" 1989 nr 1(54)

relacji paradygmatycznych nie mieszczą się w grupie logicznych relacji równoznaczności lub podrzędności i nadrzędności zakresowej.

I tak "Słownik terminów z dziedziny informacji naukowej" podaje: "Często relacje kojarzeniowe rozumiane są jako wszelkie relacje paradygmatyczne, oprócz relacji rodzaj-gatunek i część-całość /10, s. 149/.

Norma GOST 7.25-80 zatwierdza następującą definicję relacji kojarzeniowej: "Relacja kojarzeniowa jest połączeniem innych relacji nie wchodzących w skład relacji hierarchicznych i relacji synonimii. Dopuszcza się włączenie do relacji kojarzeniowych wszystkich typów relacji z wyjątkiem relacji synonimii i relacji rodzaj-gatunek /2, s. 9/.

Relacjami kojarzeniowymi są więc stosunki: system - element systemu, część - całość, przyczyna - skutek, proces - urządzenie, proces - materiał, surowiec - wyrób, wyrób - cecha, wyrób - proces wytwórczy, nauka - jej obiekty, nauka - jej przedstawiciele, przedmiot - cecha, kierunek - jego realizatorzy, kontynuatorzy /3, s. 67-68/.

Relacje kojarzeniowe "wyrażają związki nie między pojęciami, a między samymi przedmiotami (technologiczne, systemowe i inne wzajemne zależności)" /3, s. 67/.

Takie podejście do relacji kojarzeniowych, wygodne ze względów pragmatycznych, podkreśla ich odrębność i różnice w stosunku do relacji logicznych, nie pokazuje ich cech wspólnych, co w jakiś sposób utrudnia efektywne modelowanie wszelkich relacji, a jest to potrzebne do podwyższenia jakości systemów informacyjno-wyszukiwawczych.

Warto zauważyć, że w początkowym okresie rozwoju systemów informacyjno-wyszukiwawczych podejmowano próby przeniesienia całej

złożoności struktur kojarzeniowych z języka naturalnego do języków informacyjno-wyszukiwawczych /6, s.435-446/. Jednakże później - tworząc tezaury - zwracano coraz mniejszą uwagę na uniwersalność mechanizmu kojarzenia w mowie ludzkiej. Podejście do struktur kojarzeniowych coraz bardziej upraszczano, "technologizowano", co w efekcie wpłynęło na dzisiejsze ograniczenia w rozumieniu relacji kojarzeniowych. Trzeba podkreślić, że różnice w podejściu do rozumienia skojarzeń w porozumiewaniu naturalnym i w systemach informacyjno-wyszukiwawczych nie stanowią tylko problemu terminologicznego. Uniemożliwia to dostrzeżenie w systemach informacyjno-wyszukiwawczych rzeczywistych prawidłowości i podobieństw, a dokładniej tej samej podstawy procesów kojarzeniowych w języku naturalnym i języku informacyjno-wyszukiwawczym. Zagadnienie to jest bardzo ważne dla tradycyjnych, a szczególnie ważne dla zautomatyzowanych systemów informacyjnych.

#### STRUKTURY KOJARZENIOWE W MOWIE LUDZKIEJ

Intensywne badania eksperymentalne nad strukturami kojarzeniowymi prowadzone są mniej więcej od początku naszego wieku. Sprawdzają się one do podania spisu słów-stymulatorów (S) z prośbą o wskazanie dla każdego z nich słów, które kojarzą się odpowiadającemu z S. Jest to tzw. reakcja (R), czyli odpowiedź na słowo-stymulator S. Oto przykład artykułu słownikowego wybranego ze "Słownika norm kojarzeniowych języka rosyjskiego" /9/. Stymulatorem jest słowo Drug - przyjaciel, na które otrzymano 209 odpowiedzi: tovarišč (39), wrağ (30), vernyj (29), chorošij (16), moj (10), nedrug, blizkij (5), nastojaščij, staryj (4), brat, dorogoj, nadežnyj, predannyj, prijatel' (3), zakadyčnyj, lučšij, ljubimyj, on, sobaka (2), bol'soj, večnyj, v nužde, volk, dalekij, devuška, detstva, dobryj,

toroga, edinstvennyj, želannyj, žensčcina, ljubov', mal'čik, milyj, Miska, muž, navsegda, ne veritsja, net, obščij, otkrovennyj, parta, pervyj, plochoj, podlost', podrug, predatel', prichodit', protivnik, serdečnyj, sestra, siniy, sobaka - drug čeloveka, čestnyj, čudesa (1) /9, s.94/.

Artykuł ten przedstawia pole kojarzeniowe i daje możliwość zbadania struktury słownictwa człowieka. Przy czym jak widać niektóre odpowiedzi powtarzają się stosunkowo często (to standardowe, stereotypowe reakcje, wyrażające stałe związki i stanowiące centralną część pola kojarzeniowego). Ale oprócz odpowiedzi standardowych są też odpowiedzi pojedyncze, przypadkowe. Wyrażają one reakcje indywidualne, sytuacyjne.

Z treściowego punktu widzenia skojarzenia dzielą się na trzy podstawowe grupy:

- a) skojarzenia paradygmatyczne, w których dwa słowa S i R związane są relacjami synonimii (dom - zdanie [budynek], drug - tovarišč) lub antonimii (dorogoj - deševyj, drug - vrag), relacjami rodzaj - gatunek lub gatunek - rodzaj (djadja - mužčina, idti - peredvigatsja, komnata - spal'nja, drug - on). (łatwo zauważyć, że tutaj termin "paradygmatyczny" używany jest przez psychologów i psycholingwistów jako synonim terminu "logiczny").
- b) skojarzenia syntagmatyczne, w których dwa słowa S i R związane są relacją współwystępowania w wyrażeniach (w formach słownikowych lub tekstowych), np. zanjatie - dopolnitel'nye, drug - vernyj, zapisat' - muzyka, djadja - sdelat, dat'-ruku, den' - prošel, devočka - malen'kaja, gulat' - v parke.
- c) skojarzenia tematyczne, w których dwa słowa S i R występują jako słowa kluczowe charakteryzujące jakąś sytuację, np.: djadja - futbol, drug - parta, zemlja - agronom, les - lyży. Są to relacje, któ-

re uwzględnia się w tezaursach i często określa bliżej, np.: jako "część - całość" itp. ogólnie nazywane relacjami kojarzeniowymi.

Niektórzy /np. 4, s.45/ uważają, że relacje paradygmatyczne i syntagmatyczne, to związki między słowami, natomiast relacje tematyczne, to związki między sytuacjami i zjawiskami. Jednakże lepiej byłoby powiedzieć, że wszystkie relacje kojarzeniowe mają jednako-  
wy charakter, jednakową naturę - są to zawsze związki między sytuacjami wyrażonymi w słowach. Aczkolwiek jest oczywiste, że związek z sytuacją może być mniej lub bardziej bezpośredni, "skondensowany" w znaczeniu słowa, które uogólnia wielość sytuacji jakiegoś typu. Potwierdzają to przeprowadzone eksperymenty /7,8/.

Zapytywano podawano pary słów wzięte ze słownika /9/ jako pary S i R różnych typów, prosząc o napisanie zdań, w których byłaby użyta podana para. Zdania te w sposób obiektywny ilustrowały rzeczywiste związki istniejące między słowami S i R, obrazując tym samym mechanizm kojarzenia. Wykryte związki odpowiadają trzem grupom sytuacji:

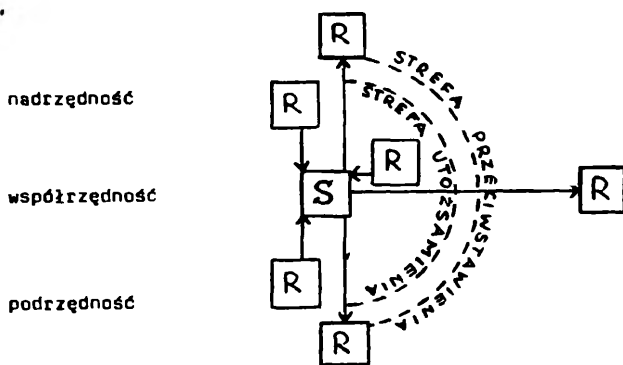
I. Słowo S wydaje się słuchającemu jako ogólne oznaczenie jakiejś sytuacji (lub typu sytuacji) jednym słowem, faktycznie jest to słowo - tekst. Pytany podaje odpowiadające słowu S słowo R oznaczające tę samą sytuację: [ S ] - [ R ], np. dom - zdanie [ budynek ]. Są to w rzeczywistości relacje paradygmatyczne wyrażające logiczne zależności synonimii, antonimii, rodzaju-gatunku, gatunku-rodzaju. Przyczym, jeśli słowa S i R są stosunkowo częste, standardowe także jako człony relacji paradygmatycznych, to abstrahowanie od konkretnych sytuacji i wrażenie, że kojarzone są słowa - tylko jako słowa, a nie oznaczone nimi sytuacje jest bardzo silne. Jednakże jest to tylko wrażenie. Eksperymenty wykazały, że zdania zbudowane ze słów



S-R połączonych relacją paradygmatyczną w 2/3 przypadków wrażają struktury, które praktycznie nie są spotykane w innych typach skojarzeń. Są to następujące struktury:

a) struktura "słowo S jest (nie jest) słowem R", np. moe ljubimoe delo moja rabota; Prijeti ešče ne značit vernutsja.

b) struktura "słowo S i słowo R są X" lub "słowo S i słowo R nie są X", to jest podane słowa S i R w strukturze zdania są członami jednorodnymi, np. Zemlja i voda - naši istočniki žizni. Słowa S i R w paradygmatycznej parze stanowią oddzielne teksty, a zdania zbudowane z takich słów są pewnego rodzaju "metatekstami" opisującymi inną, bardziej ogólną sytuację niż ta, która jest opisana za pomocą oddzielnych słów S i R. W tej nowej sytuacji znaczenia słów S i R są w jakiś sposób zrównane, zestawione w celu ich wzajemnego potwierdzenia lub na odwrót - przeciwstawienia. Przy czym to samo słowo R przez odpowiadającego może być uznane za tożsame z S (porównaj moe ljubimoe zanjatie - moja rabota) traktowane jako synonim sytuacyjny lub przeciwstawne S (porównaj s tobój ni dela ni raboty) traktowane jako antonim. Takie przeciwstawienia mogą przebiegać zarówno w poziomie (stosunki koordynacji, współrzędności) jak i w pionie (stosunki nadrzędności i podrzędności). Ilustruje to poniższy rysunek:



Związki te są podstawą budowy sieci odsyłaaczy w języku informacyjno-wyszukiwawczym. Bardziej dokładne zróżnicowanie relacji rodzajowo-gatunkowych przedstawia sytuacja następująca:

II. Słowo S wydaje się odpowiadającemu częścią znaczenia frazy użytej na określenie jakiejś sytuacji (lub typu sytuacji). I wtedy wskazuje on drugie słowo, które jakby dopełnia słowo S w owej niepełnej frazie (np. para S-R devočka - malen'kaja), co może być interpretowane jako S - devočka, R - devočka malen'kaja. Tę psychologiczną sytuację potwierdzają słowa reakcje typu: S - kurs, R - kurs lekcij, S - načalo, R - chorošee načalo. Innymi słowy, w tych przypadkach słowu S przeciwstawia się nie słowo R, ale całą frazę (syntagmę) R, co wyrazić można jako  $\boxed{S}$  -  $\boxed{(S)R}$ . Stąd też nazwa tego typu skojarzeń - skojarzenia syntagmatyczne (składniowe). W rzeczywistości owe skojarzenia syntagmatyczne dwóch słów pokazują relacje paradygmatyczne między słowem S i całą frazą S-R. Zdania zbudowane z takich składniowych par słów w 96% przypadków wyrażają struktury, w których słowa S i R są rzeczywistymi frazami, np.: Moja babuška živet v bol'šom gorode (gorod - bol'šoj). Takie zdania wyrażają tekst, w którym S i R występują jako pojedyncze frazy, nazywające jakiś fragment opisywanej sytuacji. (Często są to stałe związki frazeologiczne, ale mogą to być też połączenia doraźne). W istocie chodzi tu już o struktury kojarzeniowe obejmujące nie tylko słowa, ale większe struktury składniowe (frazy, zdania). Ze względu na swoją analityczną formę wyrażają one bardziej dokładnie zróżnicowanie rodzajowe sytuacji oznaczanej słowem S.

Na tego typu skojarzeniach składniowych oparta jest budowa haasei przedmiotowych wielowyrzawowych i rozwiniętych, np.:

Woda

Woda destylowana

Woda pitna

Hałas

Hałas - zwalczanie

Hałas - zwalczanie - górnictwo

Hałas - zwalczanie - przemysł

Hałas - zwalczanie - transport

[Przykłady zostały zmienione - przyp. tłum]. Jeszcze większe zróżnicowanie rodzajowo-gatunkowe przedstawia trzecia sytuacja:

III: Słowo S jest dla odpowiadającego częścią analitycznie podzielonego wielowyrazowego oznaczenia sytuacji (lub typu sytuacji). Słowo R jest słowem kluczowym dla tej sytuacji, najważniejszym, najbardziej trafnym. Przy czym także słowo S może być użyte jako słowo kluczowe R. Słowa S i R w połączeniu stanowią minimalny zestaw słów kluczowych, który określa jakąś sytuację. Można to wyrazić jako  $S - (S)...R$ . Te skojarzenia nazywane są tematycznymi. W rzeczywistości skojarzenia tematyczne odzwierciedlają relacje paradygmatyczne między słowem S i zestawem słów kluczowych "S - R".

To, że skojarzenia tematyczne pozwalają odtworzyć pewne sytuacje potwierdzają eksperymenty. Tematyczne pary skojarzeniowe, które okazują się rzeczywistym odpowiednikiem zestawu słów kluczowych ograniczają liczbę typów sytuacji, do których oznaczenia zostały użyte. Najczęściej jest to jedna lub dwie najbardziej typowe sytuacje. Tak, np. 90% zdań, w których wykorzystano słowa pary tematycznej gazeta - zavrtrak [śniadanie] opisuje dwa typy sytuacji:

a) "Za zavrtrakom papa cital gazetu", "Ja ne ljublju, kogda za zavrtrakom citajut gazetu".

b) "Emu žena zavrtrak v etu gazetu zavernula" itp.

Na tego typu skojarzeniach opiera się budowa hasła przedmiotowego rozwiniętego, np.: Książka - ZSRR - bibliografia, Choroby - leczenie

nie - metody niekonwencjonalne [przykłady zmienione - przyp.tłum.]. Typ ten skojarzeń jest też podstawą budowy sieci odsyłaczy uzupełniających typu:

Dewaluacja zob. też Rewaloryzacja

Barok - architektura

zob. też

Leningrad - Pałac Zimowy

Rastrelli Bartolemeo Francesco

Warto podkreślić, że w każdym konkretnym przypadku operacje u-  
tożsamienia lub przeciwstawienia S i R mogą realizować się tak jak  
w relacjach paradygmatycznych lub tak jak w relacjach syntagmaty-  
cznych lub tak jak w relacjach tematycznych. Zależy to od oceny u-  
żytkownika. Na przykład, jeśli skorzysta on z odsyłaczy podanych  
wyżej, to tym samym uznaje w tej konkretnej sytuacji za quasi -  
-synonimy (to jest tożsame). Jeśli nie skorzysta w pełni z przed-  
stawionych odsyłaczy, to znaczy, że hasła te uznał za quasiantoni-  
my (to jest przeciwstawne). Tak więc podczas wyszukiwania może być  
taka sytuacja, że te same hasła zostaną potraktowane raz jako toż-  
same, a innym razem jako przeciwstawne. Jest to uzależnione od  
użytkownika.

Takie szerokie podejście do struktur kojarzeniowych ukazuje,  
że w systemach informacyjno-wyszukiwawczych charakter kojarzeniowy  
mają nie tylko te relacje, które zwykle nazywać się kojarzeniowy-  
mi, ale także te, które nazywane są relacjami paradygmatycznymi i  
w ogóle wszelkie relacje (w tym także i syntagmatyczne).

Podstawowe rodzaje struktur kojarzeniowych w języku natural-  
nym i języku informacyjno-wyszukiwawczym ilustruje tablica:

Związki paradygmatyczne tekstów ujawnione podczas eksperymentów kojarzeniowych	Rodzaje tekstu w którym występuje słowo R	Nazwa relacji między słowami S i R w eksperymencie kojarzeniowym	Nazwa relacji między jednostkami leksykalnymi w jiw analogicznych do relacji między S i R
Tekst <span style="border: 1px solid black; padding: 2px;">S</span> - tekst <span style="border: 1px solid black; padding: 2px;">R</span>	słowo	paradygmatyczne	paradygmatyczne logiczne
Tekst <span style="border: 1px solid black; padding: 2px;">S</span> - tekst <span style="border: 1px solid black; padding: 2px;">SR</span>	fraza	syntagmatyczne	syntagmatyczne
Tekst <span style="border: 1px solid black; padding: 2px;">S</span> - tekst <span style="border: 1px solid black; padding: 2px;">S...R</span>	zestaw słów kluczowych	tematyczne	paradygmatyczne kojarzeniowe

Podsumujmy uwagi o naturze pola kojarzeniowego.

1. Pole kojarzeniowe - to nie prosty zbiór słów R podanych w odpowiedzi na słowo S, ale zbiór różnych tekstów R (lub fragmentów tekstów) które w ten czy inny sposób wchodzą w relacje paradygmatyczne ze słowem S będącym swego rodzaju peryfrazą (najczęściej gatunkową) w stosunku do słowa S. W zależności od typu tekstu R powstają różne typy struktur kojarzeniowych między słowem S a słowem R.

2. Struktury kojarzeniowe, realizujące związki słowa S z tekstami R (lub ich fragmentami) aktualizują się w mowie ludzkiej z różną częstotliwością. Dlatego w polu kojarzeniowym wydzielić można część centralną (jądro), w której występują najczęstsze struktury (stereotypowe) i część peryferyjną, w której występują struktury mniej częste (doraźne, sytuacyjne).

Złożoność analizy struktury pola kojarzeniowego języka naturalnego związana jest po pierwsze z tym, że częstotliwość występowania struktur od centrum do peryferii stopniowo zmniejsza się, ale wszelkie ostre rozgraniczenia są tu niemożliwe. Po drugie ranga struktur kojarzeniowych może się zmieniać wraz ze zmianą grup respondentów. Wreszcie po trzecie granice peryferii pola praktycznie

trudno jest ustalić, ponieważ zawsze istnieje możliwość jego uzupełnienia związana choćby z powiększeniem grupy respondentów. 3. Te same potencjalne logiczne różnice między składnikami struktur kojarzeniowych (to jest między słowami S i tekstami R) w polu kojarzeniowym mogą się wypuklać (quasiantonimia) lub neutralizować (quasisynonimia). A to znaczy, że charakter relacji paradygmatycznych między składnikami pola kojarzeniowego w języku naturalnym nie jest ostry i stały, ale zmienny, zależny od szczegółów sytuacji.

#### DIALOG UŻYTKOWNIKA Z SYSTEMEM INFORMACYJNYM JAKO WSPÓŁDZIAŁANIE STRUKTUR KOJARZENIOWYCH

Do tej pory rozpatrywaliśmy struktury kojarzeniowe w systemie informacyjno-wyszukiwawczym z punktu widzenia typologii relacji kojarzeniowych zachodzących między elementami języka informacyjno-wyszukiwawczego. Teraz interesuje nas, jak realizują się te struktury w procesach językowych w systemie.

Podkreślmy jeszcze raz, że za formami tekstu ukryte są procesy komunikowania. Innymi słowy - w formie tekstu w "skondensowanej" postaci odzwierciedlone są wszystkie aspekty i trudności procesu komunikowania. I jeśli w podejściu strukturalnym widzi się tekst jako strukturę statyczną, to w podejściu funkcjonalnym za ową statyką próbuje się zobaczyć i zbadać, zrekonstruować dynamikę procesu komunikowania.

System informacyjno-wyszukiwawczy spełnia te same warunki, które występują podczas procesu komunikowania w języku naturalnym. Przede wszystkim nie jest to tylko bezosobowa struktura językowa i bezosobowy tekst zbudowany na jej podstawie, ale związek dwóch ludzi. Tekst ten istnieje, ma sens tylko dlatego, że może być włą-

czony w akt komunikowania, umożliwiającą związek między nadawcą (tu: indeksatorem) a odbiorcą (tu: użytkownikiem systemu, próbującym za pomocą tego testu zrealizować swoje potrzeby informacyjne).

Ten, kto włada językiem włada też związkami paradygmatycznymi, które istnieją między elementami pola kojarzeniowego języka naturalnego (przede wszystkim jego centralną częścią). Elementy te mogą być zarówno strukturami niepodzielnymi (wyrazami) jak i podzielnymi (frazami lub zestawami słów kluczowych). Dlatego podczas odbierania tekstu podanego przez nadawcę, odbiorca może aktualizować odpowiednie pole kojarzeniowe (przede wszystkim jego centralną część) i zidentyfikować każdy składnik tekstu jako składnik pola kojarzeniowego, który może podlegać peryfrazie w ramach relacji paradygmatycznych tego pola. Znaczący to, że jednostki leksykalne lub ich kombinacje zawsze są wykorzystywane podczas komunikowania jako składniki odpowiednich pól kojarzeniowych.

Znajomość standardowych struktur pola kojarzeniowego z możliwością zastosowania peryfrazy - to znajomość języka znacznie wyższa niż znajomość tylko zasobu leksykalnego. Przy tym w komunikowaniu w języku naturalnym związki paradygmatyczne w polu kojarzeniowym mają charakter dynamiczny i w zależności od sytuacji mogą się realizować jako quasisynonimy lub quasiantonimy.

W języku informacyjno-wyszukiwawczym dochodzi do sprzeczności między potrzebą różnorodności wyrażania charakterystycznego dla komunikowania ludzkiego a koniecznością wyraźnego wydzielenia formy tekstu (sformułowania charakterystyki wyszukiwawczej dokumentu). Stosuje się tu więc środki zastępcze w postaci całego aparatu odsyłaczowego lub tworzenia odpowiednich słowników. Przy czym środki te mogą być podawane w sposób jawny w postaci algorytmów (tak jest w klasycznych systemach deskryptorowych) lub w postaci "podpowiedzenia", wskazania sposobu [odsyłacze - przyp. tłum.].

Aby w ogóle doszło do "spotkania" użytkownika z poszukiwanym autorem (w tym celu tworzy się katalogi) potrzebna jest zbieżność tekstów w dialogu użytkownik - indeksator. Jest to oczywiste, aczkolwiek nie osiąga się tego ani prosto ani łatwo. Co jest powodem tego stanu rzeczy?

Po pierwsze powodem może być odmienność w rozpoznaniu i wybraniu przedmiotu i jego cech.

Po drugie przedmiot może być nazwany w sposób adekwatny lub nieco uogólniony.

Po trzecie wybrany przedmiot można wyrazić w postaci hasła prostego [złożonego z samego tematu - przyp. tłum.], które stanowić będzie cełą charakterystykę wyszukiwawczą lub w postaci kilku haseł szczegółowych, które w sumie utworzą charakterystykę wyszukiwawczą lub w postaci hasła rozwiniętego [złożonego z tematu i przynajmniej jednego określnika - przyp. tłum.].

Użytkownik formułuje swoje pytanie wychodząc od najbardziej "naturalnych", znanych mu sposobów nazywania interesujących go obiektów. Są to przede wszystkim sposoby zgodne z normami języka naturalnego. Przy czym użytkownik ma zawsze wiele możliwości wyrażenia swego pytania.

Wszystko to stwarza pewną sytuację kolizji i rozbieżności między efektami pracy (to jest tekstami) indeksatora (który zna dany język informacyjno - wyszukiwawczy i stosuje określone schematy budowy charakterystyki wyszukiwawczej) a oczekiwaniami użytkownika (nawet jeśli zna on zasady danego języka informacyjno-wyszukiwawczego). Trzeba też pamiętać, że struktura tekstu wtórnego, jakim jest charakterystyka wyszukiwawcza, zawiera jednostki leksykalne, które zwykle spotyka się w tekście dokumentu. Przeniesienie ich do charakterystyki wyszukiwawczej podczas indeksowania zwiększa ryzyko



ich nieadekwatnego użycia. A ponieważ między indeksatorem a użytkownikiem systemu nie istnieje dialog "naturalny", a łączność między nimi utrzymywana jest tylko za pośrednictwem języka informacyjno-wyszukiwawczego, to możliwość rozbieżności w sformułowaniu charakterystyki wyszukiwawczej i zapytania informacyjnego wzrasta.

Mechanizm rozbieżności między tekstem indeksatora a tekstem użytkownika ma ten sam charakter, co mechanizm odmienności odpowiedzi w języku naturalnym otrzymywanych podczas przeprowadzonego eksperymentu kojarzeniowego.

Zadaniem sztucznych języków informacyjno-wyszukiwawczych jest wydzielenie najważniejszych (dla wyszukiwania), standardowych związków, odzwierciedlających najbardziej standardową i dlatego trafną, część pola kojarzeniowego i włączenie ich do systemu (m.in. w postaci sieci odsyłaczy). W ten sposób zapobiega się najczęściej spotykanym i najbardziej typowym przypadkom rozbieżności między tekstem charakterystyki wyszukiwawczej a tekstem pytania informacyjnego. Wprowadzenie takich specjalnych środków (jak sieć odsyłaczy) do systemu informacyjnego daje możliwość w miarę szybkiej korekty pytania podczas "dialogu" użytkownika z systemem. Minimalny dialog to jedna para: pytanie - odpowiedź. Jeżeli jednak pojawia się jakaś niejasność w informacji potrzebne jest uściślenie pytania. W ten sposób dochodzi do sformułowania pytań dodatkowych i w efekcie do realizacji rozwiniętego dialogu, złożonego z większej ilości pytań i odpowiedzi. Dialog ten pozwala dojść do najbardziej adekwatnego sformułowania zapytania, w szczególności zaś uściślić zbiór jednostek leksykalnych, aktualizowanych w danym polu kojarzeniowym. Sformułowanie głównego zapytania informacyjnego jak i zapytań uzupełniających (także odpowiedzi) jest ściśle związane z umiejętnością aktualizacji struktury pola kojarzeniowego i operowania nim

przez pytającego (odpowiadającego). Swoboda w operowaniu polem kojarzeniowym pozwala rozmawiającym dostosować się do siebie podczas dialogu na tyle, że proces komunikowania jest realizowany z dużym stopniem dokładności.

Pytanie informacyjne podczas dialogu z systemem można precyzować dwoma sposobami:

- a. wzbogacając je, podając jego warianty, zwiększając liczbę relekcji kojarzeniowych, rozszerzając strefę poszukiwań,
- b. uściślając je, wprowadzając ograniczenia w relacjach kojarzeniowych, ograniczając strefę poszukiwań.

To wszystko użytkownik może robić sam w zależności od swego rozeznania, wiedzy, doświadczenia, znajomości systemu itd. Może pomagać mu w tym sam system, ale do tego powinny w nim być ódzwierciedlone najbardziej prawdopodobne sposoby (wzory) takiego dialogu i opracowane środki jego realizacji.

#### TYPY ROZBIEŻNOŚCI MIĘDZY JEONOSTKAMI LEKSYKALNYMI PYTANIA I JEONOSTKAMI LEKSYKALNYMI CHARAKTERYSTYKI WYSZUKIWAWCZEJ. TYPY KOMPENSACYJNYCH STRUKTUR KOJARZENIOWYCH

Rozpatrzmy teraz typy sytuacji, które wywołują różne rozbieżności między tekstem pytania a tekstem charakterystyki wyszukiwawczej i spróbujmy wskazać i zinterpretować różne kompensacyjne struktury wypracowane w języku informacyjno-wyszukiwawczym.

Jak wiadomo tekst pytania i tekst charakterystyki wyszukiwawczej może być całkowicie zbieżny, zbieżny częściowo lub całkowicie rozbieżny.

1. Zbieżność całkowita oznacza, że użytkownik zwracając się do systemu z pytaniem nie ma żadnych problemów od strony formalnej z otrzymaniem dokumentów odpowiadających pytaniu. Nie są potrzebne

żadne pytania uzupełniające. Oczywiście jest to najprostszy, "najszczęśliwszy" przypadek, sprawiający, że proces komunikowania z systemem jest nadzwyczaj łatwy, a w rezultacie - najbardziej efektywny. Sytuacje takie zwykle zdarzają się, gdy w pytaniu użyte są albo wyrażenia bardzo ogólne, albo bardzo konkretne, które dokładnie odzwierciedlają potrzebę informacyjną użytkownika i przychodzą mu łatwo na myśl podczas formułowania pytania, a równocześnie występują w systemie. Przy tym cechą charakterystyczną takich wyrazów jest ich ostry zakres, wyraźnie oddzielający je od innych. Są to wyrażenia, które w normalnych warunkach praktycznie nie mogą być pomyłone z innymi wyrażeniami (np.: Gruzlicą, Traktory, Etyka itp.).

2. Praktycznie pełna zbieżność jednostek leksykalnych pytania i jednostek leksykalnych charakterystyki wyszukiwawczej pod względem znaczenia, choć przy odmiennych formach, jest typową synonimią. W tym przypadku jednostki leksykalne z zapytania i jednostki leksykalne z charakterystyki wyszukiwawczej pozostają wobec siebie w takiej zależności jak paradygmatyczna para kojarzeniowa połączona relacją równoznaczności.

Jeżeli tekst pytania i tekst charakterystyki różnią się, to w procesie wyszukiwania potrzebna jest korekta pytania, którą można stosunkowo łatwo przeprowadzić. Po pierwsze dlatego, że użytkownik może prowadzić wyszukiwanie według znanych mu synonimów. Po drugie takie standardowe połączenia jednostek leksykalnych mogą być odzwierciedlone w samym języku informacyjnym, gdzie jednostki leksykalne połączone relacją równoznaczności w systemie (katalogu) występują jako odsyłacze całkowite (oznaczone wyrażeniem zobacz). Stanowią one rodzaj odpowiedzi użytkownikowi i ułatwiają mu zadanie, ustalając formę obowiązującą w systemie, wybraną spośród innych form (synonimów). Jest to zwykle wyrażenie najczęściej spoty-

kane, uznane za priorytetowe wśród synonimów, jakby najbliższe użytkownikowi, najbardziej prawdopodobne do użycia na oznaczenie danej sytuacji.

3. Częściowa zbieżność zakresów jednostek leksykalnych pytania i jednostek leksykalnych charakterystyki wyszukiwawczej może mieć taki charakter, jak paradygmatyczna para kojarzeniowa połączona relacją "gatunek - rodzaj". W takich sytuacjach w procesie korekty użytkownik może sam zmienić swoje pytanie, wykorzystując termin rodzajowy zamiast terminu oznaczającego gatunek, może też użyć obu terminów. Taki związek jednostek leksykalnych jest wprowadzony do systemu w postaci odsyłaczy uzupełniających (zobacz też /5, s.93/) lub odsyłaczy bardziej wyspecyfikowanych typu: zobacz pod:, zobacz od: [w polskiej praktyce te sposoby odsyłania są stosowane tylko w niektórych katalogach - przyp. tłum.]. Są to odsyłacze podpowiadające użytkownikowi, również ułatwiające mu posługiwanie się językiem systemu, wzbogacające strefę, w której się obraca. Jedną z najbardziej typowych sytuacji jest tutaj poszukiwanie imion własnych [ogólniej wszystkich nazw jednostkowych - przyp. tłum.] typu nazwisk, marek samochodów, wydarzeń historycznych itp.

4. Częściowa zbieżność zakresów jednostek leksykalnych pytania i charakterystyki wyszukiwawczej może też mieć taki charakter jak para kojarzeniowa połączona relacją rodzaj - gatunek. W takich sytuacjach liczba możliwych związków kojarzeniowych w systemie wyraźnie wzrasta. W procesie korekty pytania użytkownik może sam je zmienić, sprecyzować wykorzystując terminy gatunkowe zamiast terminu rodzajowego lub wykorzystać wszystkie. Taki związek jednostek leksykalnych oznaczających rodzaj i gatunek jest odzwierciedlony w systemie w postaci odsyłaczy uzupełniających hierarchi-

cznych (zobacz też). Przy czym postać takiego odsyłacza może być różna: dla każdej pary rodzaj - gatunek samodzielny odsyłacz, dla jednego terminu rodzajowego - kilka terminów oznaczających gatunki, dla jednegp terminu rodzajowego - wszystkie terminy oznaczające gatunki [odsyłacz uzupełniający wyliczający - przyp. tłum], dla jednego terminu ródzajowego - podanie typu odpowiadających terminów gatunkowych [w polskiej terminologii - odsyłacze uzupełniające orientacyjne - przyp. tłum.] /5, s.90-92; 11, s. 91-93/.

5. O wiele bardziej złożone są przypadki, gdy rozbieżności między jednostkami leksykalnymi pytania i charakterystyki wyszukiwawczej wynikają stąd, że jednostki te nazywają obiekty różnych, chociaż związanych ze sobą treściowo, tematycznie, klas, np.: część - całość. W takich przypadkach w relacjach zachodzących między tymi jednostkami leksykalnymi przejawiają się relacje między bardziej szczegółowymi słowami kluczowymi, które dopiero w sumie charakteryzują dany przedmiot w całości. Odnosi się to także do relacji typu rodzaj - gatunek, które zachodzą nie między jednostkami leksykalnymi pytania i charakterystyki wyszukiwawczej, lecz między jednostkami leksykalnymi pytania i zestawem słów kluczowych, który jest swego rodzaju makrostrukturą, obejmującą obydwie te jednostki leksykalne. Co się tyczy stosunku jednostek leksykalnych pytania i charakterystyki wyszukiwawczej, to mają się one tak do siebie, jak pary tematyczne w opisanym wcześniej eksperymencie kojarzeniowym (III).

W takich przypadkach, gdy w pytaniu jedna jednostka leksykalna w rzeczywistości powinna być rozszerzona do zestawu słów kluczowych, możliwości precyzowania, uszczegółowiania przy charakteryzowaniu poszukiwanego przedmiotu są wystarczająco szerokie i wielorakie. Najbardziej standardowe przypadki, opisujące różne sytuacje

za pośrednictwem zestawu słów kluczowych mogą tworzyć tak zwane szerokie kompleksy haseł przedmiotowych zorganizowanych za pomocą odpowiedniej sieci odsyłaczy uzupełniających. Uogólnienie lub uszczegółowienie sytuacji (typów sytuacji) jest znacznie mniej znormalizowane, standardowe niż opisane w przypadkach 2,3,4 i o wiele bardziej różnorodne zarówno jeśli chodzi o kierunki, jak i stopień uszczegółowienia.

Rozpatrzone zjawiska jeszcze raz pokazują, że wszystkie związki zachodzące między jednostkami leksykalnymi wprowadzone do systemu informacyjnego, zarówno w szerokich kompleksach haseł przedmiotowych (za pośrednictwem sieci odsyłaczy), jak i w tak zwanych wąskich kompleksach (za pośrednictwem haseł przedmiotowych rozwiniętych) są związkami kojarzeniowymi o charakterze paradygmatycznym, syntagmatycznym lub tematycznym. Wszystkie one wyrażają związki między tekstami lub między elementami tekstów, dając w przybliżeniu jednakową trafność [cel'nost'] i dlatego wchodzą w pole kojarzeniowe słów, mających tak czy inaczej związki z ową trafnością.

Tak więc, jeśli mówić nie o formalnym stosunku odpowiedności jednostek leksykalnych pytania i charakterystyki wyszukiwawczej (relewantność), a o stosunkach treści, kryjących się za tymi jednostkami (pertynencja), to problem powinien dotyczyć nie zwykłej odpowiedności jednostek leksykalnych, ale odpowiedności tekstów lub ich fragmentów, odzwierciedlających tę trafność i dlatego odnoszących się do tego samego pola kojarzeniowego.

Właśnie dlatego w systemach informacyjnych wybiera się zasób leksykalnych jednostek wzajemnie powiązanych, należących do strefy centralnej danego pola kojarzeniowego. Przy tym struktura języka informacyjno-wyszukiwawczego nakierowana jest przede wszystkim na

najbardziej standardową, centralną, o najwyższym wskaźniku częstotliwości część pola kojarzeniowego, co powoduje zbieżność pól kojarzeniowych użytkownika i zbiorów dokumentów pogrupowanych według kompleksów haseł przedmiotowych najbardziej prawdopodobnych dla najczęstszych i najbardziej typowych sytuacji wyszukiwawczych.

Wykorzystanie wszystkich środków formalnych ułatwiających wykrycie związków kojarzeniowych danych stymulatorów pozwala wykryć i podpowiedzieć użytkownikowi cały zestaw jednostek leksykalnych, tworzących w danym systemie informacyjnym pole kojarzeniowe danej jednostki leksykalnej, która może być mniej lub bardziej zgodna z rzeczywistymi potrzebami użytkownika (problem kompletności). Ponieważ jednak te same jednostki leksykalne mogą wchodzić w różne pola kojarzeniowe, możliwy jest szum informacyjny (problem dokładności wyboru).

Jak widać wszystkie zjawiska, z którymi stykamy się w procesie informacji, związane z trudnościami otrzymania właściwego efektu poszukiwań, oparte są na tym samym mechanizmie struktur kojarzeniowych, co stwarza możliwości dodatkowych manipulacji tymi strukturami z myślą o ich organizacji i funkcjonowaniu.

#### STRUKTURY KOJARZENIOWE JĘZYKA INFORMACYJNO-WYSZUKIWAWCZEGO JAKO EFEKT MODELOWANIA STRUKTUR KOJARZENIOWYCH JĘZYKA NATURALNEGO

Różnorodność wielu czynników, określających złożoność typów struktur kojarzeniowych w języku informacyjno-wyszukiwawczym i wyobrażeń o budowie tego języka wyjaśnia, dlaczego jest tak niejednakowe wykorzystanie tych struktur w różnych wersjach systemów informacyjnych opartych na języku haseł przedmiotowych: od prawie ich braku w prostych indeksach przedmiotowych (tworzonych z samych tematów uszeregowanych alfabetycznie) do skomplikowanego systemu

takich struktur w rozbudowanych dużych katalogach przedmiotowych. Przy tym wykorzystanie tych struktur nie zawsze jest jednakowe. Regułą jest, że tam, gdzie katalog przedmiotowy funkcjonuje obok systematycznego (np. w Bibliotece im. Sałtykowa-Szczedrina w Leningradzie) są one odzwierciedlone w sposób o wiele bardziej fragmentaryczny niż tam, gdzie katalog przedmiotowy jest jedynym katalogiem rzeczowym (np. w Bibliotece INION w Moskwie) /12/. Różnie też może być ukazana w różnych katalogach przedmiotowych cała gama tematycznych struktur kojarzeniowych. Zwykle wykorzystuje się 3-4 typy takich struktur /3, s.69/.

Wprowadzenie do systemu informacyjnego jednego typu struktur kojarzeniowych pociąga za sobą rozwój struktur innego typu. Tak na przykład sprawa inwersji, która jest szeroko wykorzystywana w katalogach przedmiotowych [polskich nie - przyp. tłum]. Inwersja umożliwia tworzenie tzw. wąskich kompleksów haseł przedmiotowych, ale jednocześnie zwiększa liczbę synonimów, co powoduje rozbudowę odsyłaczy całkowitych [relacja równoznaczności między jednostką leksykalną postą i zinwertowaną - przyp. tłum.], wpływając tym samym na złożoność struktury języka haseł przedmiotowych.

W tym kontekście także nieco inaczej ocenia się rolę aparatu odsyłaczowego w języku haseł przedmiotowych, gdyż funkcją odsyłaczy jest nie tylko po prostu porządkowanie związków między jednostkami leksykalnymi (hasłami przedmiotowymi) w strukturze języka czy w zbiorze informacyjnym (np. w katalogu przedmiotowym). Aczkolwiek z punktu widzenia idealnych wymagań wobec aparatu odsyłaczowego zawsze możemy zarzucić mu niekompletność odzwierciedlenia relacji, to jednak w rozpatrywanym planie aparat odsyłaczowy spełnia w rzeczywistości funkcję "rekomendującą", podpowiadając użytkownikowi dodatkowe (fakultatywne) środki optymalizacji procesu wyszu-



kiwania w przypadku wystąpienia różnorodnych werbalnych środków indeksowania.

To, co powiedziano wyżej pozwala zrozumieć, dlaczego aparat odsyłaczowy jest systemem złożonym i niejednorodnym pod względem wyrażania stosunków semantycznych i przypisywanych mu zadań komunikatywnych. Każdy z fragmentów tego systemu ma swoją specyfikę i jest do pewnego stopnia autonomiczny. Dlatego można go wprowadzić do systemu (lub pominąć go) niezależnie od innych fragmentów danego systemu odsyłaczowego. Stąd też różnorodność aparatu odsyłaczowego w zależności od tego, jakie relacje kojarzeniowe są w nim odzwierciedlone, jakie związki jednostek leksykalnych ma "podpowiadać" użytkownikowi, a tym samym wpływać na elastyczność strategii wyszukiwawczej.

Pole kojarzeniowe zawarte w języku informacyjno-wyszukiwawczym ułatwia użytkownikowi sformułowanie całego fragmentu centralnej części pola kojarzeniowego danej jednostki leksykalnej. Sformowanie (i sformułowanie) tego centrum pola kojarzeniowego rozszerza możliwości sprecyzowania nazwy poszukiwanego przedmiotu, zasygnalizowanego w pytaniu. Przy tym te same rodzaje relacji kojarzeniowych (przede wszystkim tematycznych) mogą być wprowadzane do języka informacyjno-wyszukiwawczego jako środki paradygmatyczne (aparat odsyłaczowy) lub jako środki syntagmatyczne (hasło przedmiotowe rozwinięte). Czasem też przyjmuje się umowne, konwencjonalne rozwiązania dla niektórych zjawisk, które mają charakter kojarzeniowy. W tym sensie typizacja w języku haseł przedmiotowych (typy określników, typy relacji w aparacie odsyłaczowym) jest tylko odbiciem dążenia do ujednoczenia wyobrażeń o relacjach kojarzeniowych.

Jakkolwiek jednak byłby rozbudowany aparat odsyłaczowy we współczesnych systemach informacyjnych opartych na języku haseł

przedmiotowych ogólnie rzecz biorąc ma on ograniczone możliwości. Wiąże się to z elastycznością sformułowań użytkownika (odwołującego się do języka naturalnego) i schematycznością języka sztucznego, operującego ograniczonym zasobem jednostek leksykalnych i powiązań między nimi odzwierciedlonych w budowie haseł przedmiotowych i odsyłaczach, które proponuje się użytkownikowi.

Jest oczywiste, że im bardziej znormalizowane są związki danej jednostki leksykalnej w polu kojarzeniowym języka naturalnego, tym bardziej prawdopodobne, że te związki trafią do języka informacyjno-wyszukiwawczego. Przy tym im bliżej centrum pola kojarzeniowego, tym wyraźniejsze są związki paradygmatyczne i syntagmatyczne.

Wykrycie i wprowadzenie najbardziej standardowych związków z pola kojarzeniowego przy opracowywaniu języka informacyjno-wyszukiwawczego praktycznie może opierać się na:

- a. dobrej orientacji pracownika w priorytetach systemu nauk (włączając w to orientację w klasyfikacjach biblioteczno-bibliograficznych) i na odzworowywaniu w języku informacyjno-wyszukiwawczym relacji zachodzących w tym systemie,
- b. dobrej orientacji pracownika w danych empirycznych, to jest obserwacja użytkownika (zwykle jest to możliwe w małych bibliotekach, gdzie ta sama osoba opracowuje dokumenty, prowadzi katalog i udziela informacji),
- c. intuicji pracownika, wykorzystującego swoje osobiste doświadczenia pracownicze, które w sposób mniej lub bardziej trafny odtwarzają jego rozeznanie w systemie nauk i pytaniach użytkownika.

Ostatecznie jednak formułowanie struktury kojarzeniowej w języku informacyjno-wyszukiwawczym przebiega pod wpływem pól kojarzeniowych języka naturalnego. Realizuje się to zarówno podczas pracy indeksatora, jak i podczas pracy użytkownika.

Rozwój pola kojarzeniowego w języku informacyjno-wyszukiwawczym (j iw) to jest zwiększenie jego "mocy", zwielokrotnienie liczby komponentów i relacji nie ma zasadniczo granic o charakterze strukturalnym, uwarunkowanych samym charakterem skojarzeń. Wszystkie występujące ograniczenia relacji w polach kojarzeniowych j iw, to ograniczenia paradygmatyczne wynikające z realnych możliwości (zwykle ekonomicznych) prowadzenia systemu informacyjnego. Przy zmianie tych możliwości możliwe są też zmiany "mocy" pól kojarzeniowych języka.

#### O PERSPEKTYWACH ROZWOJU STRUKTUR KOJARZENIOWYCH W SYSTEMACH INFORMACYJNYCH

To, co powiedziano o strukturach kojarzeniowych, odnosi się do systemów tradycyjnych.

Jeśli chodzi o systemy zautomatyzowane, to na razie wykorzystuje się w nich te same tradycyjne metody, aczkolwiek nieco zmodyfikowane, co wiąże się z nowymi możliwościami technicznymi. Jednakże już dziś wiadomo, że istnieją możliwości zupełnie nowych rozwiązań. I tu rola struktur kojarzeniowych języka naturalnego w procesie wyszukiwania wyraźnie wzrasta. A to znaczy, że mechanizm wykrycia ich i wykorzystania w języku informacyjno-wyszukiwawczym wymaga jakościowo nowego podejścia.

Możliwości optymalizacji sprowadzają się po pierwsze do rozszerzenia peryferii pola kojarzeniowego w systemie informacyjnym i po drugie w zwiększeniu elastyczności wykorzystania składników pola kojarzeniowego i ich związków w organizacji dialogu użytkownika z systemem informacyjnym. I jedno i drugie wiąże się z "naturalizacją" systemu informacyjnego i może być osiągnięte przy szerokim wprowadzeniu techniki komputerowej.

Dziś brak jest jeszcze dokładnych danych dotyczących pól kojarzeniowych w języku naturalnym, nawet dla terminów specjalnych, najbardziej wykorzystywanych w procesie wyszukiwania. Dla sprawnego działania systemów informacyjno-wyszukiwawczych potrzebne jest zbudowanie dynamicznego (nie statycznego jak obecnie) systemu pól kojarzeniowych. Czynione są w tym zakresie pierwsze kroki, zarówno jeśli chodzi o typowanie elementów struktur kojarzeniowych, jak i strukturalizacji dynamicznego pola kojarzeniowego na podstawie metod statystycznych /1, s.37-38, 44/.

Wprowadzenie do pamięci komputera zbioru rozwiniętych tekstów, opracowanie programów rozpoznawania i analizowania relacji kojarzeniowych jednostek leksykalnych w tych tekstach pozwoli zarówno zobiektywizować jak i sformalizować, a tym samym ułatwić wykrywanie standardowych stref pól kojarzeniowych, jak też bardziej uelastyczyć, w sposób zindywidualizowany wydzielać konieczne fragmenty pól kojarzeniowych metodą dialogu - od razu w procesie zautomatyzowanego wyszukiwania informacji.

2.02.1989 r.

Tłumaczenie i adaptacja Jadwiga Sadowska

## Literatura

1. AVETIŞJAN D.D. Statističeskie metody pri rešenii prikladnyh zadatč dokumental'nogo poiska. Avtoref. dis.... dra techn. nauk. Novosibirsk 1975 48 s.

2. GOST 7.25-80. Tezaurus informacionno-poiskovyj odnojazyčnyj. Pravila razrabotki, struktura, sostav i forma predstavlenija, Vved. 01.01.82. 16 s.
3. GRININA R.F. Predmetnyj katalog kak informacionno-poiskovaja sistema. Leningrad 1979 81 s.
4. KLIMENKO A.P. Leksičeskaja sistemnost' i ee psiholingvističeskoe izučenje. Minsk 1974 108 s.
5. KRUGLIKOVA V.P. Predmetizacija proizvedenij pečati. Obšč. metodika. Moskva 1967 173 s.
6. MICHAJLOV A.I., ČERNYJ A.I., GILJAREVSKIJ R.S. Osnovy informatiki. Moskva 1968 756 s.
7. DVCINNIKOVA I.G. Tekstoobrazajuščaja rol' verbal'nyh asociativnyh struktur. Avtoref. dis....kand. filol. nauk. Leningrad 1986 14 s.
8. SACHARNYJ L.V., DVCINNIKOVA I.G. Asociativnye pary kak teksty. V: Materialy VIII Vsesojuznogo simpoziuma po psiholingvistike i teorii komunikacii (Tezisy dokladov). Moskva 1985, s.126-128.
9. SLOVAR' associativnyh norm russkogo jazyka. Moskva 1977, 192 s.
10. SLOVAR' terminov po informatike na russkom i anglijskom jazykach. Moskva 1971 360 s.
11. SUVOROVA V.P. O nekotoryh voprosach ssyločno-spravočnogo apparata predmetnogo kataloga (k itogam razrabotki GOSTov po predmetizacii) V: Predmetnyj poisk v tradicionnyh i netracionnyh informacionno-poiskovyh sistemach, Vyp. 7. Leningrad 1987 s. 82-96
12. Televidenije (spisok predmetnyh rubrik). Materialy k tezasuru po obščestvennym naukam. Moskva, 1975 152 s.

THE ASSOCIATIVE STRUCTURES IN NATURAL AND INFORMATION RETRIEVAL  
LANGUAGE (On the Example of the Subject Headings Language)

Summary

There are presented in the article the associative relationships occurring in information retrieval language (IRL). They are characterized from the point of view of the universal associative mechanisms specific for natural languages.

The associative field of natural language is analyzed. Three groups of associations are assigned: the paradigmatic, syntagmatic and the subject ones. The performed associative experiments are described. The role of the associative structures in the construction of the reference apparatus as well as in the dialog of user with information retrieval system is discussed. The associative nature of the all relationships between lexical units of the IRL is stressed.

АССОЦИАТИВНЫЕ СТРУКТУРЫ В ЕСТЕСТВЕННОМ ЯЗЫКЕ И В ПРЕДМЕ-  
ТИЗАЦИОННОМ ИЛЯ

Резюме

В статье делается попытка обосновать представление всех типов смысловых связей в предметизационном илЯ как результата моделирования универсального механизма ассоциативных связей, проиизывающего смысловые структуры в речевой деятельности на базе еЯ. Рассматривается ассоциативное поле еЯ и приводятся результаты трёх экспериментов. На основе экспериментов выделены три типа отношений: парадигматические, синтагматические, тематические,

которым в ипЯ соответствуют отношения парадигматические логические, синтагматические и парадигматические ассоциативные: все они по характеру являются ассоциативными.

Представлен диалог абонента и ипс как взаимодействие ассоциативных структур. Анализируются типы несовпадений ле запроса и ле пода и типы компенсационных ассоциативных структур в предметизационных ипЯ.

ELŻBIETA ARTOWICZ

Wydział Neofilologii UW

BOLESŁAW SZOMAŃSKI

Instytut Organizacji Systemów Produkcyjnych PW

**SYSTEM INFORMACJI O METODACH I TECHNIKACH ORGANIZATORSKICH  
W STEROWANIU JAKOŚCIĄ (SIMTOJ)**

Zakres prac nad budową systemu. Badania potrzeb informacyjnych użytkowników systemu. Uczestnicy systemu na etapie projektowania i wstępnej eksploatacji. Zasoby informacyjne. Skomputeryzowana baza danych. Struktura funkcjonalna i organizacyjna systemu. Problemy projektowania języka informacyjno-wyszukiwawczego. Zastosowane środki techniczne. Współpraca z użytkownikami.

Instytut Organizacji Systemów Produkcyjnych Politechniki Warszawskiej realizuje temat badawczy CBBP 02.20.PI. "Wybrane zagadnienia jakości wyrobów w przedsiębiorstwach i zapleczu naukowo-badawczym przemysłu elektromaszynowego", wchodzący w skład centralnego programu badawczego "Wybrane zagadnienia podstawowych problemów współczesnej metrologii oraz technologii i konstrukcji systemów i urządzeń pomiarowo-kontrolnych". Zespół specjalistów, reprezentujących wyższe uczelnie, zaplecze naukowo-badawcze i liczne zakłady przemysłowe zajmuje się przede wszystkim techniczno-organizacyjnymi zagadnieniami sterowania jakością. Sondażowe badania w przedsiębiorstwach /12/ wykazały nieznaczny zakres znajomości metod i technik organizatorskich, a także niewielką liczbę

---

"Zagadnienia Informacji Naukowej" 1989 nr 1(54)



praktycznych wdrożeń. Przyczyną niedostatecznej znajomości wiedzy teoretycznej w praktyce jest rozproszenie informacji zawartych w publikacjach o różnym stopniu dostępności, objętości i szczegółowości opisu. Pracownicy przedsiębiorstw bezpośrednio zainteresowanych poprawą jakości wyrobów metodami organizacyjnymi praktycznie nie dysponują zorganizowanymi serwisami informacyjnymi, udostępniającymi syntetyczną informację o metodach i technikach organizatorskich, dostosowaną do potrzeb poszczególnych użytkowników z przemysłu i zaplecza naukowo-badawczego.

Stan ten uznano za podstawę do rozpoczęcia prac nad budową Systemu Informacyjnego o Metodach i Technikach Organizatorskich w Sterowaniu Jakością (SIMTOJ). Projektowanie i eksperymentalne uruchamianie systemu przy współpracy jego przyszłych użytkowników stanowi część wymienionego tematu badawczego, obejmującego ponadto zagadnienia teoretyczne jakości (kwalitologii), teorii organizacji i zarządzania, ergonomii, normalizacji i innych dyscyplin szczegółowych.

#### ZAKRES PRAC

Stosowaną przez wiele lat praktyką projektowania systemów inżynierskich było przyjęcie założenia, że każdy system jest wbudowany w hierarchiczną strukturę SINTO, ograniczającego się przede wszystkim do informacji dokumentacyjnej. Założenie to odzwierciedlają także opublikowane materiały i zalecenia metodyczne w zakresie języków informacyjno-wyszukiwawczych. Problemy potrzeb użytkowników, wzmiankowane w zasadzie tak w pracach metodycznych, jak i w dokumentacjach projektowych poszczególnych systemów, nie doczekały się w rzeczywistości kompleksowego opracowania.

Zespół autorski projektu SIMTOJ stał przed zadaniem złożonym, miał bowiem przedstawić koncepcję systemu, jego projekt techniczny oraz doprowadzić do szybkiego eksperymentalnego wdrożenia części systemu, aby przekonać wszystkich aktualnych i potencjalnych uczestników systemu o celowości prowadzenia dalszych prac nad niekonwencjonalnym, w świetle dotychczasowej praktyki, rozwiązaniem. Zadania tego zespołu sformułowano następująco:

1. Określenie bieżących i perspektywicznych celów systemu wraz z określeniem kolejności ich realizacji,
2. Zorganizowanie współpracy z użytkownikami informacji, będącymi równocześnie jej twórcami, zasilającymi zbiory systemu.
3. Zaprojektowanie struktury funkcjonalnej systemu oraz struktury organizacyjnej.
4. Opracowanie koncepcji języka informacyjno-wyszukiwawczego i metodycznych założeń do opracowywania informacji wejściowej.
5. Opracowanie formatów opisu metod i technik organizatorskich (MiTO) na wejściu i wyjściu systemu, przy wykorzystaniu pakietu Mikro COS ISIS.

Specyfika procesu projektowania polega na tym, iż większą część składu zespołu autorskiego stanowią użytkownicy i twórcy informacji źródłowej, określający bezpośrednio swoje potrzeby informacyjne, odwzorowywane następnie w odpowiednich częściach dokumentacji projektowej.

#### POTRZEBY INFORMACYJNE UŻYTKOWNIKÓW SIMTOJ

Sondażowe badania wykorzystania wiedzy teoretycznej, na temat MiTO przeprowadzone w wybranych przedsiębiorstwach przemysłu elektromaszynowego wykazały - jak wspomniano - stosunkowo słabą znajomość literatury przedmiotu, wynikającą z rozproszenia i nie-

dostępności źródeł oraz z ograniczeń czasowych specjalistów zatrudnionych w służbach jakości. Stąd też konieczne było wstępne rozpoznanie potrzeb, stanowiące punkt wyjścia do określenia informacji wyjściowej systemu, którego zadaniem jest udostępnianie informacji o MiTO dotyczących jakości. Pierwszym etapem rozpoznawania potrzeb użytkowników, nie znających w praktyce regularnych serwisów polegał na zebraniu zapytań, jakie chcieliby kierować do systemu. Zgromadzony zestaw ponad 100 zapytań wykazał znaczne zróżnicowanie potrzeb użytkowników, odzwierciedlające ich sferę działania (badania, dydaktyka, produkcja) oraz różne wyobrażenia o możliwościach systemów informacyjno-wyszukiwawczych, zwłaszcza jeżeli chodzi o rodzaj wydawanej informacji.

Wśród użytkowników reprezentujących placówki naukowo-badawcze przeważały zapytania o zestawienia bibliograficzne publikacji źródłowych o MiTO. Część zapytań wyrażała potrzeby, których zaspokojenie wymagałoby dysponowania systemem inteligentnym (ekspertowym), potrafiącym podać algorytm postępowania przy rozwiązywaniu określonego problemu heurystycznego w starowaniu jakością oraz wykazać optymalne z tego punktu widzenia metody. Wyrażano także zainteresowanie opisem faktograficznym MiTO dla celów dydaktycznych z rozbudowaną informacją na temat efektów stosowania w konkretnych warunkach zakładu przemysłowego.

Użytkownicy reprezentujący sferę produkcji przemysłowej lub jej zaplecze badawcze interesowali się bądź konkretnymi metodami, o których mieli pewną ogólną orientację, bądź wyrażali zapotrzebowanie na system inteligentny z rozwiniętym modułem konwersacyjnym typu "jadłospisowego" (menu driven), który pomagałby wyszukać informację o metodach według różnorodnych problemów jakości, obszarów produkcji oraz zróżnicowanych szczebli zarządzania procesem

produkcyjnym. Zakres określanych problemów jakości wyrobów również był silnie zróżnicowany ze względu na rodzaj produkowanych wyrobów i ich przeznaczenie.

#### UCZESTNICY SIMTOJ

Uczestnikami systemu na etapie projektowania i eksploatacji, poza zespołem autorskim i obsługą techniczną, są:

- pracownicy jednostek sterowania jakością w zakładach przemysłowych, zainteresowani jedynie wyszukiwaniem informacji w systemie,

- pracownicy jednostek sterowania jakością w zakładach przemysłowych, korzystający ze zbiorów informacyjnych systemu i zasiłający je informacjami o wynikach wdrożenia poszczególnych MiTO, a w szczególności cenną informacją niepublikowaną.

- autorzy nowych oryginalnych MiTO, twórcy informacji źródłowej,

- pracownicy naukowo-dydaktyczni jednostek badawczych Politechniki Warszawskiej i innych placówek naukowych,

- autorzy opisów MiTO dla potrzeb systemu,

- studenci Instytutu Organizacji Systemów Produkcyjnych PW.

Koordinacją współpracy z autorami informacji źródłowej i (lub) syntetycznej zajmuje się operator systemu i podległa mu służba informacyjna, a w szczególności służba odpowiedzialna za gromadzenie informacji na temat doświadczeń stosowania poszczególnych MiTO podczas sterowania jakością w przedsiębiorstwach. Operator systemu ma współpracować ponadto z zespołem autorskim w zakresie aktualizacji języka informacyjno-wyszukiwawczego i materiałów metodycznych.

## ZBIORY INFORMACYJNE SYSTEMU

W trakcie realizacji tematu "Wybrane zagadnienia jakości wyrobów w przedsiębiorstwach przemysłowych i zapleczu naukowo-badawczym przemysłu elektromaszynowego" w latach 1987 - 1988 /12/ przygotowano bibliografię dostępną w kraju literatury na temat MiTO, opublikowanej w latach 1960 - 1987. Zawiera ona ponad 1300 pozycji reprezentujących różny poziom szczegółowości opisu metod i technik organizatorskich, których nazwy zawiera lista słów kluczowych (ponad 2000), obejmująca także nazwy synonimiczne. Ponadto opracowano wykaz 230 norm dotyczących zagadnień jakości wyrobów w omawianej dziedzinie, obowiązujących w Polsce i innych krajach świata oraz zalecanych przez RWPG, ISO i EOQC<sup>1/</sup>.

Informację źródłową systemu stanowią:

- monografie z zakresu metodologii organizacji o profilu teoretycznym, podejmujące problemy definicji podstawowych terminów, a także prezentujące różnorodne typologie (klasyfikacje) metod,
- monografie poświęcone charakterystyce poszczególnych klas MiTO, zawierające ponadto opisy z przykładami zastosowań /2/.
- monografie poświęcone opisowi pojedynczych metod i ich aspektom wdrożeniowym /2/,
- opracowania na temat MiTO przy omawianiu problemów innowacji, racjonalizacji i projektowania /3/,
- praca popularyzatorskie, materiały dydaktyczne /5, 7/,
- monografie o profilu metodologicznym w zakresie teorii organizacji i zarządzania /8/,
- wydawnictwa encyklopedyczne,
- wydawnictwa stowarzyszeń wyższej użyteczności - na przykład oddziałów oraz zarządów Towarzystwa Naukowego Organizacji i Kierownictwa, instytutów i ośrodków badawczych (Instytutu Administracji

<sup>1/</sup>European Organisation for Quality Control

i Zarządzenia, Instytutu Pracy, Instytutu Filozofii i Socjologii PAN) oraz uczelni wyższych,

- artykuły z czasopism naukowych poświęcone zagadnieniom teoretycznym lub aplikacjom poszczególnych MiTO,
- oryginalne, niepublikowane opisy MiTO,
- dokumentacja służb sterowania jakością w przedsiębiorstwach przemysłowych dotycząca stosowanie MiTO oraz posiadanych doświadczeń.

Z punktu widzenia sposobu reprezentacji na informacją źródłową składa się informacja tekstowa oraz informacja graficzna w postaci diagramów, schematów, wykresów ilustrujących algorytmy postępowania przy zastosowaniu poszczególnych MiTO, a także programy komputerowe.

#### SKOMPUTERYZOWANA BAZA DANYCH SIMTOJ

Zgodnie z założeniami systemu obiektami, o których gromadzona jest informacja są metody i techniki organizatorskie. Analiza terminologiczna prac monograficznych w omawianym zakresie, porównanie różnorodnych definicji "metody" i "techniki" doprowadziła do wniosku, że wobec niekonsekwencji autorów i braku consensusu w sprawie cech dystynktywnych, różniących znaczenie terminu "metoda" i "technika" przyjmuje się ich łączne traktowanie jako nazwy obiektu systemu - MiTO (Metody i Techniki Organizatorskie). Systemowa nazwa danej MiTO jest podstawą organizacji zbioru informacji źródłowej oraz skomputeryzowanej bazy danych, w której przechowywane są syntetyczne informacje faktograficzne oraz dane o dokumentach źródłowych, zawierających opis szczegółowy, a także o jednostkach, które wdrożyły daną MiTO oraz o autorze opisu. Przy takim założeniu w bazie danych SIMTOJ gromadzi się informacje faktograficzno-

-dokumentacyjne, zaś charakterystyka wyszukiwawcza MiTO zawiera trzy poziomy opisy odpowiadające trzem rodzajom relewancji w systemie, zgodnie z ustalonymi potrzebami użytkowników:

- 1) opis encyklopedyczny (dokumentacyjny),
- 2) opis identyfikacyjny (faktograficzny),
- 3) opis szczegółowy (źródłowy).

Opis encyklopedyczny (dokumentacyjny) odwzorowuje następujące cechy dystynktywne MiTO:

- a) nazwę systemową MiTO,
- b) cechy bibliograficzne dokumentów zawierających szczegółowy opis MiTO oraz informacje o efektach stosowania przez konkretnych użytkowników,
- c) zwięzły opis dokumentacyjny MiTO o objętości do 1 strony tekstu,
- d) informacje o wdrożeniach MiTO.

Opis identyfikacyjny (faktograficzny), odwzorowuje pozajęzykowe cechy MiTO, to jest atrybuty obiektów bazy danych SIMTOJ uznane za relewantne z punktu widzenia problemów rozwiązywanych przez użytkowników. Pozajęzykowe atrybuty MiTO są odwzorowywane w strukturze pola semantycznego języka informacyjno-wyszukiwawczego.

Elementami opisu identyfikacyjnego są także:

- cechy terminologiczne MiTO, takie jak: nazwy synonimiczne, nazwy metod nadrzędnych i podrzędnych zakresowo, mereologicznie oraz nazwy metod skojarzonych,

- dane o autorze charakterystyki wyszukiwawczej oraz opisu encyklopedycznego, a także terminy kolejnej aktualizacji opisu.

Opis szczegółowy - jest to pełny opis metody i techniki organizatorskiej, umożliwiający jej zastosowanie w praktyce;

opis ten powinien być wykonywany dla nowych metod albo metod, co do których nie istnieje ogólnodostępna literatura w języku polskim. W pozostałych przypadkach rolę opisu szczegółowego powinny spełniać pozycje źródłowe lub odbitki części z tych pozycji.

Struktura zbiorów informacyjnych SIMTOJ przedstawiona jest na rys. 1.

#### STRUKTURA FUNKCJONALNA SYSTEMU

Docelowo system będzie składał się z sześciu podsystemów (modułów) odpowiadających poszczególnym procesom i etapom przetwarzania informacji w systemie:

Moduł 1 - informacja źródłowa: gromadzenie i ewidencjonowanie informacji wejściowej, źródłowej.

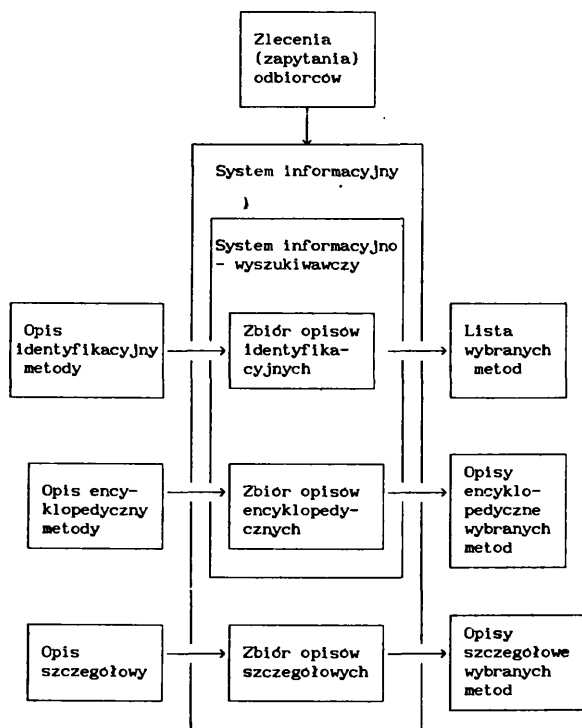
Moduł 2 - opis szczegółowy: współpraca z twórcami informacji źródłowej (autorami opisów i zapleczem naukowo-badawczym przemysłu), przygotowywanie opisów szczegółowych.

Moduł 3 - opis identyfikacyjny i encyklopedyczny: opracowanie charakterystyki wyszukiwawczej MiTO, przygotowanie opisu identyfikacyjnego na podstawie instrukcji stosowania języka informacyjno-wyszukiwawczego oraz ich wstępna weryfikacja.

Moduł 4 - baza danych SIMTOJ wprowadzanie, wyszukiwanie, aktualizacja zapisanych w bazie opisów MiTO.

Moduł 5 - współpraca z użytkownikami zewnętrznymi: obsługa informacyjna użytkowników na podstawie wyazukanych opisów MiTO oraz źródłowych opisów szczegółowych, w tym przygotowanie kopii dokumentów źródłowych, schematów i wykresów, a także udostępnianie informacji nieopublikowanej o wynikach stosowania MiTO w praktyce przemysłowej oraz zapleczu naukowo-badawczym przemysłu.





Rys 1. Schemat struktury zbiorów informacyjnych SIMTOJ

Moduł 6 - badania i rozwój systemu: analiza potrzeb i wymagań użytkowników; korekta i usprawnianie języka informacyjno-wyszukiawczego oraz usprawnianie wykorzystywanych środków informatycznych i organizacyjnych.

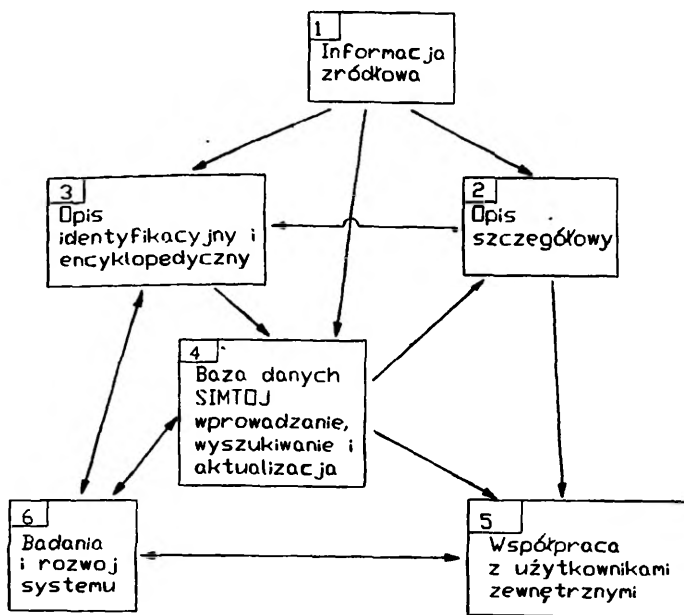
Założono, że SIMTOJ będzie systemem o strukturze gwiazdистой (patrz rys. 2), tj. w ujęciu Z. Pawlaka będzie systemem z centralną pamięcią, w której przechowywana są zbiory opisów szczegółowych MiTO oraz ich charakterystyki wyszukiawcze /9/. Aktualizacja i rozwój języka informacyjno-wyszukiawczego (i odpowiednio zbiorów informacji w skomputeryzowanej bazie danych) będzie realizowana przez operatora systemu i służby informacyjne. Równocześnie użytkownicy będą mogli korzystać ze zbiorów charakterystyk wyszukiawczych, udostępnianych na nośnikach maszynowych (dyskietkach), w zakresie pełnym lub ograniczonym, stosownie do potrzeb. Język informacyjno-wyszukiawczy będzie udostępniany użytkownikowi w zakresie niezbędnym do realizacji procesów wyszukiwania.

#### JĘZYK INFORMACYJNO-WYSZUKIWAWCZY

Projektowanie języka informacyjno-wyszukiawczego SIMTOJ poprzedzono analizą literatury z zakresu systemów informacji faktograficznej oraz literatury przedmiotu na temat metod i technik organizatorskich.

Zadaniem systemu jest dostarczanie użytkownikowi w odpowiedzi na zapytanie dwu rodzajów informacji:

- pozajęzykowej wiedzy faktograficznej o cechach obiektów opisywanych w systemie,
- metainformacji o cechach dokumentów, w których zawarta jest informacja o obiektach (informacji dokumentacyjnej).



Rys. 2. Schemat funkcjonalny SIMTOJ

Praktyka projektowania języków informacyjno-wyszukiwawczych (JIW) i dostępne zalecenia metodyczne opiera się na rozłącznym traktowaniu języków stosowanych w systemach faktograficznych i dokumentacyjnych /4/. Nie wydaje się to słuszne, ponieważ o przydatności danego JIW w systemie decyduje nie przyjęte a priori jego przeznaczenie, lecz odwzorowana w nim wiedza o cechach obiektów językowych i/lub pozajęzykowych i o ich związkach wyznaczających w sumie strukturę pola semantycznego. Potwierdza to zresztą L. Bielicka w pracy poświęconej JIW prostych systemów faktograficznych typu "obiekt - cecha", charakteryzując słowniki tych języków jako przypominające strukturę tezaury /4/.

Formułowanie koncepcji JIW w systemie wiązało się z następującymi wymaganiami i ograniczeniami.

1. Potrzeba udostępniania użytkownikom zbiorów skomputeryzowanej bazy danych do lokalnego wykorzystania, co wymagało rozbudowy języka konwersacyjnego (wyszukiwawczego), który umożliwiałby wyszukiwanie informacji w trybie "jadłospisowym" (menu driven), istotnym z punktu widzenia użytkowników nie mających praktyki posługiwania się skomputeryzowanymi serwisami informacyjnymi. Ponadto zakłada się możliwość wyszukiwania charakterystyk wyszukiwawczych MiTO według cech (lub ich kombinacji) z góry zadanych, z pominięciem przewidzianego schematu dialogu, na przykład według nazwy systemowej MiTO, celu wyszukania pełnego opisu, według cechy obszaru zastosowania w procesach produkcyjnych itp.

2. Duże zróżnicowanie zainteresowań użytkowników z punktu widzenia specjalizacji zawodowej oraz aplikacyjne lub teoretyczne przeznaczenie wyszukanej informacji. Wynika stąd konieczność zapewnienia zindywidualizowanego sposobu reprezentacji wiedzy w systemie, to jest ustalenia cech relewantnych MiTO odpowiadających różnym profilom zainteresowań.

W tym celu przeprowadzono wspomnianą sondszą analizę przewidywanych zapytań, formułowanych przede wszystkim przez przedstawicieli przemysłu. Na tej podstawie podjęto opracowanie formatów opisów MiTO na wejściu i wyjściu systemu, przy uwzględnieniu jądrowego lub indywidualnego trybu wyszukiwania informacji.

Z ustaleniem repertuaru cech identyfikacji MiTO w opisie faktycznym wiązała się konieczność zdefiniowania "metody" i "techniki" ze względu na możliwość ich rozróżnienia na poziomie opisu systemowego. Analiza literatury przedmiotu w zakresie teorii organizacji i zarządzania, prakseologii i innych dziedzin pokrewnych wykazała brak ostrej definicji tych terminów, a ponadto duża dowolność w nazywaniu obiektów systemu "metodą" lub "techniką", przy intuicyjnym ich rozumieniu jako "sposobu postępowania". Należy zauważyć, że definicja metody w "Słowniku języka polskiego" /11/ jest uwikłana w błędne koło: metoda to sposób postępowania, a sposób to metoda, forma postępowania. Propozycje niektórych autorów przyjęcia jako cechy rozróżniającej metodę od techniki konieczność instrumentalizacji działania w przypadku techniki nie znajdują potwierdzenia w praktyce terminologicznej literatury przedmiotu. Metody rozumiane jako modele działania ze względu na cel nazywane są często technikami, co stanowi źródło trudności w klasyfikacji lub typologizacji MiTO.

Z tego też względu przyjęto, że język informacyjny będzie odzworowywać cechy obiektów nazwanych umownie MiTO, ustalone według następujących kryteriów semantycznych, nawiązujących do teorii przypadków głębokich C.J. Fillmore'a /6/.

MiTO jest czynnością, instrumentem lub zasadą, która może być: prosta lub złożona, posiada określonego AGENSA, PACJENSA,

BENEFICJENTA, jest zrelatywizowana do pewnego STANU (jego korekcji lub tworzenia), ma zastosowanie uniwersalne lub ograniczone. Musi być opisana w pewnym JĘZYKU i zrelatywizowana do CZASU (przeszłego, teraźniejszego lub przyszłego).

Wymienione przykładowo kryteria różnicowania MiTO ujmowane w postaci binarnych opozycji są podstawą ustalenia sieci relacji paradygmatycznych, odwzorowywanych w leksyce JIW i ilustrujących pewne zależności klasyfikacyjne między obiektami systemu. Z punktu widzenia organizacji paradygmatycznej JIW monohierarchiczne ujmowanie MiTO okazuje się niesatysfakcjonujące, ponieważ zależności między nimi mają charakter mozaikowy, co wymaga polihierarchicznej struktury leksyki języka, pozwalającej przypisywać metodom różne wyrażenia określające ich cechy relewantne ze względu na miejsce (obszar) stosowania, działający podmiot i poddawany działaniu przedmiot, efekt działania, przeznaczenie wyników działania oraz cechy użytkowe, stopień trudności, wymaganą wiedzę, instrumentalizację, koszty oraz inne ograniczenia.

Wyrażenia raprezentujące cechy dystynktywne MiTO według wymienionych ogólnych kryteriów tworzą słownik deskryptorów, docelowo tezaurus o fasetowej organizacji pola semantycznego.

Indeksowanie MiTO odbywa się według zasad indeksowania ankietowego, między innymi ze względu na możliwość przygotowania charakterystyk wyszukiwawczych przez studentów Instytutu Organizacji Systemów Produkcyjnych w ramach zajęć dydaktycznych.

Przyjmuje się, że wymienione zasady organizacji leksyki w słowniku deskryptorów umożliwią w przyszłości reprezentację wiedzy w systemie w formie sieci semantycznych i przekształcenie SIMTOJ w

system inteligentny (ekspertowy) przy zastosowaniu odpowiedniego oprogramowanie.

#### ŚRODKI TECHNICZNE SYSTEMU

Podstawowym urządzeniem w fazie projektowania i eksperymentalnego uruchamiania systemu (bazy danych) jest mikrokomputer klasy IBM AT o następującej konfiguracji:

- mikroprocesor Intel 80286,
- pamięć operacyjna 1 MB RAM,
- 1 napęd dyskietek 5.25 cala (1.2 MB),
- 1 napęd dyskietek 5.25 cala (360 kB),
- dysk twardy (40 MB),
- karta graficzna EGA (256 KB pamięci RAM),
- monitor barwny EGA,
- drukarka Star NX 15; docelowo drukarka laserowa.

Do wprowadzania, aktualizacji i wyszukiwania opisów MiTD wykorzystuje się pakiet Micro CDS ISIS.

Urządzeniem uzupełniającym do udostępnienia informacji źródłowej (opisów szczegółowych) jest kopiarka kserograficzna.

#### WSPÓŁPRACA Z UŻYTKOWNIKAMI

Charakteryzowany system jest elementem działalności konsultacyjnej i badawczej na poziomie badań podstawowych, prowadzonej przez Instytut Organizacji Systemów Produkcyjnych Politechniki Warszawskiej. W związku z tym informacja wyszukiwana w SIMTOJ stanowi część konsultacji udzielanych przez specjalistów z Instytutu jednostkom produkcji i zaplecza naukowo-technicznego przedsiębiorstw przemysłowych. Między użytkownikami i operatorem systemu istnieje więc sprzężenie zwrotne, u-

możliwiającej aktualizację charakterystyk wyszukiwawczych MiTO w części opisu identyfikacyjnego, dotyczącej praktycznej oceny efektywności MiTO. Oznacza to ponadto ścisłe powiązanie modułów w SIMTOJ.

Funkcje modułu 5 - współpraca z użytkownikami zewnętrznymi są następujące:

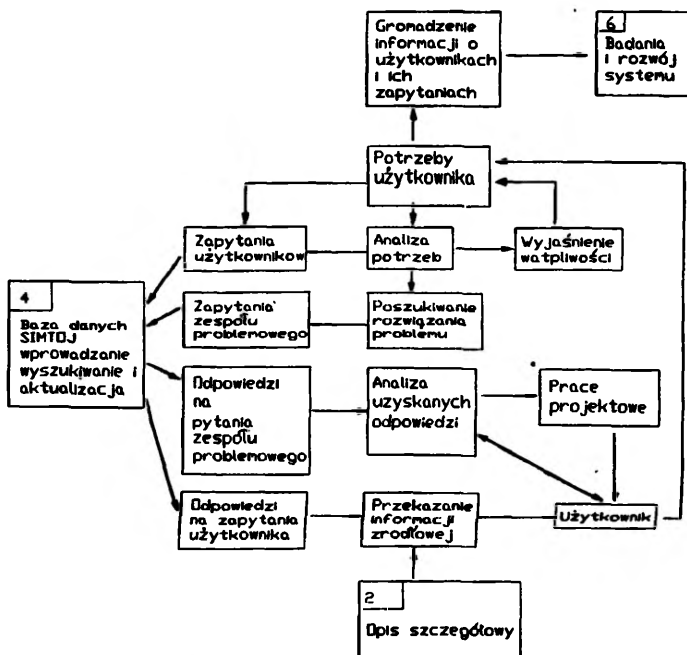
- analiza zapytań przy korzystaniu z centralnej bazy danych,
- formułowanie instrukcji wyszukiwawczych na podstawie zapytań,
- udostępnianie wyszukanej informacji wraz z ewentualną poszerzoną konsultacją,
- gromadzenie informacji o profilach zainteresowań użytkowników,
- udzielanie konsultacji merytorycznych w zakresie określonym w zapytaniu użytkownika.

Schemat modułu 5 współpracy z użytkownikiem ilustruje rysunek 3.

\* \* \*

System SIMTOJ jest projektowany jako narzędzie wspomagające przepływ między sferą badań podstawowych i przemysłem, stanowiące element warsztatu pracowników nauki i sposób pozyskiwania syntetycznej informacji faktograficznej w sterowaniu jakością w przemyśle elektromaszynowym. Sondażowe badania potrzeb użytkowników, reprezentujących przedsiębiorstwa przemysłowe, wykazały stosunkowo niewielkie zainteresowanie systemami informacji dokumentacyjnej jako niesatysfakcjonującymi z punktu widzenia potrzeby posiadania syntetycznej informacji bibliograficznej, niezbędnej w przemyśle oraz wskazania źródeł, gdzie można informację tę uzupełnić. Zapotrzebowanie na faktograficzno-dokumentacyjne systemy informacyjno-wyszukiwawcze wpływa na wymagania wobec języka informacyjno-wyszukiwawczego, to jest zmianę proporcji w odwzorowanej wiedzy tak, aby projektowany język mógł służyć do wyszukiwania informacji dokumentacyjnej kompletarnej wobec informacji faktograficznej, oczekiwanej przez użyt-





Rys. 3. Schemat modułu współpracy z użytkownikami

kownika. Ze zmianą zakresu i rodzaju informacji przetwarzanej w systemie wiąże się także rozbudowa języka wyszukiwawczego o moduł konwersacyjny wraz ze zróżnicowaniem formatów informacji wyjściowej.

#### Literetura

1. ALTSZULER G.S. Algorytm wynalazku. Warszawa 1974.
2. ANTOSZKIEWICZ J. Metody heurystyczne. Warszawa PWE 1982.
3. APPELT H.G., KRAMER F. Innowacje w przemyśle. Warszawa WNT 1978.
4. BIELICKA L. Języki informacyjno-wyszukiwawcze typu obiekt - atrybut prostych systemów faktograficznych. Prace IINTE 1984 nr 49.
5. CHWAŁEK B. Techniki organizatorskie: Gliwice. Wydawnictwo Politechniki Śląskiej 1980,
6. FILLMORE G.J. The case for case. W: Bach E., Harms R. Universals in linguistics theory. New York 1968.
7. MACHIŃSKA H., MALINOWSKI A. Wprowadzenie do technik decyzyjnych i organizatorskich. Warszawa. Wydawnictwa Uniwersytetu Warszawskiego 1985.
8. MARTYNIAK Z. Organizatoryka. Warszawa PWN 1987.
9. PAWLAK Z. Systemy informacyjne. Podstawy teoretyczne. Warszawa WNT 1983.
10. PAWLAK Z., TROCKI M. Metody organizatorskie. Warszawa SGPiS 1986.
11. Słownik języka polskiego. red. Szymczak M. Warszawa PWN 1978.
12. Wybrane zagadnienia jakości wyrobów w przedsiębiorstwach i zapleczu naukowo badawczym przemysłu elektromaszynowego. Raport z pracy badawczej CPBP 02.20.PI Warszawa Instytut Organizacji Systemów Produkcyjnych PW. Część I. Rok 1987, Część II. Rok 1988 (msszynopsis).

15.05.1989 r.

THE INFORMATION SYSTEM ON ORGANIZATIONAL METHODS  
AND TECHNIQUES IN QUALITY CONTROL (SIMTOJ)

Summary

The article is devoted to the characteristics of designing of the Information System on Organizational Methods and Techniques in Quality Control (SIMTOJ) at the Institute of Organization of Production Systems, Warsaw Polytechnics within the research problem CBBP 02.20.PI.

There are described: investigation of the information needs of the future users of the system and their participation in works at the stage of designing and of the preliminary implementation. The information resources of the system are outlined together with the computerized data base. The data base contains the factual and bibliographic data worked-out at three levels of description: the encyclopaedic, identificatory and the detailed one. The functional and organizational structure of the system is presented. The special attention is paid to the problem of designing of the information retrieval language. The technological tools of the system and the principles of cooperation with users are mentioned.

ИНФОРМАЦИОННАЯ СИСТЕМА ОБ ОРГАНИЗАТОРСКИХ МЕТОДАХ И ТЕХНИКАХ  
ДЛЯ СТИМУЛИРОВАНИЯ КАЧЕСТВОМ (SIMTOJ)

Резюме

Статья посвящена характеристике системы об организаторских методах и техниках для стимулирования качеством (SIMTOJ) разрабатываемой в Институте организации производственных систем,

аравского политехнического института в рамках исследовательской программы СВВР 02.20.Р1.

писаны: исследование информационных потребностей будущих потребителей системы и их участие в работах на этапе проектирования предварительного внедрения. Охарактеризованы информационные фонды системы, включая автоматизированную базу данных. База одержит фактографические и библиографические данные обработанные а трёх уровнях описи: энциклопедическом, идентификационном детальном. Представлены: функциональная и организационная структура системы.

Особое значение придаю проблеме разработки информационно-поискового языка. Указаны технологические средства системы и принципы кооперации с потребителями.



## MATERIAŁY I PRZYCZYNKI

DOROTA GŁĄZEK

Biblioteka Główna i Ośrodek  
Informacji Naukowo-Technicznej PWr.

### MOŻLIWOŚĆ ZWIĘKSZENIA EFEKTYWNOŚCI WYSZUKIWANIA INFORMACJI W JĘZYKU SWOBODNYCH SŁÓW KLUCZOWYCH

Słownik swobodnych słów kluczowych systemu informacji o zakończonych pracach naukowych Politechniki Wrocławskiej. Propozycja poprawy efektywności wyszukiwania informacji w systemie metodą eliminowania błędów w słowach kluczowych oraz tworzenia zbioru ciągów terminologicznych o zmniejszających się zakresach.

Politechnika Wrocławska jest jedną z nielicznych w kraju uczelnią posiadającą zautomatyzowany system informacji o zakończonych pracach naukowych swoich pracowników /3/. Jest to system SINT/NB, jeden z podsystemów systemu APIN (Automatyzacja Przetwarzania Informacji Naukowej), wdrożony w 1972 r. Oprócz dokumentowania prac i generowania zestawień wspomagających sterowanie badaniami naukowymi w uczelni, zajmuje się on także rozpowszechnianiem w kraju informacji o jej osiągnięciach naukowych. W trakcie wieloletniej eksploatacji tego systemu w Politechnice Wrocławskiej utworzono bazę danych na taśmie magnetycznej zawierającą prawie 70 tysięcy zapisów. Każdy z zapisów informuje o jednej zakończonej pracy naukowej, z podaniem jej pełnego opisu bibliograficznego oraz infor-

---

"Zagadnienia Informacji Naukowej" 1989 nr 1(54)

macji dodatkowych, takich jak np. analiza dokumentacyjna czy słowa kluczowe.

Do indeksowania prac wprowadzanych do bazy danych zastosowano język swobodnych słów kluczowych, rozwiązanie to uznano za optymalne dla systemu. System był w kraju pionierski, nie było możliwości praktycznego zweryfikowania rozwiązań znanych z literatury.

W trakcie prac nad założeniami projektowymi systemu ustalono, że obowiązującym językiem informacyjnym będzie język swobodnych słów kluczowych a indeksowanie będzie wykonywane przez autorów prac. Decyzja ta spowodowana była przede wszystkim ówczesnym słabym zaawansowaniem w kraju prac nad tezaurusemi z zakresu dziedzin leżących w kręgu zainteresowania wielowydziałowej uczelni oraz koniecznością szybkiego wdrożenia systemu.

Język swobodnych słów kluczowych jest językiem informacyjno-wyszukiwawczym najbardziej spośród wszystkich języków zbliżonym do naturalnego, nie posiada wad języków sformalizowanych, takich jak np. trudności w wyrażeniu terminów najnowszych, ale nie ma też ich zalet ułatwiających sformalizowane wyszukiwanie informacji. Obecnie w bibliotekach i ośrodkach informacji dysponujących zautomatyzowanymi systemami wyszukiwania informacji, poza tradycyjnymi katalogami systematycznymi i przedmiotowymi, najczęściej stosuje się właśnie język słów kluczowych /1/. Jest to także zgodne z założeniami zapisanymi w dokumencie programowym "System Informacji Naukowej, Technicznej i Organizacyjnej SINTO". Wprowadzenie indeksowania swobodnymi słowami kluczowymi - podyktowane koniecznością - po wieloletniej eksploatacji systemu okazało się decyzją trafną. Przy ocenie efektywności systemu należy jednak pamiętać o tym, że system SINT/NB w założeniu przeznaczony był do wspomagania zarządzania badaniami naukowymi i nie był (bądź był, ale tylko w sposób marginalny) wykorzystywany do wyszukiwań tematycznych.

Obecnie, między innymi na skutek znaczącego wzrostu zasobu wprowadzonych prac, coraz częściej zdarzają się zamówienia na retrospektywne wyszukiwanie informacji tematycznej. Są one realizowane przez inny podsystem, za pośrednictwem przekazywanej tam bazy danych SEBAN, dostępnej jedynie w trybie selektywnej dystrybucji informacji. Mamy jednak świadomość strat informacji spowodowanych m.in. niedostatkami istniejącego modelu indeksowania prac.

Tworzeniem słów kluczowych projektanci obarczyli użytkowników, zakładając słusznie, że są oni najbardziej kompetentni będąc jednocześnie autorami prac. Prace, których opisy są wprowadzane do systemu dotyczą najnowszych odkryć i technologii, a ich słownictwo jest dopiero przyswajane przez język polski (najczęściej z literatury anglojęzycznej). Po uwzględnieniu wszystkich powyższych faktów nasuwał się wniosek, że kontrolę słownictwa należy przenieść na etap wyszukiwania informacji. Próba przeanalizowania słownika słów kluczowych została dokonana pod kierunkiem doc. Bożenny Bojar jako praca dyplomowa na Studium Podyplomowym Informacji Naukowej /2/.

#### ANALIZOWANY SŁOWNIK SŁÓW KLUCZOWYCH

Przy tworzeniu badanego słownika założono ograniczenie analizowanych prac do około jednego tysiąca, zdecydowano także zawęzić tematykę prac - przeanalizowano prace wykonane w Instytucie Technologii Elektronowej Politechniki Wrocławskiej w latach 1979-1987. Ważnym zagadnieniem przy wyborze określonej tematyki słownika był fakt, że już w trakcie pobieżnego przeglądu można było zauważyć istnienie w wydzielonej tematycznie grupie słownika spory zbiór słów kluczowych, które mogą mieć zastosowanie we wszystkich lub w wielu dziedzinach wiedzy. Owało to rękojmię reprezentatywności wybranej próby także dla innej tematyki. Niestety



oprogramowanie systemu nie zapewniło możliwości przeanalizowania częstości występowania poszczególnych słów kluczowych.

Jedną z pierwszych czynności było zidentyfikowanie błędów w indeksowaniu. Błędy te podzielono na 9 typów, które dokładnie omówiono poniżej.

Typ 1. Niejednolita gramatycznie forma tego samego terminu:

a) Różnice w strukturze syntaktycznej terminów złożonych

Przykłady: WIĄZKA ELEKTRONÓW            LUTOWANIE PRÓŻNIOWE  
              WIĄZKA ELEKTRONOWA        LUTOWANIE W PRÓŻNI

b) Używanie liczby pojedynczej lub mnogiej dla tego samego terminu

Przykłady: CERMET                        CZUJNIK TEMPERATURY  
              CERMETY                      CZUJNIKI TEMPERATURY

Typ 2. Różny szyk wyrazów w wyrażeniu złożonym:

Przykłady: REZYSTOR CIENKOWARSTWOWY    WARSTWA GRUBA  
              CIENKOWARSTWOWY REZYSTOR    GRUBA WARSTWA  
  
              KONDENSATOR CIENKOWARSTWOWY    WARSTWA CIENKA  
              CIENKOWARSTWOWY KONDENSATOR    CIENKA WARSTWA

Typ 3. Używanie lub pomijanie w nazwie pewnych elementów:

a) całych wyrazów

Przykłady: SPEKTRUM AUGERA                TRAWIENIE JONOWE  
              SPEKTRUM ELEKTRONÓW AUGERA    TRAWIENIE WIĄZKĄ JONÓW

b) jednego z członów wyrazów złożonych

Przykłady: TECHNOLOGIA ELEKTRONOWA  
              TECHNOLOGIA ELEKTRONOWIĄZKOWA  
              SPAWARKA ELEKTRONOWA  
              SPAWARKA ELEKTRONOWIĄZKOWA

Typ 4. Niekonsekwentne używanie spolszczonej pisowni wyrazów pochodzących z języka obcego:

Przykłady: MIKROSKOP SCANNINGOWY

MIKROSKOP SKANINGOWY

MIKROSKOP SKENINGOWY

Typ 5. Używanie różnych terminów dla określenia tego samego denotatu:

Przykłady: TECHNIKA PRÓŻNI

OPTYKA ELEKTRONOWA

WYTWARZANIE PRÓŻNI

OPTOELEKTRONIKA

Typ 6. Używanie terminów nieprecyzyjnych:

Przykłady: URZĄDZENIA BARDZO WYSOKIEJ PRÓŻNI

TESTOWANIE POŁĄCZEŃ

Typ 7. Tworzenie długich fraz, w których skład wchodzi samodzielne słowa kluczowe:

Przykłady:

DYFRAKCJA ODBICIOWA ELEKTRONÓW O DUŻYCH ENERGIACH (RHEED)

Występują tu samodzielne słowa kluczowe:

DYFRAKCJA ODBICIOWA

RHEED

ELEKTRON

STRUKTURA KRYSTALOGRAFICZNA ARSENIKU GALU

Występują tu samodzielne słowa kluczowe:

STRUKTURA KRYSTALOGRAFICZNA

ARSENEK GALU

Typ B. Dznaczenie związków chemicznych:

a) symbolami lub odpowiednikami słownymi

Przykłady: AL2O3

ZNO

TLENEK GLINU

TLENEK CYNKU

b) stosowanie polskiej lub łacińskiej nazwy pierwiastka w nazwie związku

Przykład: AZOTEK GLINU  
AZOTEK ALUMINIUM

Typ 9. Błędy powstałe w trakcie wprowadzania danych:

a) "obcięcie" części tekstu

Przykłady: EN EK CYNKU                      ĄZKA ELEKTRONOWA  
zamiast: TLENEK CYNKU      zamiast: WIĄZKA ELEKTRONOWA

b) wprowadzenie tekstu stanowiącego symbole kodowe

Przykłady: REZYSTANCJEA1Ø                      POLEROWANIE CHEMICZNOE1Ø  
zamiast: REZYSTANCJE      zamiast: POLEROWANIE CHEMICZNE

c) błędy literowe

Przykłady: IMPLANTACJA IONÓW  
zamiast: IMPLANTACJA JONÓW  
MASKA RENTGENOLIGOGRAFICZNA  
zamiast: MASKA RENTGENOLITOGRAFICZNA

d) wprowadzenie informacji stanowiącej zawartość pól o innych etykietach

Przykłady:

IRENA            - zawartość podpole H w polu o etykiecie 1ØØØ lub 7ØØØ  
CZARCZYŃSKI-    - zawartość podpole A w polu o etykiecie 1ØØØ lub 7ØØØ

Już pobieżne przeanalizowanie wpływu błędów w tworzeniu słów kluczowych na efekty wyszukiwania informacji pozwala zauważyć, że nie wszystkie błędy mają podobną wagę. Np. użycie słowa kluczowego TRANZYSTOR lub TRANZYSTORY może nie mieć wpływu na wyszukiwanie w systemie pozwalającym na maskowanie końcówek. Terminem wyszukiwanym w takim wypadku będzie TRANZYSTO\$, gdzie \$ - symbol maskowania.

Jednym z wymiernych efektów tej części przeprowadzonej pracy było powstanie wskazówek metodycznych, które powinny doprowadzić do wyeliminowania dużej liczby błędów w indeksowaniu.

## METODA POLEPSZENIA EFEKTYWNOŚCI WYSZUKIWANIA INFORMACJI WEDŁUG SŁÓW KLUCZOWYCH

Mimo wymienionych licznych błędów indeksowania i związanej z tym jakości zgromadzonego słownika zaproponowano metodę, której zastosowanie powinno spowodować polepszenie efektywności wyszukiwania informacji w istniejącym zbiorze.

Inspiracją do opracowania zaproponowanej metody była praca Jadwigi Sambor /4/. Autorka przytacza tam stwierdzenie Altmana i Kinda, że systemy klasyfikacji oparte na zhierarchizowanym układzie pojęć najbardziej odpowiadają strukturze ludzkiej pamięci, człowiek najlepiej poznaje i zapamiętuje dane pojęcie poprzez umieszczenie go we wcześniej już znanej strukturze pojęć zhierarchizowanych.

Jeśli uporządkujemy terminy zgodnie z definicjami słownikowymi, to wraz ze wzrostem wskaźnika oddalenia terminu od początku ciągu maleje jego zakreślenie znaczeniowe (wzbogaca się treść), a także maleje liczba terminów znajdujących się na tym samym poziomie szczegółowości.

Wniosek jest taki, że istnieje możliwość uporządkowania słów kluczowych zgodnie z zasadą wynikającą z ich definicji. Zaproponowano i wykonano uporządkowanie terminów występujących w słowniku w tzw. ciągu terminologiczne wyrażen o zmniejszających się zakresach znaczeniowych i coraz bogatszej treści.

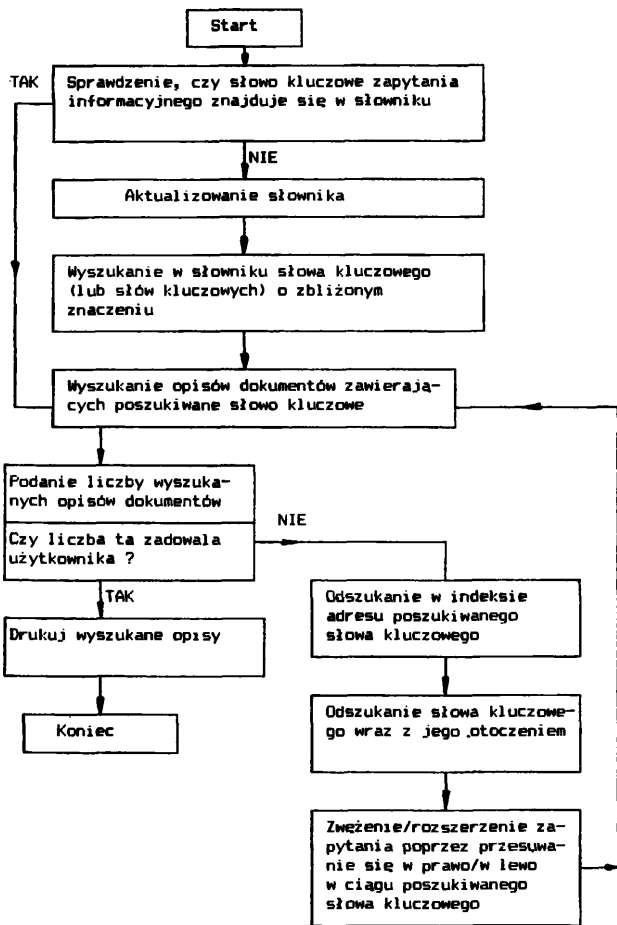
Przykłady:

PRÓŻNIA - WYSOKA PRÓŻNIA - ULTRAWYSOKA PRÓŻNIA

TECHNIKA PRÓŻNI - TECHNIKA WYSOKIEJ PRÓŻNI

OYFUZJA - TEORIA OYFUZJI - OYFUZYJNOŚĆ - DROGA OYFUZJI = DROGA

OYFUZJI NOŚNIKA - ŚREDNIA DROGA OYFUZJI - DŁUGOŚĆ DROGI OYFUZJI

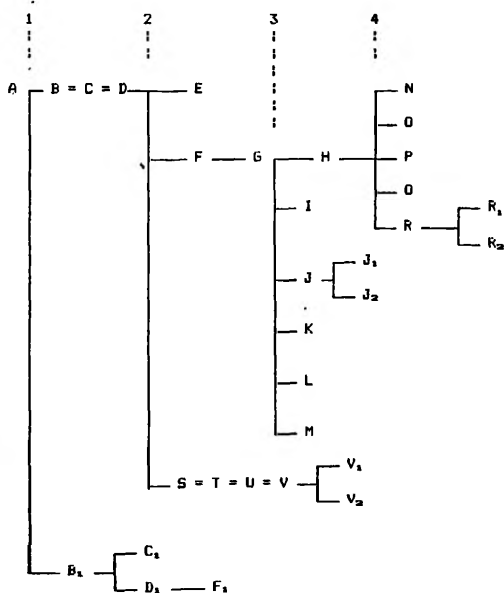


Rys. 1 Schemat wyszukiwania informacji z zastosowaniem zbioru ciągów terminologicznych wyrażeń o zmniejszających się zakresach znaczeniowych.

Z ostatniego przykładu widać, że dopuszczono istnienie kilku terminów na tym samym poziomie szczegółowości danego ciągu terminologicznego (DROGA DYFUZJI = DROGA DYFUZJI NOŚNIKA). Dodatkowo zaproponowano utworzenie indeksu określającego adres danego słowa kluczowego w zbiorze, tzn. możliwości odnalezienia go w odpowiednim ciągu terminologicznym. Przy wymienionych założeniach organizacyjnych program wyszukiwania informacji powinien działać zgodnie ze schematem zamieszczonym na rys. 1.

### Przeprowadzony eksperyment

Przeprowadzono eksperyment uporządkowania 1527 słów kluczowych w ciągu terminologiczne wyrazów o zmniejszających się zakresach znaczeniowych. Utworzono 63 ciągi terminologiczne o bardzo zróżnicowanej liczbie terminów w jednym ciągu (2 - 33). Na rys. 2 przedstawiono schemat ciągu o maksymalnej liczbie terminów.



Rys. 2 Schemat ciągu terminologicznego wyrazów o maksymalnej liczbie terminów.

W trakcie tworzenia ciągów przyjęto pewne odstępstwa od reguły ścisłego "zawierania się" zakresów znaczeniowych, co zostało zrekompensovane poprawieniem dokładności wyszukiwania informacji. Przyjęto założenia:

wytwarzanie X	zawiera się w X
urządzenie do wytwarzania X	zawiera się w X
urządzenie pomocnicze do X	zawiera się w X
teoria o X	zawiera się w X

Wydzielone też zostały następujące grupy słów kluczowych, z których nie tworzą ciągów terminologicznych:

- wzory i symbole pierwiastków i związków chemicznych  
(uporządkowano je alfabetycznie)

- słowa kluczowe nie dające się uporządkować w ciągi terminologiczne.

Przykłady: AUTOMATYZACJA

CHARAKTERYSTYKA ROBOCZA

- słowa kluczowe zbyt ogólne, nieprzydatne do wyszukiwania informacji

Przykłady: FIZYKA NIEZAWODNOŚCI

ELEKTRONIKA CIAŁA STAŁEGO

- słowa kluczowe błędnie utworzone, nieprzydatne do wyszukiwania informacji

Przykłady: MARKOW                      powinno być: ŁAŃCUCH MARKOWA

LUB      PROCES MARKOWA

- słowa kluczowe błędnie zapisane w zbiorze

Przykłady: GNTAL

IMPLANTACJA IONÓW

Na zakończenie eksperymentu przeprowadzono próbę wyszukiwania informacji z zastosowaniem zaproponowanej metody. Wyszukiwanie zrea-

lizowano (manualnie) dla kilku pytań. W celu zilustrowania wyników jedno z nich zostanie omówione bardziej szczegółowo.

Hasło wyszukiwania: ROZPRASZANIE WSTECZNE ELEKTRONÓW

Indeks alfabetyczny kieruje do ciągu, którego pierwszy człon stanowi termin: DETEKCJA ELEKTRONÓW.

Odnajdujemy w jego strukturze termin ROZPRASZANIE WSTECZNE ELEKTRONÓW i stwierdzamy, że dla wyrażenia tego samego pojęcia były używane dodatkowo następujące terminy:

1. ROZPRASZANIE WSTECZNE
2. WSTECZNE ROZPRASZANIE
3. WSTECZNE ROZPRASZANIE ELEKTRONÓW

Wszystkie dokumenty zawierające słowa kluczowe 1-3 zostałyby pominięte przy tradycyjnej metodzie wyszukiwania informacji. Dodatkowo ze struktury zbudowanego ciągu można odczytać, że istnieją następujące słowa kluczowe o szerszym zakresie znaczeniowym:

1. SYGNAŁ EWR
2. SYGNAŁ ELEKTRONÓW WSTECZNIE ROZPROSZONYCH
3. STRUMIEŃ WSTECZNY
4. WSTECZNIE ROZPROSZONE ELEKTRONY
5. DETEKTOR EWR
6. DETEKTOR ELEKTRONÓW
7. DETEKTORY ELEKTRONÓW
8. DETEKCJA ELEKTRONÓW

o tym samym poziomie szczegółowości zakresu znaczeniowego:

1. WSPÓŁCZYNNIK WSTECZNEGO ROZPRASZANIA
2. WSPÓŁCZYNNIK ROZPRASZANIA WSTECZNEGO
3. MECHANIZM ROZPRASZANIA
4. MACIERZ ROZPROSZENIA
5. STOŻEK WSTECZNEGO ROZPRASZANIA.



### **Podsumowanie wyników wyszukiwania**

Przy wyszukiwaniu zgodnie z hasłem ROZPRASZANIE WSTECZNE ELEKTRONÓW (jeśli założymy, że jedno słowo kluczowe występuje tylko jeden raz) otrzymujemy 1 dokument. Po zastosowaniu ciągu terminologicznego i zidentyfikowaniu terminów o tym samym zakresie znaczeniowym otrzymujemy jeszcze 3 dokumenty. W odszukanym ciągu terminologicznym identyfikujemy jeszcze 3 dokumenty ze słowami kluczowymi będącymi terminami o tym samym poziomie szczegółowości zakresu znaczeniowego oraz 5 dokumentów ze słowami kluczowymi będącymi terminami o węższym zakresie znaczeniowym. W sumie o 11 dokumentów więcej niż byłoby to możliwe bez zastosowania proponowanej metody.

Brak możliwości podania informacji o częstości występowania słów kluczowych uniemożliwia pełną ocenę uzyskanych wyników. Jednak nawet przy bardzo ostrożnej ocenie osiągniętych rezultatów wyniki są obiecujące i warte dalszych badań, oczywiście po ulepszeniu możliwości programowych systemu.

### **Podsumowanie wyników pracy**

Stwierdzono możliwość znacznej poprawy kompletności i trafności wyszukiwania informacji przy istniejącym modelu indeksowania prac naukowych swobodnymi słowami kluczowymi. Możliwości te istnieją w dwu kierunkach:

- wyeliminowanie błędów w słowach kluczowych,
- utworzenie dodatkowego zbioru ciągów terminologicznych wyrażań o zmniejszających się zakresach znaczeniowych oraz zaprojektowanie programu wyszukiwania informacji działającego z uwzględnieniem utworzonego zbioru ciągów.

Należy podkreślić, że przydatność zaproponowanej metody sprawdzono w specyficznych warunkach uczelnianego systemu informacji

o własnych pracach naukowych. Można twierdzić, że sprawdzi się ona w systemach podobnego typu, natomiast oczywiście wydaje się, że nie jest to metoda uniwersalna, przydatna w warunkach przeciętnego systemu informacji naukowej.

#### Literatura

1. BIELICKA L.A., ŚCIBOR E. Tendencje rozwoju języków informacyjnych w Polsce do 1990 i 2000 roku. APIO 1987 R.32 nr 2 s.8-15
2. GŁAZEK D. Metoda polepszenia efektywności wyszukiwania informacji w systemie stosującym język swobodnych słów kluczowych. Raporty Biblioteki Gł. i OINT Politechniki Wrocławskiej 1987 ser. SPR nr 9B (maszyn. powiel)
3. Przegląd zautomatyzowanych systemów biblioteczno-informacyjnych Politechniki Wrocławskiej. Skrypt. Warszawa CİNTE 1984 Materiały Szkoleniowe nr 35
4. SAMBOR J. O budowie tzw. ciągów definicyjnych (na materiale definicji leksykalnych). Biuletyn Polskiego Towarzystwa Językoznawczego 1986 z. 40 s.151-165.

28.04.1989 r.

#### POSSIBILITY OF INCREASING THE EFFICIENCY OF INFORMATION RETRIEVAL BY MEANS OF THE FREE KEY WORDS

##### Summary

There is characterized the dictionary of the key words used in the functioning information system on the completed research at the Wrocław Polytechnics. The author analyses the errors occurring in indexing of documents. The proposal of a method of information retrie-

val is presented. The method consists in constructing of the set of terminological strings the range of which is decreasing. The conducted experiments are described. The article ends with the conclusions concerning the possibility of using of the suggested method in other information systems.

ВОЗМОЖНОСТЬ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ПОИСКА ИНФОРМАЦИИ  
НА ЯЗЫКЕ СВОБОДНЫХ КЛЮЧЕВЫХ СЛОВ

Резюме

Охарактеризован словарь ключевых слов используемый в действующей системе о законченных научных исследованиях во Вроцлавском политехническом институте. Автор анализирует ошибки индексирования. Предложенный метод заключается в создании набора терминологических цепей с уменьшающимися объемами. Статья закончена выводом относящимся к возможности использования предлагаемого метода в других информационных системах.

PETER ŹAK

Biblioteka Narodowa

## MIĘDZYNARODOWY FORMAT WYMIANY DANYCH BIBLIOGRAFICZNYCH

Międzynarodowy format wymiany danych bibliograficznych - UNIMARC Manual. Historia powstania i zawartość podręcznika. Struktura i zawartość formatu UNIMARC (etykieta rekordu, bloki informacyjne). Forma zapisu danych. Załączniki. Zalety formatu.

Nadrzędnym punktem odniesienia dla rozumienia kwestii tworzenia i upowszechniania międzynarodowych formatów wymiennych dotyczących informacji bibliograficznej są powiązane ze sobą programy głównie IFLA: Powszechnej Dostępności Publikacji (Universal Availability of Publications - UAP), Powszechnej Rejestracji Bibliograficznej (Universal Bibliographic Control - UBC) oraz Międzynarodowy Program MARC (International MARC Programme - IMP). Prace nad formatem UNIMARC stanowią integralną część realizacji zwłaszcza dwóch ostatnich programów, które zostały w 1987 roku - ze względu na wspólne cele, przedmiot i środki działania - połączone w jeden program, tzw. IFLA UBCIM Programme (IFLA Universal Bibliographic Control and International MARC Programme).

Wiodącym ośrodkiem realizacji tego programu jest Biblioteka Brytyjska w Londynie.

Jednym z kluczowych zagadnień programu jest zapewnienie wymiennalności informacji bibliograficznych między poszczególnymi kra-

---

"Zagadnienia Informacji Naukowej" 1989 nr 1(54)

jami. Jeśliby pominąć skądinąd istotną sprawę zakresu wymiany (informacji o różnych rodzajach dokumentów)<sup>1/</sup>, to efektywność wymiany warunkują z jednej strony porównywalność danych zależna od spójności przepisów i zwyczajów opracowywania materiałów<sup>2/</sup>, zaś z drugiej strony ujednoczona forma przekazu (odpowiednio sformatowane dane na odpowiednim nośniku).

Głównym celem UNIMARC jako formatu międzynarodowego mieszczącego w sobie wszystkie elementy opisu (i nie tylko) jest więc ułatwienie wymiany danych bibliograficznych między narodowymi ośrodkami bibliograficznymi w postaci czytelnej przez programy komputerowe. Jednocześnie UNIMARC może posłużyć jako model, wzorcowe rozwiązanie problemu określania treści, struktury i formy danych bibliograficznych (częściowo też bibliotecznych) w kształcie umożliwiającym automatyczne przetwarzanie we własnym systemie krajowym. Świadczy o tym fakt, że UNIMARC stał się w wielu krajach podstawą formatu narodowego<sup>3/</sup>. Inaczej mówiąc stanowi on już w dużym stopniu sprawdzony punkt odniesienia, narzędzie tworzenia formatów opisu dokumentów tam, gdzie narodowe formaty bibliograficzne nie zostały jeszcze opracowane lub dopracowane, gdzie buduje się automatyzowane systemy bibliograficzne i biblioteczno-informacyjne za-

---

1/ Na przykład w polskiej bibliografii narodowej bieżącej druków zwartych nie są dotychczas uwzględnione tzw. materiały nieksiążkowe, a także publikacje z tzw. "drugiego obiegu".

2/ W Polsce brakuje nowych norm opisu dla wydawnictw ciągłych oraz artykułów, stosuje się nieco inne zasady opisu dla ISDS oraz bibliografii narodowej.

3/ Zob. International Guide to MARC Database and Services. National Magnetic Tape Services. Ed. by D. Wolf and F. Conrad. Frankfurt a. Main 1984. Przyjęcie UNIMARC za podstawę formatu MARC-BN zadeklarowała również Biblioteka Narodowa w Warszawie (s. 117).

kładające wymianę informacji oraz współpracę przy tworzeniu baz danych o dokumentach.

UNIMARC będący z natury formatem wymiannym nie stawia żadnych warunków, nie formułuje wskazówek odnośnie rekordów i zawartych w nich danych przetwarzanych w poszczególnych systemach krajowych. Dostarcza jednak takich zaleceń wóczas, kiedy dane bibliograficzne mają podlegać wymianie międzynarodowej. Od danego ośrodka krajowego prowadzącego wymianę zależy więc, czy decyduje się na przyjęcie własnego formatu (a więc szerbką konwersję danych na międzynarodowy format wymienny), dostosowanie własnego już istniejącego formatu do UNIMARC czy wręcz na przejście UNIMARC w całości. Jeśli jednak podejmuje się tego typu decyzje, to względy racjonalne - ekonomiczne, techniczne i inne - każdą dostrzec potrzebę maksymalnego dostosowania własnego formatu narodowego do formatu wymiennego<sup>4/</sup>. W przypadku Polski sprawa teoretycznie wydaje się nieco łatwiejsza, ponieważ nie został tu jeszcze stworzony system wewnętrznej wymiany danych na nośnikach maszynowych. Nie mamy więc do czynienia z jakimś powszechnie i dłużej stosowanym formatem krajowym (lub formatami lansowanymi przez duże biblioteki), co by utrudniało adaptację UNIMARC. Należy bowiem pamiętać, iż znajdujemy się w nieco innej sytuacji niż swego czasu niektóre kraje zachodnie jak Stany Zjedno-

---

<sup>4/</sup>Warto tu przypomnieć, że już w 1978 r. rozpoczęto wstępne rozmowy nt. wymiany rekordów bibliograficznych na nośnikach maszynowych pomiędzy Biblioteką Brytyjską oraz Biblioteką Narodową w Warszawie. Z ramienia British Bibliographic Services brał w nich udział Richard E. Coward, który również uczestniczył rok później w sympozjum "Rola kontaktów międzynarodowych w działalności współczesnych bibliotek" zorganizowanym w Warszawie. W swoim referacie wyraził nadzieję dalszej współpracy oraz przedstawił problemy i trudności związane z międzynarodową wymianą rekordów w sieci MARC. W odniesieniu do Polski wiele zdawartych tam stwierdzeń wydaje się nadal aktualne. Zob. R. E. Coward: Międzynarodowa sieć MARC: zarządzanie i rozwój. Warszawa, Uniwersytet Warszawski, Instytut Bibliotekoznawstwa i Informacji Naukowej 1979 [maszyn.powieł].

czony czy Wielka Brytania. Tam zaczynano od tworzenia krajowych sieci katalogowania w postaci czytelnej dla maszyny (MARC), powstały one w końcu lat sześćdziesiątych i dopiero nieco później rozpoczęła się wymiana w skali międzynarodowej.

UNIMARC opublikowano po raz pierwszy w 1977 r. jako rezultat prac grupy roboczej IFLA kierowanej przez Henriette D. Avram. Drugie wydanie pojawiło się w 1980 r. za sprawą Sally H. Mc Callum. Trzy lata później - w wyniku upowszechniania się formatu i zaistniałej potrzeby jego jednolitej interpretacji - wydano UNIMARC Handbook, opracowany przez Alana Hopkinsona z udziałem S.H. Mc Callum oraz Stephena P. Davisa. Ze względu na wzrastające zainteresowanie tym formatem ośrodków narodowych, potrzebę objęcia nim również innych materiałów niż druki zwarte, ciągłe oraz kartograficzne, a także ze względu na bieżące ustalenia ISBD, Komitet Doradczy MARC (International MARC Advisory Committee) w myśl wytycznych Międzynarodowego Programu MARC (IMP) zatwierdził w 1985 r. podjęcie prac nad UNIMARC Manual. Ustalono, iż UNIMARC Manual zostanie opracowany w Bibliotece Brytyjskiej przy współpracy z pracownikami Biblioteki Kongresu w Waszyngtonie. W ten sposób powstał w 1987 r. UNIMARC Manual<sup>5/</sup>, który dzięki poszerzeniu i połączeniu poprzednich opracowań UNIMARC Format oraz UNIMARC Handbook zawiera zaktualizowany format wraz z wykładnią oraz instrukcją.

Należy podkreślić, że format został najpierw przetestowany w Bibliotece Brytyjskiej, Bibliotece Kongresu oraz innych ośro-

---

<sup>5/</sup> UNIMARC Manual. Ed. by Brian P. Holt with the assistance of Sally Mc Callum and A. B. Long. London: IFLA Universal Bibliographic Control and International MARC Programme, British Library Bibliographic Services 1987, 482 s.

kach narodowych; dopiero potem zatwierdzono proponowane zmiany czy uzupełnienia i wprowadzono je do formatu. W tym miejscu jedna uwaga - otóż w przedmowie do UNIMARC Manual czytamy, że chodzi tu o ostateczne określenie formatu ("definitive statement of the UNIMARC Format"). Stwierdzenia tego typu są nieco mylące, gdyż prace nad formatami mają charakter ciągły. Również w przypadku UNIMARC grupa ekspertów czuwa nad tym, by był on formatem "żywym" i otwartym, to znaczy uwzględniającym istotniejsze zmiany mogące rzutować na jego adekwatność - jak np. nowelizację przepisów katalogowania, a pojawiają się nowe rodzaje dokumentów, lokalne potrzeby systemowe itp. Zgodnie z tym przewiduje się nowe wydanie UNIMARC już w 1990 r.

UNIMARC Manual można podzielić na dwie części, które uzupełniają 14 załączników oraz indeks przedmiotowy.

Jeśli by nie liczyć krótkiej przedmowy to na pierwszą część składają się: wprowadzenie oraz trzy kilkunastonicowe rozdziały (zatytułowane: Organizacja podręcznika; Struktura formatu; Etykieta rekordu i pola danych - informacje ogólne), gdzie zawarto niezbędne informacje dotyczące całego opracowania-podręcznika oraz samego formatu. Ułatwiają one posługiwanie się podręcznikiem, przybliżają pojęcie formatu, rekordu oraz określają podstawy ich tworzenia. Czytelnik może więc zapoznać się z uwzględnionymi normami, podstawowymi definicjami (poza typowymi definicjami bibliograficznymi), sposobem charakterystyki pól, znakami używanymi przy tworzeniu rekordu, logiczną i fizyczną strukturą rekordu, wstępną charakterystyką jego składników (etykieta, tablicą adresów, segmentem pól zmiennej długości), zasadami wiązania pól i całych rekordów, kwestią powtarzalności danych i w ogóle z ich rodzajami oraz formą, użyciem znaków specjalnych, przestankowych itp.



Istotne są następujące stwierdzenia:

1. UNIMARC przeznaczony jest do opisu wydawnictw zwartych i ciągłych, materiałów kartograficznych, audiowizualnych, mikroform oraz w perspektywie dla zbiorów informacji na nośnikach maszynowych.

2. UNIMARC jest dostosowany do przepisów ISBD (International Standard Bibliographic Description) stanowiących podstawą norm narodowych.

3. Przestrzega on zaleceń normy ISO 2709 określającej wymogi w zakresie wymiany informacji bibliograficznych na nośnikach maszynowych.

Brakuje jednak bliższego wyjaśnienia odnośnie tzw. "Authorities Format", na który się autorzy podręcznika powołują zarówno we wstępie jak i innych miejscach<sup>6/</sup>.

Druga, zasadnicza część podręcznika dotyczy struktury i zawartości formatu UNIMARC - etykiety rekordu (tzw. lidera) oraz poszczególnych bloków informacyjnych obejmujących pola stałej i zmiennej długości, w których zawarto informacje identyfikacyjne, zakodo-

---

<sup>6/</sup> Informacje na temat tego formatu można znaleźć w powielonym materiale: UNIMARC/Authorities. Universal MARC Format for Authorities. Draft. IFLA Steering Group for Authorities Format, 1988 - 01-015. Z informacji zawartych w powyższym projekcie stanowiącym obecnie przedmiot szerokiej konsultacji wynika, iż "Authorities Format" jest przeznaczony dla ujednoliconych i innych form haseł głównych. W oparciu o UNIMARC oraz Guidelines for Authority and Reference Entries (GARE) umożliwia tworzenie różnych rekordów powiązanych z rekordami opisu dokumentów w formacie UNIMARC. Na podstawie "Authorities Format" można tworzyć rekordy:

- haseł ujednoliconych
- haseł odsyłaczowych (skierowujących)
- haseł ogólnych wyjaśniających (naprowadzających)
- danych (haseł) równoległych
- danych (haseł) w innym alfabecie
- haseł złożonych,

przy czym hasła mogą być zawarte w pierwszym typie rekordu przeznaczonego dla haseł ujednoliconych i następnie z niego generowane lub - w zależności od potrzeb - tworzone odrębnie jako inne rekordy.

wane oraz tekstowe. Pominiemy lidera, który jest niezbędną częścią rekordu o stałej długości 24 znaki i pełni określone funkcje systemowe przy przetwarzaniu rekordu<sup>7/</sup>, skoncentrujemy się natomiast na omówieniu informacji podzielonych na bloki oraz na sposobie zapisu danych.

W UNIMARC przyjęto następujące bloki informacyjne:

- 0 - - Informacje identyfikacyjne
- 1 - - Informacje zakodowane
- 2 - - Opis bibliograficzny dokumentu
- 3 - - Uwagi
- 4 - - Informacje o powiązaniach rekordu
- 5 - - Tytuły
- 6 - - Opis rzeczowy dokumentu
- 7 - - Informacje związane ze strefą odpowiedzialności
- 8 - - Blok do użytku międzynarodowego
- 9 - - Blok do użytku narodowego

Bloki są podzielone na pola oznaczone trzycyfrowymi etykietami począwszy od 001 do 999, zaś pola zawierają podpola oznaczone jednoznakową etykietą - separatorem (małą literą lub cyfrą). Kompletny wykaz pól według głównych bloków zamieszczono przed ich opisem (s. 18-19), zawiera również nieliczne, stosownie oznaczone pola zdezaktualizowane oraz zarezerwowane.

---

<sup>7/</sup> W UNIMARC nie sprecyzowano w jaki sposób tworzony jest lider w tej części, która nie jest generowana automatycznie (w przypadku kiedy wymienne rekordy bibliograficzne tworzone są od razu w formacie UNIMARC). Widocznie rozwiązanie problemu pozostawia się użytkownikowi formatu, który ma np. możliwość wprowadzenia do formatu dodatkowego pola z niezbędnymi dla lidera danymi. Por. Z. Mośczyńska-Pętkowska: System SABINA. Opis wydawnictw zwartych. Warszawa 1982 s.10-12, gdzie do formatu MARC-BN zamieszczono pole 002 - Etykieta rekordu.

Blok informacji identyfikacyjnych zawiera numer rekordu, numery ISSN, ISBN oraz numer bibliografii narodowej. Dzięki nim można dokument łatwo zidentyfikować.

Blok informacji zakodowanych obejmuje dane zapisane w sposób symboliczny, mówiące o typie publikacji, języku, alfabecie, miejscu opublikowania, okresie jej ukazywania się, przeznaczeniu czytelniczym, zawartości. Wzbogacony jest odpowiednimi wykazami kodów języków, krajów i innych potrzebnych oznaczeń. Pola bloku zawierają dane o stałej długości, które są zakodowane za pomocą znaków alfanumerycznych i znajdują się na stałych miejscach w określonym ciągu znaków. Inaczej mówiąc o znaczeniu danego kodu decyduje pozycja, na której został zapisany. Informacje z tego bloku oraz z bloku 0 - - pełnią funkcję kontrolno-sterującą przy przetwarzaniu, służą również na potrzeby szybkiego wyszukiwania umożliwiając bezpośrednio uzyskanie informacji o cechach formalnych i treściowych dokumentu. Znajdują się tu również informacje nie należące do tradycyjnego opisu (jak data założenia rekordu, kod rekordu modyfikowanego).

Blok ten należy do trudnych i może w praktyce sprawiać trudności interpretacyjne. Zastanawia między innymi odmienne traktowanie książek i czasopism jako rodzajów materiału (zawartych tekstów). W przypadku książek (w polu 105) określa się łącznie typ wydawnictwa i jego zawartość za pomocą trzech znaków, zaś dla czasopism przewiduje się tylko jeden znak określający dominującą cechę czasopisma jako źródła informacji (reference material).

Następny blok - opis bibliograficzny takich trudności nie przysparza, ponieważ jest wyjątkowo przejrzysty. Zawiera pola, które odpowiadają w przepisach ISBD strefom: tytułu i oznaczenia odpowiedzialności, wydania, strefie specjalnej, adresu wydawniczej

go, opisu fizycznego oraz serii. Poprawność wypełniania pól bloku będzie więc zależała głównie od stopnia znajomości zasad opisu bibliograficznego.

Kolejny blok - uwagi - również odpowiada strefie w ISBD. Stwarza możliwość wyrażania uwag w sposób ogólny lub szczegółowy. Użytkownik formatu może blok uwag wypełniać według własnego uznania, ponieważ blok ten nie jest przeznaczony do wyszukiwania. Uwagi dotyczące tytułów mogą być generowane z bloku 5 - -.

Blok tytułów zawiera informacje o zależnościach między tytułem właściwym a tytułami pochodzącymi z innych miejsc dokumentu niż karta tytułowa. Jest tu także miejsce dla tytułu ujednoliconego, kluczowego, dodanego lub przetłumaczonego przez katalogującego i in. Blok przeznaczony jest do celów wyszukiwawczych.

Blok informacji o powiązaniach rekordu dotyczy takich spraw jak kontynuacja, wchłonięcie, tłumaczenie z ... lub na .. zależności hierarchiczne itp. Korzysta się tu z metody "osadzania pól" ("embedded fields") branych z innych rekordów, do których się odsyła zamieszczając także ich numer identyfikacyjny. Blok ten można jednak wypełniać nieco inaczej, zaznaczając tylko informacje o samych związkach opisu, a nie o powiązaniach między fizycznymi rekordami.

Blok opisu rzeczowego uwzględnia kilka języków informacyjno-wyszukiwawczych, mianowicie Klasyfikację Dziesiętną Deweya, Uniwersalną Klasyfikację Dziesiętną, PRECIS<sup>B/</sup>, Klasyfikację Biblioteki Kongresu, hasła przedmiotowe. Nie uwzględnia natomiast języków wykorzystywanych tylko w Polsce - np. Polskiej Klasyfikacji Tematycznej, Rubrykatora MSINT. Poza tym należy pamiętać o różni-

---

<sup>B/</sup> Preserved Context Indexing System (W. Brytania)

cach w stosowaniu języka haseł przedmiotowych, a także tzw. klasyfikacji własnych bezznakowych, powstaje więc problem częściowego dostosowania bloku do praktyki polskiej.

Blok informacji związanych ze strefą odpowiedzialności uwzględnia podział na pierwsze i następne oznaczenia. Bierze też pod uwagę najnowsze przepisy dotyczące haseł głównych jak i możliwość nie stosowania tych haseł w ogóle. Rodzaj odpowiedzialności (współpracy) oznaczony jest odpowiednimi kodami, których wykaz jest podany. Wydaje się jednak, że koncepcja bloku 7 - - kryje w sobie pewną niekonsekwencję. Otóż nazwa całego bloku (Intellectual Responsibility Block), zamieszczone w nim przykłady oraz ogólnie przyjęta zasada przestrzegania przepisów ISBD sugerują, iż blok ten powinien dotyczyć osób lub ciał korporatywnych biorących w różnym stopniu odpowiedzialność za intelektualną zawartość dokumentu. Jednakże z zastosowania innego nazewnictwa dla rodzajów odpowiedzialności niż w polu 200 (pole dotyczące tytułu oraz określenia odpowiedzialności w opisie bibliograficznym), bardzo obszernego wykazu kodów współpracy (np. dla adresata listów, dedykacji, wydawcy, właściciela prac wydawniczych, autora utworu stanowiącego podstawę dla dzieła innego autora opisywanego w rekordzie) jak i z komentarza do bloku 7 - - i pół wynika, że w bloku uwzględniono również inne osoby, ciała zbiorowe - takie, które są w różny sposób związane z dokumentem, niekoniecznie w sensie współudziału w jego tworzeniu. W takim ujęciu odpowiedzialność intelektualna rozumiana jest bardzo szeroko w zależności od zakładanych rozmaitych potrzeb wyszukiwawczych, zaspokajaniu których blok służy. Najistotniejszą sprawą jest określenie przez dany ośrodek narodowy rodzajów informacji mających stanowić przedmiot wyszukiwania, a w tym zakresie format nie stawia ograniczeń.

Blok zarezerwowany do użytku międzynarodowego w tej chwili dotyczy tylko ISDS.

Ostatni blok przeznaczony jest na informacje lokalne, jak np. sygnatura, akcesja, proveniencja, uwagi o egzemplarzu, nakład itp. Z kolei w przypadku czasopism można tu zamieścić informacje o zasobie. Blok wymaga opracowania przez poszczególne ośrodki krajowe; dzięki odpowiedniemu oprsowaniu bloku 9 - UNIMARC może pełnić również funkcję formatu katalogowego.

Obok struktury formatu i rodzajów zawartych w nim danych równie istotna jest forma ich zapisu. UNIMARC stosuje dwie formy - informacje zapisywane są w polach i podpolach oraz we wskaźniksch. Informacja podana we wskaźnikach jest zawsze zakodowana za pomocą znaku numerycznego, informacja w polach i podpolach może być różnie zakodowana lub tekstowa. Pola, podpola mogą być powtarzalne w zależności od potrzeb, natomiast wskaźniki danego pola są dwa i przeznacza się na nie po jednym znaku. Jeśli wskaźnik nie został zdefiniowany, wypełniany jest spacją, w innym przypadku znakiem "|", który ma również specyficzne zastosowanie w przypadku celowego nieuwzględnienia informacji zakodowanych w polach lub podpolach. W UNIMARC elementy opisu bibliograficznego zapisywane są w logicznej kolejności ustalonej przepisami, co znacznie ułatwia pracę wypełniającemu formularz.

Ów dogodny dla bibliografa sposób zapisu oraz pojemna i zarazem przejrzysta struktura formatu, umożliwiająca określenie wszystkich informacji o dokumencie, to dwie podstawowe cechy UNIMARC, które przemawiają za jego upowszechnianiem. Inną, nie mniej ważną cechą jest jego otwartość wobec przepisów, rozwiązań czy zwyczajów przyjmowanych w różnych krajach, o czym między innymi świadczy przyjęty sposób zapisu oraz mała liczba pól obowiązkowo-

wych<sup>9/</sup>, różne możliwości wypełniania bloków 3 - -, 5 - -, czy 7 - -, stosowanie znaku "|". Te cechy pozytywne w pewnym sensie przesłaniają jego główny mankament, mianowicie UNIMARC do chwili obecnej - jak zresztą wszystkie formaty wywodzące się z MARC - nie rozwiązał problemu opisu dokumentu wielotomowego (opisu wielopoziomowego)<sup>10/</sup>.

Wprowadzającą oraz opisową część podręcznika uzupełnia 14 załączników - wykazów, zestawów i tabel opatrzonych stosownym komentarzem:

A - Kody języków (3 znaki literowe),

B - Kody krajów (wykazu używanych dwuznakowych kodów literowych nie podano, odsyłając czytelnika do normy ISO 3166).

C - Kody rodzajów odpowiedzialności (związków osób i ciał korporacyjnych z dokumentem; 3 znaki numeryczne),

D - Kody obszarów geograficznych (7 znaków literowych i kresek).

E - Kody chronologiczne (wyrażające okres którego publikacja dotyczy; 4 znaki alfanumeryczne),

F - Kody geograficzne dla pola 131 dotyczącego materiałów kartograficznych i ich pomiarów (2 znaki literowe),

---

9/ UNIMARC przewiduje następujące pola obowiązkowe:

001 - numer identyfikacyjny rekordu

100 - ogólne dane dotyczące rekordu i jego zawartości

101 - język dokumentu (w zakresie podpól mających zastosowanie)

123 - informacje zakodowane o materiałach kartograficznych (skala, dane matematyczne)

200 - tytuł oraz oznaczenie odpowiedzialności (obowiązkowe tylko podpole \$ a - tytuł właściwy)

206 - strefa specyficzna dla materiałów kartograficznych - dane matematyczne

801 - dane dotyczące tworzenia rekordu

10/ Jak np. w Common Communication Format, gdzie zastosowano technikę łączenia tzw. segmentów (prostych opisów) w pionie i poziomie. (The CCF: Common Communication Format. Ed. by P. Simons and A. Hopkinson. Paris 1984).

(Przejęte wykazy kodów A, C, D i E zostały opracowane w Bibliotece Kongresu na użytek USMARC, zaś wykaz F jest stosowany dla potrzeb UKMARC).

G - Kody słowników i schematów klasyfikacyjnych zastosowanych przy przedmiotowym opracowaniu dokumentów.

H - Kody przepisów katalogowania przestrzeganych przy sporządzaniu opisów.

I - Tabela wartości pól, wskaźników oraz podpól.

J - Repertuar znaków służących do przedstawienia informacji w systemie przetwarzania danych (sposób kodowania informacji).

K - Wskazówki dotyczące dokumentacji towarzyszącej wymianie taśm.

L - Przykłady pełnego opisu w formacie. (Uwzględniają także elementy równoległe oraz powiązania między rekordami. Przykład 4 został nawet rozpisany w kształcie w jakim znajduje się na taśmie magnetycznej. Przykłady częściowe dotyczące stosowania poszczególnych pól są zawarte w części opisowej podręcznika).

M - Bibliografia (raczej wykaz norm międzynarodowych oraz innych dokumentów związanych z UNIMARC).

N - Adresy instytucji (wskazujące, gdzie można uzyskać potrzebne wydawnictwa i materiały).

Podręcznik posiada także indeks przedmiotowy obejmujący m.in. wszystkie elementy opisu (pól).

Wskazując na bogate wyposażenie podręcznika w różne pomoce nie sposób ominąć innego, chyba największego jego atutu, mianowicie jednolitej i klarownej wykładni formatu, opisu prawie wszystkich jego pól - począwszy od definicji pola i kończąc na zamieszczonych przykładach zapisu danych. Przykładów jest dużo, czasami - gdzie to dopuszczalne - uwzględniają one różne możliwości (wer-



sje) wypełniania pól przez ośrodki bibliograficzne. Zdarzają się tu jednak błędy drukarskie sprowadzające się najczęściej do pomyłek w oznaczeniu podpól.

UNIMARC Manual zasługuje na szersze omówienie i sam format na głębszą analizę (na tle innych formatów bibliograficznych), przeprowadzoną pod kątem jego przydatności jako formatu narodowego<sup>11/</sup>. To ważny temat bezpośrednio związany z tworzeniem krajowych zautomatyzowanych systemów biblioteczno-informacyjnych i współpracą międzynarodową, który powinien być na bieżąco penetrowany w polskiej literaturze fachowej.

15.05.1089 r.

THE INTERNATIONAL FORMAT FOR THE EXCHANGE  
OF BIBLIOGRAPHIC DATA

Summary

The author characterizes the works on UNIMARC format carried within the program of IFLA "Universal Bibliographic Control and International MARC (UBCIM). The history of working out of the format is presented. The structure and the content of the UNIMARC MANUAL is described. There are also discussed: the structure and

<sup>11/</sup> Zob. np. H. Vodičková: K otázce struktury jmenného záznamu. Knižnický Zborník 1987 s.34-67. Opracowanie jest częstkową analizą porównawczą struktury identyfikacyjnego opisu dokumentów stosowanej w ISBD/G/, NTP MCINT, Reference Manual UNESCO, CCF, normie i formatach używanych w CSRS oraz radzieckiej normie GOST 7.1-84

the content of the UNIMARC format (the label of the record, data blocs), the form of the recording of data and the appendices. The advantages of the discussed format are indicated.

**МЕЖДУНАРОДНЫЙ ФОРМАТ ДЛЯ ОБМЕНА БИБЛИОГРАФИЧЕСКИМИ  
ДАНЫМИ**

**Резюме**

Автор характеризует работы по созданию формата UNIMARC проводимые в рамках IFLA "Universal Bibliographic Control and International MARC (UBCIM). Представлена история разработки формата. Описаны структура и содержание пособия по формату UNIMARC (маркер участка, блоки информации), форма записи данных и приложения. Указаны преимущества обсуждаемого формата.



JADWIGA SADOWSKA

Biblioteka Narodowa

#### UWAGI NA TEMAT FORMATU POLMARC

Organizacja prac nad formatem USMARC. Międzynarodowy format wymienny UNIMARC. Uwagi na temat formatu Biblioteki Narodowej MARC-BN i prac nad formatem POLMARC.

Format opisu dokumentu jest pewnego rodzaju strukturą umożliwiającą zapisanie informacji o dokumencie w taki sposób, aby były one czytelne dla programów komputerowych. Problem ten pojawił się wraz z wprowadzaniem do bibliotek nowej techniki, w latach sześćdziesiątych w Stanach Zjednoczonych, Kanadzie, Wielkiej Brytanii, RFN, nieco później, bo w latach siedemdziesiątych w innych krajach, między innymi w Polsce.

Pierwszym, a jednocześnie najbardziej znanym i rozpowszechnionym, był format opracowany w drugiej połowie lat sześćdziesiątych przez Bibliotekę Kongresu w Waszyngtonie - LC MARC (MACHINE READABLE CATALOGUING LIBRARY OF CONGRESS). Od tego czasu powstało wiele odmian narodowych i obecnie termin MARC należy traktować jako ter-

min rodzajowy wobec takich terminów jak UK MARC, CANMARC, UNIMARC i innych. Ich związek z formatem MARC często zaznaczany jest już w samej nazwie: AUSMARC (Australia), CANMARC (Kanada), danMARC (Dania), UK MARC (Wielka Brytania), NORMARC (Norwegia), SWEMARC (Szwecja). (Mamy nadzieję, że do tej grupy nazewniczej dołączy także POLMARC).

Oprócz odmian narodowych powstał także UNIMARC- Universal MARC Format przygotowywany od połowy lat siedemdziesiątych pod auspicjami IFLA. Trzeba powiedzieć, że o ile początkowo wzorem dla formatów narodowych był format amerykański, o tyle obecnie staje się nim UNIMARC. Według danych zawartych w "International Guide to MARC Databases and Services" jest on podstawą dla formatów: włoskiego, japońskiego, tajwańskiego, chińskiego, węgierskiego, polskiego (MARC-BN). Jest oczywiste, że ci, którzy opracowywali swoje formaty narodowe wcześniej, wzorowali się na formacie Biblioteki Kongresu (jest to bardzo widoczne w formacie angielskim i kanadyjskim), ci, którzy robili to później brali pod uwagę UNIMARC. Wiele też bibliotek narodowych, w tym także takich krajów jak Australia, Kanada, Wielka Brytania, Stany Zjednoczone zgodziło się stosować UNIMARC jako format wymienny. To zadecydowało o uznaniu za priorytetowe prac nad konwersją formatów narodowych na UNIMARC. Niezależnie od tego od końca lat siedemdziesiątych prowadzone są, pod auspicjami UNESCO, prace nad Common Communication Format<sup>1/</sup>. Wszystko jednak wskazuje na to, że międzynarodowym formatem wymiennym, a także wzorcem dla formatów narodowych ma szansę stać się UNIMARC.

---

<sup>1/</sup> zob. Common Communication Format. Paris 1988; P. Murasik: Uniwersalny format wymiany danych bibliograficznych. ZIN 1988 nr 2(53) s.165-186.

## 1. USMARC - format Biblioteki Kongresu

W pracach nad formatem Biblioteki Kongresu wydziela się trzy etapy<sup>2/</sup>:

1966 - 1968 MARC Pilot Project, zwany inaczej MARC I,

1968 - 1974 zmodyfikowany LC MARC, znany jako MARC II

1974 - wersja określana jako USMARC.

Od początku okres pierwszy traktowany był jako pilotażowy i rozpoznawczy. Był to właściwie eksperyment przeprowadzony w Bibliotece Kongresu i kilkunastu innych bibliotekach. W 1968 r. do prac włączył się Oddział Informacji i Automatyzacji Amerykańskiego Stowarzyszenia Bibliotekarzy (Information Science and Automation Division), podejmując szereg przedsięwzięć związanych z wdrażaniem MARC II. W 1969 r. rozpoczęto dystrybucję opisów książek angielskojęzycznych - około 1000 rekordów tygodniowo, włączając kolejno książki w językach alfabetu łacińskiego. W 1970 roku opublikowano format dla czasopism i map, w 1971 dla filmów, w 1973 dla rękopisów i druków muzycznych. Obecnie formatów jest 7: dla książek (BK), filmów (FI), rękopisów (MS), map (MP), druków muzycznych (MU), czasopism (SE), zbiorów danych (DF - Machine Readable Data Files Format).

Prace nad formatami amerykańskimi wpłynęły na prace normalizacyjne. Struktura MARC została przyjęta jako norma narodowa i międzynarodowa. W 1973 r. opublikowano normę ISD 2709, która była odpowiednikiem normy amerykańskiej ANSI Z 39.2.

W 1974 r. w ramach Amerykańskiego Stowarzyszenia Bibliotekarzy (American Library Association - ALA) powstał komitet, znany

---

<sup>2/</sup> zob. MARC for library use: understanding the USMARC formats. White Plains 1984.

obecnie jako MARBI (Committee on Representation in Machine-Readable Form of Bibliographic Information). Komitet ten, spełniający rolę ciała doradczego, jest reprezentowany przez przedstawicieli kilku oddziałów ALA.

Formaty są publikowane przez Bibliotekę Kongresu i ona jest odpowiedzialna za ich ostateczny kształt. Jednakże nie ona sama decyduje o wprowadzanych zmianach. Są one dyskutowane na otwartych sesjach, w których poza Biblioteką Kongresu i komitetem MARBI, biorą udział przedstawiciele Narodowej Biblioteki Medycznej, Narodowej Biblioteki Rolniczej, Biblioteki Narodowej w Ottawie oraz przedstawiciele służb bibliograficznych innych zainteresowanych bibliotek i instytucji. Zgłaszane propozycje są wcześniej weryfikowane przez Bibliotekę Kongresu. Ona też przygotowuje wstępne propozycje rozwiązań, które następnie są dyskutowane w gronie przedstawicieli wymienionych wcześniej instytucji. Każdego roku odbywają się 4 sesje przeglądowe. Zaakceptowane zmiany są drukowane i rozsyłane przez Bibliotekę Kongresu do zainteresowanych.

Tak w ogólnym zarysie wygląda procedura aktualizacji formatów amerykańskich. Jak widać przywiązuje się do tego wielką wagę. Prace prowadzone są w sposób systematyczny. Wiadomo, kto jest odpowiedzialny, do kogo należy zgłaszać propozycje, od kogo należy oczekiwać pomocy metodycznej. Nadanie odpowiednich ram organizacyjnych pracom nad formatem USMARC, sprawia, że zainteresowanie nim nie słabnie, a korzystający z niego nie tylko mogą liczyć na format, ale także na omówienia i podręczniki metodyczne.

## 2. UNIMARC - Universal MARC Format<sup>3/</sup>

UNIMARC pomyślany jest jako międzynarodowy format wymienny przeznaczony do opisu książek, czasopism, map, druków muzycznych i materiałów audiowizualnych (non-book materials). Może on być także podstawą dla formatów narodowych. Trzeba jednak pamiętać, że format ten uwzględnia przede wszystkim to, co jest ważne z punktu widzenia potrzeb wymiany informacji. Nie uwzględnia i nie podaje sposobów wyrażenia danych lokalnych, co między innymi odróżnia go od formatów narodowych. Przewiduje natomiast na tego typu informacje blok ostatni, ale bez podania jakichkolwiek rozwiązań. W tym sensie UNIMARC jest uboższy w stosunku do USMARC i podobnych.

UNIMARC opracowany jest przez grupę roboczą, której koordynatorem jest International Federation of Library Associations and Institutions (IFLA). W skład tej grupy wchodzi głównie przedstawiciele bibliotek narodowych, a przewodniczy jej Henriette Avram z Biblioteki Kongresu, autorka pierwszej wersji formatu amerykańskiego i aktywna uczestniczka wszelkich prac związanych z USMARC. Daje to w jakimś stopniu gwarancję, że zachowany jest związek obu formatów, a jednocześnie, że UNIMARC wolny jest od błędów i niedociągnięć formatu amerykańskiego, który od lat sześćdziesiątych był wiele razy poprawiany i usprawniany. Trzeba też powiedzieć, że w ciągu 10 lat UNIMARC był trzy razy wydawany: I wyd.- 1977, II wyd. - 1980, III wyd. - 1987. Przy czym wydanie ostatnie jest połączeniem formatu z podręcznikiem, co znakomicie ułatwia jego wykorzystywanie. Dodajmy, że UNIMARC uwzględnia przepisy ISBD/M/, ISBD/S/, ISBD/CM/, ISBD/PM/, ISBD/NBM/.

---

<sup>3/</sup> Ostatnie wydanie: UNIMARC manual. London 1987



Porównując bardzo pobieżnie UNIMARC, USMARC oraz pokrewne zauważyć można różnice w rozkładzie informacji w poszczególnych blokach oraz różnice w podejściu do niektórych informacji. Tak na przykład informacje związane z opisem rzeczowym UNIMARC grupuje w jednym bloku - i jest to wygodniejsze. Nieco inaczej traktowana jest strefa odpowiedzialności oraz tytuły. Wydaje się, że sposób przyjęty w UNIMARC jest bardziej przejrzysty (dotyczy to zwłaszcza tytułów) i prostszy. UNIMARC stosuje dla różnych typów tytułów osobne pola, USMARC dla części pola, a dla części wskaźniki. W formacie UNIMARC zrezygnowano z tzw. haseł głównych zamieszczonych w bloku 100 - (występują one w bloku 7--). Nie zamierzamy tu dokonywać szczegółowego porównania, ponieważ nie jest to naszym celem, wydaje się jednak, że można uznać format UNIMARC za bardziej przejrzysty niż np. USMARC. Być może na wrażenie to wpływa postać edycyjna obu formatów, korzystniejsza w publikacjach poświęconych formatowi UNIMARC. Pamiętajmy jednak, że UNIMARC jest tylko formatem bibliograficznym i cała strefa informacji własnych, lokalnych wymaga dostosowania do potrzeb narodowych. Niewątpliwie atutem UNIMARC jest też stała opieka ze strony IFLA.

### 3. MARC-BN

Już na początku lat siedemdziesiątych zainteresowano się w Polsce amerykańskim formatem MARC. w 1972 roku zaadaptowano go w Bibliotece i Ośrodku INTE Politechniki Wrocławskiej. Format ten znany był jako APIN/MARC (Automatyczne Przetwarzanie Informacji Naukowej). Pojawiły się też opracowania mające charakter uproszczonych tłumaczeń, m.in. w Ośrodku Przetwarzania Danych Biblioteki Narodowej została opublikowana praca J. Dody-Szargut: "Analiza formatu MARC LC opisu wydawnictw ciągłych". Bez wątpienia w owym czasie

była to znacząca pozycja, przybliżająca czytelnikowi polskiemu istotę formatu MARC.

Bardziej systematyczne prace nad formatem podjęto w Bibliotece Narodowej na przełomie lat siedemdziesiątych i osiemdziesiątych, a efektem ich były dwa formaty - dla wydawnictw zwartych i ciągłych opracowane przez Z. Moszczyńską-Pętkowską<sup>4/</sup>. Były one zgodne z normą ISO 2709. Formaty te nazwane MARC-BN zostały przygotowane z myślą o bibliografii narodowej, a ich związek z Biblioteką Narodową podkreślony był w samej nazwie.

Wzorem dla MARC-BN był UNIMARC i powiązania te są wyraźne, aczkolwiek widać też pewne wpływy innych formatów, a także rozwiązania własne. Nie wnikając w szczegóły można powiedzieć, że ogólny rozkład bloków informacyjnych jest taki jak w UNIMARC, ale wprowadzono tak zwany blok haseł głównych tak jak w USMARC, UK MARC, CANMARC. Z kolei wydaje się, że przyjęto własne rozwiązania jeśli idzie o wskaźniki, inaczej rozwiązano blok informacji rzeczowych, zwłaszcza opisu przedmiotowego.

Od 1986 r. według formatu MARC-BN opisywane są dokumenty publikowane w Przewodniku Bibliograficznym, a także czasopisma w katalogu alfabetycznym Biblioteki Narodowej. Podczas wdrażania format ten podlegał pewnym modyfikacjom, co jest sprawą zupełnie zrozumiałą. Szkoda tylko, że nie został on w odpowiednim czasie zweryfikowany, dopracowany i upowszechniony w innych bibliotekach. Zaniedbania w tej sprawie sprawiły, że poszczególne środowiska wraz z postępującym zaopatrzeniem bibliotek w sprzęt, przystąpiły do opracowywania własnych formatów, uznając, że format MARC-BN nie w pełni za-

<sup>4/</sup> Z. Moszczyńska-Pętkowska: System SABINA. Opis wydawnictw zwartych. Warszawa BN 1982, System SABINA. Opis wydawnictw ciągłych. Warszawa BN 1983.

spokaja ich potrzeby. Nie bez znaczenia był tu brak jakichkolwiek ram organizacyjnych związanych z aktualizacją formatu, jego rozpowszechnianiem, szkoleniem w zakresie wdrażania, a nawet pewnego rodzaju promocją.

Podkreśliśmy jeszcze raz, że format ten został wdrożony, choć prawdą jest też, że podczas wdrażania ujawniły się pewne jego niedociągnięcia, zwłaszcza związane z opisem dzieł wielotomowych (jest to jednak ogólny mankament formatu MARC), ale to nie może nikogo dziwić. Dziwić może natomiast jakby programowe ograniczenie jego stosowania do bibliografii narodowej z jednej strony, z drugiej zaś niewielkie zainteresowanie nim środowiska. Przyczyny tego stanu rzeczy na pewno są złożone i nie podejmujemy się tu ich wyjaśniania i dociekania.

Oceny żadnego formatu nie można przeprowadzać na poziomie ogólnego porównania podziału na bloki informacyjne, choć jest to na pewno pierwszy element porównania. Istotne jest to, co kryje się w poszczególnych polach i pdpolach, jak wyrażana jest ich wartość. Ważne jest, czy format stwarza możliwość jednoznacznego wyrażania informacji, łatwego dostępu do nich i prostego wykorzystania, czy nie powieliła informacji bez uzasadnionej potrzeby itd. Tak więc format to nie tylko sprawa bibliotekarza czy bibliografa, ale także informatyka i ich wzajemnego współdziałania.

#### Podsumowanie

1. Należy zwrócić uwagę na fakt, że zarówno na USMARC jak i UNIMARC składają się kilka formatów, ale nie m wśród nich osobnego formatu dla artykułu z czasopisma, fragmentu książki, poszczególnych map, melodii itp. Tego typu dokumenty opisywane są za pomocą innych formatów, tzn. przystosowania innych formatów.

2. Problemem w formatach typu MARC jest opis dokumentów wielo-  
poziomowych. Typowymi przykładami są książki wielotomowe i czaso-  
pisma z opisem tzw. zasobu, to jest poszczególnych roczników, ze-  
szytów itp., na co należy zwrócić szczególną uwagę opracowując for-  
maty w Polsce.

3. Wzorem dla formatu polskiego powinien być UNIMARC w za-  
kresie opisu bibliograficznego dokumentów i USMARC w zakresie da-  
nych katalogowych i innych lokalnych.

4. Format ten powinien być formatem uniwersalnym, nadającym  
się zarówno dla wydawnictw bibliograficznych, jak i różnych typów  
katalogów, w tym także katalogu centralnego.

5. Formaty powinny być opracowywane w sposób zintegrowany.  
Nie oznacza to wcale równoczesnego opracowywania, ale format dla  
każdego typu dokumentu musi być umieszczony w kontekście pozosta-  
łych formatów. Zapewnia to spójność, która jest widoczna zarówno  
w USMARC jak i UNIMARC, jak też jednolitą interpretację informacji  
w poszczególnych polach i podpolach.

6. Format narodowy powinien być skorelowany z formatem wy-  
miennym.

7. Trzeba mieć świadomość, że format jest strukturą stabilną,  
niemniej jednak podlegającą pewnym zmianom, wynikającym choćby  
ze zmian w opisie bibliograficznym. Dlatego musi mieć zapewnioną  
aktualizację. W żadnym razie nie wolno formatu traktować jako  
przedsięwzięcia jednorazowego o formie stałej i zamkniętej.  
Zmiany w formacie muszą uwzględniać zmiany przepisów związanych  
z opisem bibliograficznym dokumentów. Nie można jednak formatu  
dostosowywać do pojedynczych przypadków (opisów nietypowych),  
zwłaszcza że może się zdarzyć, że owa nietypowość opisów jest  
wynikiem niejednoznacznej interpretacji normy, a tym bardziej

interpretacji przepisów funkcjonujących w sposób zwyczajowy. Dotyczy to tych dokumentów, dla których nie ma polskich norm.

8. Stan formatu powinien być obiektem stałego zainteresowania i okresowych przeglądów z udziałem wielu użytkowników, bo tylko to może zapewnić utrzymanie jego jednolitego kształtu we wszystkich bibliotekach polskich. Co pewien czas format powinien być publikowany.

9. Przed upowszechnieniem format powinien być wszechstronnie przetestowany, a następnie wspomagany przez odpowiednie opracowania metodyczne i popularyzacyjne.

Przedstawione tu uwagi należy traktować jako głos w dyskusji nad sprawą formatów w Polsce. W ostatnim czasie w ramach Resortowego Programu Badawczo-Rozwojowego powierzono opracowanie formatów trzem instytucjom: Bibliotece Uniwersytetu Warszawskiego - format katalogowy dla książek i czasopism (a więc bez uwzględnienia potrzeb bibliografii narodowej), OIN PAN - format dla artykułów, Biblioteka Narodowej - format wymienny.

Należy mieć nadzieję, że prace te nie zostaną wykonane, jak wiele innych, metodą współzawodnictwa, ale we współpracy, która jest konieczna przy tego rodzaju przedsięwzięciach.

15.03.1989 r.

## REMARKS ON THE POLMARC FORMAT

### Summary

The organization of works on the POLMARC format is presented. Some characteristics of the international exchange format UNIMARC is given. The author formulates the remarks concerning the format of the Polish National Library MARC BN and the works on POLMARC.

## ЗАМЕЧАНИЯ НА ФОРМАТ POLMARC

### Резюме

Представлена организация разработки формата данных POLMARC. Приведена характеристика международного коммуникативного формата данных UNIMARC. Автор формулирует замечания относящиеся к формату польской Национальной библиотеки MARC BN и связанным с POLMARC работам.



## RECENZJE I OMÓWIENIA

### Z DOŚWIADCZEŃ KRAKOWSKIEJ PRACOWNI OIN PAN NAD TWORZENIEM SŁOWNIKÓW SŁÓW KLUCZOWYCH NAUK SPOŁECZNYCH

W 1988 r. Pracownia OIN PAN w Krakowie opublikowała pięć słowników słów kluczowych z następujących dyscyplin nauk społecznych: ekonomii politycznej, filozofii, językoznawstwa, pedagogiki i psychologii.\*

Wydane przez OIN PAN słowniki liczą łącznie 9125 słów. W rozbiciu na poszczególne dyscypliny wygląda to następująco:

- ekonomia polityczna	2124
- filozofia	1037
- językoznawstwo	1572
- pedagogika	2006
- psychologia	2386

W druku znajdują się trzy słowniki. A mianowicie słów kluczowych z: marksizmu-leninizmu, nauki o polityce (politologii) i socjologii. W trakcie opracowywania są też trzy następne, to jest słowniki dla historii, literaturoznawstwa i religioznawstwa.

Wszystkie wydane słowniki zawierają: wstęp od wydawcy, mówiący o istocie i przeznaczeniu słowników. wstęp metodologiczny autorów,

\*Słownik słów kluczowych z ekonomii politycznej. Kraków 1988 Prace OIN PAN 47 s.; Słownik słów kluczowych z filozofii. Kraków 1988 Prace OIN PAN 28 s.; Słownik słów kluczowych z językoznawstwa. Kraków 1988 Prace OIN PAN 36 s.; Słownik słów kluczowych z pedagogiki. Kraków 1988 Prace OIN PAN 60 s.; Słownik słów kluczowych z psychologii. Kraków 1988 Prace OIN PAN 46 s.



wykaz najważniejszych pozycji bibliograficznych, z którego czerpano słownictwo oraz wykaz słów kluczowych ułożonych w porządku alfabetycznym. Prócz wymienionych części niektóre słowniki zawierają (w postaci aneksów): wykazy nazwisk wybitnych przedstawicieli danej nauki, wykazy instytucji, organizacji, stowarzyszeń i wydarzeń historycznych związanych z daną dyscypliną, a także tytuły związanych z tą dyscypliną czasopism naukowych.

Jeśli idzie o formę samych słów kluczowych, to są to słowa jedno lub wielowyrazowe, z tym, że rzeczownik z reguły występuje w pierwszym przypadku liczby pojedynczej, np. **Byt** (filozofia). W liczbie mnogiej występują tylko te słowa, których użycie w liczbie pojedynczej zmianałoby ich właściwy sens (np. w filozofii **Cynik** nie znaczy to samo co **Cynicy**), lub gdy dane słowo z zasady występuje w liczbie mnogiej. Np. **Granice poznania**, **Kategorie filozoficzne** (w filozofii), **Siły Wytwórcze**, **Stosunki Produkcji** (w ekonomii) czy **Stosunki Międzyludzkie**, **Stosunki Międzynarodowe**, **Swobody Polityczne i Wybory** (w politologii i socjologii).

W słowach wielowyrazowych rzeczownik z reguły występuje na pierwszym miejscu, po nim zaś przymiotnik lub inne wyrazy. W nielicznych jednak przypadkach, tam gdzie wymagał tego naturalny szyk języka polskiego, przymiotnik wystąpił na pierwszym miejscu. Np. **Amerykańska szkoła strukturalistyczna**, **Psychologiczna realność reguł gramatycznych** (językoznawstwo), **Nierównomierny wzrost gospodarczy**, **Niezbędny czas pracy**, **Dwupoziomowy system cen** (ekonomia), **Liniowy układ programu** (pedagogika).

Spośród wielowyrazowych słów są to:

- słowa dwuwyrzowe, składające się z rzeczownika i przymiotnika, np. **Cecha psychiczna** (psychologia), **Byt absolutny** (filozofia) oraz dwa rzeczowniki, pierwszy w mianowniku, drugi w dopełniaczu, np. **Akumulacja kapitału** (w ekonomii). **Deprywacja snu** (psychologia),

- trzywyrazowe, składające się z dwu przymiotników i rzeczownika, lub odwrotnie, dwu rzeczowników i przymiotnika. Np. **System dydaktyczno-wychowawczy**, **Techniczne środki dydaktyczne** (pedagogika), **Dziecko niedostosowane społecznie** (psychologia), oraz **Dciążenie funkcjonalne fonemu** (językoznawstwo), **Bilans gospodarki narodowej** (ekonomia), **Automatyzm czynności ruchowych** (psychologia), **Zasada racji dostatecznej** (filozofia).

Niektóre trzywyrazowe słowa zawierały także przyimki i przysłówki. Np.: **Rzecz w sobie**, **Harmonia wprzód ustanowiona** (filozofia). **Praca z książką**, **Nauka o rzeczach**, **Uczenie się przez przyswajanie** (pedagogika).

Czterowyrazowe słowa i liczniejsze (pięcio- i sześciowyrazowe) słowa są rzadsze, niemniej występują w słownikach. Np. **Dowody na istnienie Boga** (filozofia), **Test rozwoju percepcji wzrokowej Mariany Frostig** (psychologia). Najwięcej występuje ich jednak w ekonomii, głównie w postaci nazw praw ekonomicznych (bp.: **Prawo wartości siły roboczej w kapitalizmie**, **Prawo nierównomiernego rozwoju ekonomicznego i politycznego**, **Prawo planowego i perspektywicznego rozwoju gospodarki socjalistycznej**). Jak więc widać, prócz rzeczowników i przymiotników są w tych słowach kluczowych również spójniki. Ale spójnik stanowi tu integralną część nazwy danego prawa.

W słownikach najwięcej jest słów dwuwyrazowych, składających się z rzeczownika i przymiotnika. Szacunkowe obliczenia upoważniają także do postawienia tezy, że w naukach starszych, w których systematyka wewnętrzna jest już dość uporządkowana, a kategorie pojęciowe bardziej utrwalone (np. w filozofii), występuje więcej słów jednowyrazowych; wielowyrazowe są nieliczne. W naukach młodszych słów wielowyrazowych jest znacznie więcej.

Wszystkie słowniki zawierają relację synonimii (tożsamości znaczenia różnych nazw), która oznaczona jest odsyłaczem całkowi-

tym (zwanym też bezwzględny) "zobacz". Zastosowano tu bardzo konsekwentnie zasadę symetrii tejsze synonimii, tzn. powrotu do nazwy wyjściowej w układzie alfabetycznym. Z tym jednakże, że odpowiednią grubością czcionki zaznaczone zostały słowa preferowane, a więc poprawniejsze z punktu widzenia ich naukowości, nowoczesności i częstotliwości występowania. Np. w filozofii za poprawiejszy uznano: **Byt absolutny**, a nie Absolut, **Historiozofia**, a nie Filozofia historii, **Uniwersalia**, a nie Powszechniki, w ekonomii: **Przemyt**, a nie Kontrabanda, **Akumulacja**, a nie Nagromadzenie, **Tendencja**, a nie Trend. W psychologii są to **Demencja**, a nie Ostępnie, **Urojenie**, a nie Deluzja, **Kara**, a nie Wzmocnienie negatywne, **Padaczka**, a nie Epilepsja. W pedagogice: **Uogólnienie**, a nie Generalizacja, **Nauczanie**, a nie Uczenie, **Samokontrola**, a nie Autokontrola.

W omawianych słownikach preferencja przy synonimach nie jest całkowita. A ściślej mówiąc nie została ona narzucona czy wymuszona w stosunku do tych terminów, które w danej nauce występują równie często i są równie poprawne pod względem naukowym. A oto kilka przykładów słów synonimicznych, w których nie zastosowano preferencji, uznając je za równorzędne. W psychologii: **Oligofrenia** i **Wiedorozwój umysłowy**, **Psychologia rozwoju człowieka** i **Psychologia biodrowalna** oraz **Life span psychology**, **Leukotomia** i **Lebotomia**, **Rozwój umysłowy** i **Rozwój intelektualny**, **Trening** i **Ćwiczenie**. W pedagogice: **Ustrój szkolny** i **System szkolny**, w filozofii: **Aksjologia** i **Teoria Wartości**, **Antropologia filozoficzna** i **Filozofia człowieka**, **Porządek i ład**, **Esencja** i **Istota**, w ekonomii: **Polityka ekonomiczna** i **polityka gospodarcza**, **Prosperita** i **Boom** oraz **Rozkwit gospodarczy**, **Przychody** i **Aktywa**.

Prócz omawianej tu synonimii w słownikach występuje też relacja pokrewieństwa, którą oznaczono odsyłaczem częściowym (względny)

"zobacz też". Np. w filozofii: **Absolutyzm**, zob. też **Relatywizm**, **Akt** zob. też. **Możliwość**, **Dusza** zob. też **Niesmiertelność**. W ekonomii: **Burżuazja** zob. też **Mieszczanństwo**, **Giełda pracy** zob. też **Pośrednictwo pracy**, **Czas pracy** zob. też **Praca**. W pedagogice: **Blokowy układ programu** zob. też **Nauczanie programowane**, **Ideał wychowania** zob. też **Pajdocentryzm**. W psychologii: **Akceptacja społeczna** zob. też **Przystosowanie społeczne**, **Amnezja** zob. też **Luka pamięciowa**, **Antropomorfizm** zob. też **Egocentryzm dziecięcy**.

W niektórych słownikach łączono za pomocą tego samego odsyłacza zobacz też (jako relacje kojarzeniowe) antonimy, a więc przeciwieństwa. Np. w filozofii: **Amoralizm** zob. też **Moralizm**, **Bezwzględność** zob. też. **Względność**, **Byt** zob. też **Nicość**, **Egoizm** zob. też **Altruizm**. W ekonomii: **Eksport** zob. też **Import**, **Bessa** zob. też **Hossa**.

Przy homonimach, a więc słowach wieloznacznych, (np. **Depresja gospodarcza** i **Depresja psychiczna**), których było jednak bardzo niewiele, właściwe znaczenie słowa podane zostało w okrągłym nawiasie. W takim samym nawiasie zamieszczano dookreślenia, to znaczy wyrazy precyzujące znaczenie, rodzaj i datę, a czasem i wywodzenie się (od określonej szkoły czy określonego uczonego) danego słowa. Np. w pedagogice słowo **Biologia** raz ma dookreślenie: nauka pomocnicza pedagogiki, drugi raz: przedmiot nauczania, słowo **Pomiar** i **Pokaz** mają dookreślenie: metoda nauczania, słowo **Bracia Szkół Chrześcijańskich** mają dookreślenie: XVII/XVIII wiek. W językoznawstwie przy wielu słowach występuje w nawiasie dookreślenie: gramatyka generatywna bądź gramatyka generatywno-transformacyjna.

Jeśli idzie jeszcze o formę słów kluczowych, to zasadą jest, iż występują one w wersji polskojęzycznej, zaś nazwy w ich pełnym brzmieniu. Skróty, jako preferowane słowa kluczowe występują tylko w przypadkach ich wielkiego upowszechnienia (np. ONZ, RWPG). Niemniej jednak od tych skrótów występuje odsyłacz całkowity zob. do

ich pełnej nazwy. W psychologii zamieszczono skrót I.I. od Iloraz inteligencji, ale nie jako słowo preferowane.

Obce nazwy zostawione zostały tylko wtedy, gdy były nieprzekładane na język polski, tzn. nie mają do tej pory adekwatnego do ich pełnego znaczenia polskojęzycznego odpowiednika. Na przykład w psychologii nie znaleziono odpowiednika do francuskich słów *Déjà vu* (już widziane), *Déjà vecu* (już przeżyte), *Jamais vu* (nigdy nie widziane). Słowo *Folie à deux* (szaleństwo we dwoje) ma dwa polskie ekwiwalenty. Mianowicie *Obłąd indukowany* i *poprawniejszy*, to jest preferowany przez autorów *Psychoza indukowana*. Angielskie terminy *Brainstorming* i *Brainwashing* mają polskie odpowiedniki jako: *Burza mózgow* i *Pranie mózgow*, ale z preferencją ich angielskojęzycznej wersji. Podobnie przy *Biofeedback* dano ekwiwalent *Sprzężenie zwrotne*, ale z preferencją tego pierwszego słowa.

W ekonomii zostawiono bez polskiego ekwiwalentu słowo *Marketing* oraz *Free alongside* i *Free on board*, których skróty: *FAS* i *FAB* uznane zostały za słowa preferowane. Słowo *Joint ventures* ma ekwiwalent: *Przedsięwzięcia wspólne*, ale preferencję zatrzymano dla tego pierwszego terminu. Podobnie jest ze słowem *Terms of trade*, które ma ekwiwalent w postaci: *Warunki handlu zagranicznego*, ale preferencję przyznano słowu angielskiemu. W filozofii bez przekładu na język polski pozostawiono słowa: *A posteriori* i *A priori*. O preferencji wszystkich wymienionych tu słów obcojęzycznych zadecydowała powszechność ich stosowania w danej dyscyplinie.

Wydane przez OIN PAN słowniki - w części dotyczącej wyboru podstawowych terminów - opracowała średnia kadra naukowa (adiunkcji), recenzowane zaś były i konsultowane przez samodzielnych pracowników nauki (przeważnie profesorów). Przy czym recenzje, o których tu mowa, dotyczyły poprawności słownictwa zarówno w jego aspekcie merytorycznym jak i też formalnym, to znaczy rygorów języka informa-

cyjnego. Tę drugą ocenę sporządzali pracownicy nauki Instytutu Informacji Naukowo-Technicznej L. A. Bielicka i E. Ścibor.

Przed przystąpieniem do prac nad słownikami Pracownia OIN PAN w Krakowie opracowała ogólną instrukcję wyboru słów kluczowych w naukach społecznych. Instrukcja ta określała cel naszej pracy, definiowała słowa kluczowe dając zarazem ich konkretne przykłady, wskazywała źródła, z których należy czerpać słownictwo<sup>1/</sup>, a także sposoby sprawdzania jego kompletności, to jest unikania "luk". Prócz tego, dla opracowujących wykazy słów kluczowych zorganizowane zostało seminarium naukowe z udziałem specjalistów od języków informacyjnych (mgr L.A. Bielicka, dr W. Pindłowa, doc.dr hab. E. Ścibor), dostarczono im też z tego zakresu literaturę naukową i istniejące (przeważnie zachodnie) tezaury dydziedzinowe. W rozmowach bezpośrednich konsultowano także robocze wersje słowników.

Wskazówki dotyczące merytorycznych kryteriów wyboru słów, które mają się znaleźć w słownikach słów kluczowych, były następujące:

- mają być to słowa ważne, charakterystyczne dla danej dyscypliny naukowej; obejmujące cały jej obszar badawczy wraz z metodologią badań;

- terminy naukowe, a nie potoczne;
- poprawne pod względem językowym;
- aktualne, a nie przestarzałe;
- polskojęzyczne, z wyjątkiem tych obcych słów, które nie mają jeszcze polskich ekwiwalentów;
- powszechnie stosowane;

---

<sup>1/</sup> Wspomagaliśmy się tu m.in. publikacją: M. Jasińska, H. Makowa: Polskie źródła informacji w naukach społecznych i humanistycznych. Warszawa 1983. Biblioteka Narodowa. Zakład Informacji Naukowej.

- o znaczeniu raczej szerszym niż węższym;
- mające znaczenie wyraźne, a nie intuicyjne;
- przydatne przy opisywaniu treści dokumentów i ich wyszukiwaniu, a więc mające wartość praktyczną dla określonego systemu informacyjnego.

Liczbę terminów pozostawiono do rozstrzygnięcia osobom sporządzającym ich wykaz.

Źródła zasobu leksykalnego dla słowników typowane były następująco:

- podstawowa literatura przedmiotu: podręczniki, monografie, rozprawy naukowe;
- słowniki i encyklopedie dziedzinowe;
- katalogi rzeczowe;
- podstawowe dla danej dyscypliny dokumenty źródłowe;
- indeksy bibliografii;
- własna wiedza o przedmiocie, dotycząca w szczególności najnowszych osiągnięć naukowych, które nie znalazły jeszcze odbicia w słownikach, encyklopediach itp.

Mimo tego, zdawałoby się bardzo starannego i szczegółowego instruktażu, pierwsze wykazy słów kluczowych, jakie otrzymaliśmy, były dość dalekie od wymagań stawianych słownikom słów kluczowych. Różniły się też one między sobą strukturą wewnętrzną. Część dyscyplin miała układ słów nie alfabetyczny, lecz gniazdowy (pojęcia szerokie i wchodzące w jego skład węższe), część miała strukturę mieszaną (układ alfabetyczny i wmontowane w niego tu i ówdzie elementy układu hierarchicznego), a jedna dyscyplina miała nawet strukturę (oczywiście niekonsekwentną) UKD. Poza tym niektóre słowa (np. Amerykańska droga budowy kapitalizmu w rolnictwie) przybierały postać całych zdań, inne zaś (wielowyrzowe) miały nienaturalny szyk (np. Cału definicja zamiast Definicja celu).

To mieszanie przez autorów wykazów układu alfabetycznego i hierarchicznego, a także obudowywanie niektórych słów licznymi określnikami i kwalifikatorami nasuwa wniosek, że w naukach społecznych jest wiele takich słów, których "ostrość" i znaczenie staje się w pełni zrozumiałe dopiero w szerszym kontekście słownym, oraz że są takie słowa, których sens łatwiej pokazać poprzez relacje z innymi - jako część i całość, odmiany, itp.

Jeśli idzie o sposób wyboru słów, to - jak wiadomo - wśród nieautomatyzowanych sposobów istnieją dwa zasadnicze. Mianowicie wybór z dokumentów pierwotnych (źródeł), opracowań monograficznych, rozpraw naukowych, itp. (metoda indukcyjna) oraz wybór słów ze słowników terminologicznych, encyklopedii, indeksów rzeczowych (metoda dedukcyjna). Otóż autorzy na ogół chętniej pracowali tą drugą metodą - gdyż jest niewątpliwie łatwiejsza. Ale tu właśnie kryła się pewna pułapka. Zebrane przez nich słownictwo było bądź to zbyt ogólne, bądź zbyt przestarzałe i na ogół brakowało w nim terminów nowych. A poza tym występowały w nim poważne "luki", tzn. niektóre problemy czy nawet działy nauki w ogóle nie miały reprezentacji słownej, bądź ją miały, lecz bardzo ubogą. Niedomagania te wyszły właściwie dopiero w drugiej fazie pracy nad zebraniem słownictwa, tj. w trakcie recenzji dokonywanych przez specjalistów z poszczególnych dziedzin.

Ważnym zagadnieniem przy wyborze słownictwa kluczowego jest wyraźne określenie zakresu danej dyscypliny. Granice bowiem pomiędzy naukami społecznymi nie są dostatecznie wyraziste. Nauki te badają czasem te same zagadnienia (jak np. historia i prawo, historia i ekonomia) i tymi samymi metodami badawczymi, różniąc się tylko innym akcentowaniem ich poszczególnych aspektów i innym wnioskowaniem. Są też nauki tzw. "nowe", (jak nauka o polityce, religioznawstwo itp.), które biorą za swój przedmiot badań tradycyjnie przynależne



do innych dyscyplin zagadnienia oraz również nauki nowe interdyscyplinarne, zwane też "stykowymi", a rodzące się na pograniczu dwu lub nawet trzech starych dyscyplin. Np. na pograniczu psychologii i fizjologii powstała psychofizjologia, na pograniczu psychologii, lingwistyki i teorii informacji - psycholingwistyka, na pograniczu psychologii i ergonomii - psychologia pracy. Jest rzeczą oczywistą, że w ślad za tym rodzą się też nowe, niejako "stykowe" pojęcia i słowa i czasem trzeba arbitralnie rozstrzygać dla której z tych nauk - starej czy nowej - są one bardziej reprezentatywne, a dla której tylko pomocnicze, bądź zupełnie peryferyjne. Problem ten komplikuje jeszcze fakt, iż wśród samych uczonych nie ma jednolitego poglądu co do charakteru tych nowych nauk "stykowych". Jedni uważają je za nauki samodzielne, inni tylko za subdyscypliny czy wręcz integralne części nauki starszej. Dotyczy to m.in. relacji pomiędzy etyką i logiką a filozofią, bibliotekoznawstwem a literaturoznawstwem.

W zebranym słownictwie uderza to, iż autorzy poszczególnych wykazów słów kluczowych dawali w zasadzie tylko trzon starej dyscypliny. Nie dawali słów z subdyscyplin i nauk pokrewnych. Np. w psychologii nie ma słownictwa z wymienionych wyżej trzech jej subdyscyplin, jak również w literaturoznawstwie są tylko słowa dotyczące teorii literatury i jej historii. Nie ma natomiast bibliotekoznawstwa, edytorstwa, tekstologii i retoryki. Powstaje w związku z tym pytanie, które trzeba będzie w dalszej przyszłości rozstrzygnąć: w jaki sposób mają być konstruowane słowniki słów kluczowych z tzw. subdyscyplin i nauk pomocniczych? Czy mają to być samodzielne wykazy słów, czy też przyporządkować je należy dyscyplinie macierzystej?

Jedną z trudniejszych spraw przy sporządzaniu słownika jest kwestia prekoordynacji, a więc stopnia podziału słów wielowyrzowych na mniejsze jednostki leksykalne. Otóż opublikowane słowniki

mają stosunkowo wysoki stopień prekoordynacji. Jesteśmy jednak zdania, że (jeśli system informacyjny wyraźnie tego nie wymaga) nie należy wywierać w tym względzie zbyt wielkiego nacisku na autorów wykazów słów kluczowych, gdyż oni mają większe wyczucie co do powszechności i częstotliwości występowania danych terminów w nauce. Np. termin z politologii: Teoria stosunków międzynarodowych mógłby być podzielony na Teoria i Stosunki międzynarodowe. Jednakże autor uznał, że jest to zbyt ważna dziedzina w sferze nauki o polityce i termin ten należy zostawić w pierwotnym jego brzmieniu. Również inni autorzy słowników wskazywali na to, że zbyt daleko idąca postkoordynacja może utrudnić użytkownikowi wyszukiwanie informacji, gdyż wymagałaby od niego bardzo dużej wiedzy w danej dyscyplinie (które słowa z sobą łączyć, aby otrzymać pożądaną informację), a poza tym posiada on pewne nawyki i pewną "gotową" aparaturę pojęciową, a nie posiada znajomości logiki systemu czy języka, którym się posługuje. Innymi słowy sugestie naukowców idą w tym kierunku, aby nie rozrywać wielowyrzowych terminów o utrwalonym już znaczeniu. Może się bowiem zdarzać i tak, że przez podział danego terminu zmieni się w sposób zasadniczy jego sens. Prawnicy np. wskazują na takie niebezpieczeństwo przy dzieleniu terminów: Zafałszowanie wyborów i Ważność głosowania. Same bowiem wyrazy Zafałszowanie i Ważność nabierają wtedy znaczenia wyłącznie potocznego. Podobnie ma się rzecz w ekonomii politycznej ze słowami Stopa akumulacji czy Poziom życia. Samo słowo stopa (zresztą homonim) nic w ekonomii nie znaczy. Ale w połączeniu z innymi występuje aż w dwunastu bardzo ważnych terminach.

Spśród dwunastu rozpoczętych słowników słów kluczowych z nauk społecznych wstrzymaliśmy prace tylko nad prawem. Okazało się bowiem, że prawo, podobnie zresztą jak historia, jest szczególnie trudne do przełożenia na język słów kluczowych. Historia ma ogromną

masę faktów i nazw własnych, prawo znów posługuje się w istocie dwoma językami: językiem prawniczym, to jest naukowym i językiem prawnym, to jest językiem norm prawnych stosowanych w praktyce. Poza tym przedmiotem regulacji prawnej jest niemal wszystko. Dotyczy ono niemal wszystkich aspektów życia zbiorowego i jednostkowego. Stąd też ogromna masa jego słów musi mieć dodatkowe kwalifikatory, gdyż bez nich będą one właściwe dla innych dyscyplin, a nie dla prawa.

Główną jednak przyczyną wstrzymania pracy nad słownikiem z prawa jest fakt, iż pracownia informatyczna przy Kancelarii Sejmowej buduje już własny czteroczęściowy System Informacji Legislacyjnej (SIL) z własnym tezaurem. Nie chcieliśmy więc dublować tej pracy, jakkolwiek opracowany przez nas słownik słów kluczowych z dziedziny prawa mógłby być bardzo pożyteczny do uzupełniania istniejącego tam tezaury.<sup>2/</sup>

Jeśli idzie o pozostałe dyscypliny, to jakkolwiek zebrane dla nich słownictwo ma jeszcze pewne mankamenty, to jednak ogólnie rzecz biorąc ocenione ono zostało bądź jako właściwe i dobre, bądź jako dobra podstawa wyjściowa do dalszych nad nim prac.

Znamienne przy tym, iż wszyscy autorzy, a także i niektórzy recenzenci wskazują na fakt, iż wybór słownictwa kluczowego okazał się w praktyce znacznie trudniejszym zadaniem, niż to sądzono przed rozpoczęciem tych prac. A ponadto, że praca nad takim wyborem słów wyraziściej niż cokolwiek uwypukla "chaos" terminologiczny w niektórych dziedzinach nauki, oraz wzrastającą tendencję do zbyt częstego posługiwania się terminami obcymi, bez większego wysiłku szukania dla nich ekwiwalentów polskich.

---

<sup>2/</sup>Zob. T. Możdżyńska, Z. Wojciechowski: System Informacji Legislacyjnej - Stan i perspektywy do 1990 r. Aktualne Problemy Informacji i Dokumentacji 1987 nr 3 s. 10-12.

My zaś, opracowujący pod względem metodologicznym wszystkie słowniki, zauważamy jak wiele w naukach społecznych jest terminów identycznych, wspólnych, choć wydaje się, iż ich zakres znaczeniowy w poszczególnych naukach nie jest identyczny. A oto przykłady takich słów wspólnych: Aktywność, Abstrakcja, Analiza, Cecha, Czynniki, Depresja, Faza, Funkcja, Grupa, Klasa, Konflikt, Metoda, Obiekt, Pojęcie, Praca, Proces, Ruch, Struktura, System, Teoria, Układ. Gdyby się więc łączyło słowniki dziedzinowe w jeden słownik scalony, musiałoby się z pewnością dawać przy tych słowach dookreślenia (jako kwalifikatory) w rodzaju: (filozofia), (socjologia), (politologia) itd.

Kończąc pragnę jeszcze podkreślić fakt, iż opublikowane słowniki nie mają charakteru normalizacyjnego. Głównym ich bowiem przeznaczeniem jest pomoc przy indeksowaniu i wyszukiwaniu dokumentów. Mają więc one charakter wybitnie pragmatyczny. Nie są one słownikami kompletnymi, choćby z powodu braku w nich słownictwa subdyscyplin poszczególnych nauk. Nie są także zbiorami zamkniętymi, lecz otwartymi, wymagającymi ciągłej aktualizacji.

Ukazanie się tych słowników jest w przeważającej części zasługą ich zespołów autorackich i recenzentów, którzy za niewielkim wynagrodzeniem, a więc z pobudek prawie wyłącznie intelektualnych, wykonali tę olbrzymią i - w naszym odczuciu - bardzo pożyteczną pracę.

. Alina Fitowa

2.02.1989 r.

## W POSZUKIWANIU MODELU NARODOWYCH SYSTEMÓW INFORMACJI

Jako motto omawianej książki\* wykorzystał G. Rózsa sformułowanie W. Hugo "Połączyć myśl kraju z sercem narodu", stanowiące równocześnie uzupełnienie i parafrazę znanego z dokumentów UNESCO stwierdzenia, iż siła narodu tkwi w jego zasobach naturalnych, źródłach energii oraz informacji. Publikacja ta jest zbiorem studiów i esejów powstałych w latach siedemdziesiątych i osiemdziesiątych. Odzwierciedlają one różnorodne aspekty poszukiwania optymalnego modelu narodowych systemów informacji w myśl przytoczonej dewizy W. Hugo. Prace zawarte w zbiorze łączy ponadto problem nadrzędny - sformułowana z dystansem i znanstwem ocena możliwości wykorzystania międzynarodowej współpracy w dziedzinie informacji naukowej do kształtowania narodowych polityk informacyjnych w poszczególnych krajach. Nieczęsto zdarza się, by autorzy publikacji podejmujących globalne problemy rozwoju informacji legitymowali się wiedzą i doświadczeniem podobnym do reprezentowanego przez dr Gy. Rózse: wykształceniem w zakresie bibliotekoznawstwa i ekonomii, gruntowną znajomością codziennej rzeczywistości biblioteczno-informacyjnej wyniesioną z wieloletniego kierowania Biblioteką Węgierskiej Akademii Nauk oraz Biblioteką ONZ w Genewie, udziałem w pracach licznych organizacji międzynarodowych zajmujących się problemami przepływu informacji. Z drugiej strony perspektywę podejmowanych problemów nie-

---

\* Rózsa György: Information: from claims to needs. National aptitudes for international co-operation in scientific information economy. Budapest: Kultura, Library of the Hungarian Academy of Sciences 1988 202 s.

jako modyfikuje punkt widzenia potrzeb własnego kraju - niewielkiego pod względem obszaru i ludności, choć reprezentującego znaczny potencjał naukowy i stojącego przed trudnym problemem określenia polityki informacyjnej stosownej do bieżących i przyszłych potrzeb nauki i gospodarki.

Perspektywy rozwoju narodowych systemów informacji rozpatruje się zwykle w kontekście opozycji "tradycyjne" biblioteki - "nowoczesne", skomputeryzowane systemy informacji. Teza autora jest odmienna.. " społeczeństwo potrzebuje wszelkich rodzajów informacji, niezależnie od jej pochodzenia i formy .../.../ ...tym, co pragnę wypuklić jest zawartość, użyteczność i wartość informacji, nie zaś kanały i techniki jej przekazu"... (s.11). Autor omawia trzy grupy problemów:

- społeczne i zawodowe aspekty rewolucji w informacji naukowej,
- rolę współpracy międzynarodowej w zakresie informacji, ilustrowaną osobistymi doświadczeniami,
- infrastrukturę informacyjną w krajach rozwijających się.

W szczegółowych referatach i esejach rozwijających wymienione problemy podkreślana jest zasada ścisłego uzależnienia rozwoju każdej dziedziny nauki i gospodarki od dostępu do relewantnej dla użytkownika informacji, niezbędnej przede wszystkim w procesach decyzyjnych na wszystkich szczeblach. Informacja naukowa jako dziedzina interdyscyplinarna jest ściśle powiązana z ekonomią, jej instytucje powinny być traktowane jako jednostki produkcyjne przynoszące dochód, przy czym jest to dochód szczególnego rodzaju - dochód intelektualny. Przy współczesnej specjalizacji w nauce i gospodarce - twierdzi autor - racjonalne podejmowanie decyzji zależy w dużej mierze od dysponowania porównywalną informacją o aktualnych tendencjach rozwoju wiedzy w krajach rozwiniętych. Pozwala ona zrównoważyć w procesie decyzyjnym potrzeby rozwoju nauki w skali narodo-

wej wynikające z jej tradycji i kultury narodu z potrzebami ekonomicznymi, długo- i krótkofalowymi. O użyteczności informacji decydują: jej treść i sposób intelektualnego przetworzenia. W szczególności w latach siedemdziesiątych powszechne było przekonanie, iż rozwój komputeryzacji serwisów informacyjnych może stanowić panaceum na wszelkie niedomagania w tym zakresie spowodowana tzw. eksplozją informacji. W pracach powstałych w tym właśnie okresie autor podkreśla, że nie technika przetwarzania czyni informację "nowoczesną", wskazując równocześnie, że do pogłębienia kryzysu informacyjnego przyczynił się między innymi lawinowy rozwój serwisów dokumentacyjnych, adresowanych do anonimowego odbiorcy (impersonalized) i faktycznie ubogich w informację. Dowodów na słuszność tego poglądu dostarcza wieloletnia praktyka we współpracy międzynarodowej w zakresie informacji nauk społecznych (MISON, ECSSID). W systemach tych należyta obsługa użytkownika może być zapewniona przez sprawnie funkcjonującą bibliotekę, szybki dostęp do dokumentów pierwotnych oraz merytoryczną kompetencję pracowników bibliotek i ośrodków informacji. Różni to zasadniczo informację w naukach społecznych od informacji w naukach ścisłych, w których transfer danych odbywa się niejako poza systemem społecznym i podlega w nieporównywalnie wyższym stopniu działaniu czynnika czasu (czynnika starzenia się informacji). Ekspozowanie absolutnego priorytetu indywidualnych potrzeb użytkownika w obsłudze informacyjnej nie oznacza w żadnej mierze deprecjonowania przez autora znaczenia technologii przetwarzania informacji, chodzi bowiem jedynie o podkreślenie jej służebnej roli w kulturotwórczym funkcjonowaniu narodowych służb informacyjnych. Racjonalne kształtowanie polityki informacyjnej powinno wyrażać się w wyważeniu proporcji w rozwoju serwisów informacyjnych dla potrzeb kultury narodowej oraz tych przeznaczonych do wymiany międzynarodowej na zasadach komercyjnych.

Warunki do prowadzenia takiej wymiany powinny stworzyć rządy poszczególnych krajów stosownie do prowadzonej polityki informacyjnej. Jak wykazały liczne eksperymenty i programy realizowane w ramach UNESCO i innych organizacji międzynarodowych nie ma uniwersalnego modelu polityki informacyjnej dla krajów rozwijających się. Wiele z opracowanych programów okazało się niekompatybilnymi lub nieprzystosowanymi do lokalnych potrzeb uwarunkowanych tradycją i specyfiką kulturową poszczególnych narodów. Autor podkreśla, iż nadmierne poleganie na technologii przetwarzania informacji przy braku organicznych podstaw sieci informacyjnych w postaci sprawnie funkcjonujących bibliotek prowadzi do swoistej "kolonizacji" informacyjnej i kulturowego uzależnienia się od krajów rozwiniętych. Podporządkowana zasadom ekonomiki informacja powinna rozwijać się na poziomie mikro- i makro-, zwłaszcza jeśli ma zapewniać obsługę w zakresie badań i rozwoju. W proponowanym, zindywidualizowanym modelu narodowych systemów informacji powstaje zapotrzebowanie na nowy typ specjalizacji wśród kadr informacyjnych (definiowany w pracy z początku lat siedemdziesiątych) - "information officer", zapewniający indywidualną obsługę pracowników badawczych, którego praca - jak zaznacza autor - zaczyna się tam, gdzie kończy się praca dokumentalisty w serwisach informacyjnych. Status zawodowy tej części kadry informacyjnej powinien być taki sam jak status pracowników badawczych, ponieważ indywidualna obsługa jest najlepszą metodą stymulowania badań w zakresie B+R.

Praktyka serwisów informacyjnych i niepowodzenia uniwersalnych programów rozwoju informacji w ostatnich latach potwierdziły słuszność poglądów autora omówionej pracy.

Elżbieta Artowicz

2.05.1989 r.





## KRONIKA

### KONFERENCJA "POWSZECHNA DOSTĘPNOŚĆ DO PUBLIKACJI"

Warszawa 24-25 kwietnia 1989 r.

W dniach 24-25 kwietnia br. odbyła się w Warszawie, zorganizowana wspólnie przez Bibliotekę Narodową i Stowarzyszenie Bibliotekarzy Polskich konferencja, poświęcona sprawom dostępności do publikacji w Polsce, w kontekście programu głównego IFLA zatytułowanego: "Powszechna Dostępność do Publikacji" (Universal Availability of Publications - UAP). Wzięli w niej udział przedstawiciele bibliotek naukowych i części bibliotek publicznych.

W pierwszym dniu konferencji, poświęconym omówieniu ogólnych problemów powszechnej dostępności do zbiorów, przedstawiono 4 referaty. W pierwszym wystąpieniu, które autorzy Stanisław Czajka i Jan Wołoz z tytułowali "Rola Biblioteki Narodowej w kształtowaniu warunków sprzyjających powszechnej dostępności do publikacji", omówione zostały kierunki działań Biblioteki Narodowej mających na celu: po pierwsze, ułatwienie dostępu do zbiorów i po drugie, eliminację ograniczeń dostępności. Zwrócono uwagę na szeroki zakres prac wydawniczych (bibliografie, katalogi, informatory), a także podejmowane prace naukowe. Efektem tych prac jest stworzenie podstawowego trzonu informatorów oraz poradników i pomocy metodycznych pozwalających rozwiązywać problemy powstające przy korzystaniu ze zbiorów. Na kanwie tych informacji zrodziło się pytanie czy Biblioteka Narodowa ma samodzielnie prowadzić tak szeroki front prac oraz dość automatycznie odpowiedź, że w dzisiejszych warunkach konieczny jest powrót do autentycznej współpracy bibliotekarzy i odejście od oczekiwania na centralne rozwiązywanie tych

problemów oraz innych, związanych ze specjalizacją bibliotek, wypożyczeniami, automatyzacją prac bibliotecznych.

"Programy główne IFLA ze szczególnym uwzględnieniem UAP" przedstawił Adam Wysocki, dając bardzo ciekawy przegląd dziejów powstania tychże programów, obecnych ich zakresów tematycznych, celów i zadań, zasad ich finansowania itd. Programy te to: najstarszy "Powszechna Rejestracja Bibliograficzna i Międzynarodowy MARC"<sup>1/</sup>; dalej "Powszechna Dostępność do Publikacji", którego zadaniem jest ujawnienie przeszkód i dążenie do pełnego udostępnienia publikacji na całym świecie; trzeci, obecnie najbardziej dynamicznie rozwijający się, dążący do wypracowania norm i poszukiwania najskuteczniejszych metod działania: "Konserwacja i Ochrona Zbiorów"; czwarty "Powszechny Przepływ Danych i Telekomunikacja", którego celem jest promowanie elektronicznego przepływu danych, tworzenie norm międzynarodowych m.in. dla łamania barier telekomunikacyjnych i program piąty, nie mający jeszcze nazwy dotyczących rozwoju bibliotek w krajach rozwijających się. Referent przedstawił również osiągnięcia tych programów i bieżącą działalność IFLA w zakresie każdego z nich.

Referat trzeci, Krystyny Ramlau-Klekowskiej: "Udział Polski w programie UBC i wpływ tych prac na usprawnienie informacji bibliograficznej" uzmysłowił słuchaczom, iż w momencie przystąpienia do jego realizacji podejmowane w Polsce działania były zgodne na ogół z pryncypiami programu. Stwarzało to szansę na jego dobrą realizację. Stan obecny w tym zakresie oceniła autorka niestety dość negatywnie przedstawiając wiele dokumentujących tę ocenę zjawisk. Do

---

<sup>1/</sup> Dawniej 2 programy: UBC - Universal Bibliographic Control i IMP - International MARC Programme obecnie wspólny: UBCIM - Universal Bibliographic Control International MARC.

najważniejszych z nich należy zaliczyć: duże opóźnienia druku części składowych naszej bibliografii narodowej, nieuwzglębnienie przez nią wszystkich druków, brak realizacji założonych celów komputeryzacji (np. wymiennalności danych), opory przed stosowaniem przyjętego ogólnie systemu MARC. Autorka zwróciła również uwagę na niedostępność w Polsce wielu prac metodycznych IFLA z tego zakresu, brak kontaktów osobistych, nieuczestniczenie w konferencjach i naradach. Pewien optymizm można wiązać z otrzymaniem przez Bibliotekę Narodową nowego komputera, co powinno pomóc w rozwiązaniu części problemów.

Lucjan Biliński poświęcił swe wystąpienie omówieniu zagadnienia "Powszechnej dostępności do materiałów bibliotecznych w świetle przepisów prawnych". Ogólnie można stwierdzić, że istnieje teoretycznie wiele przepisów pozwalających na pełną, powszechną dostępność do publikacji, ale są i takie, które tę dostępność ograniczają. Do nich należy zaliczyć np. nie w pełni precyzyjne sformułowanie przepisów o egzemplarzu obowiązkowym (sprawa tzw. materiałów poufnych), złe ich egzekwowanie w praktyce, dalej wydane w oparciu o nowelizację Ustawy o Kontroli Publikacji i Widowisk zarządzenia nr 1 i 2 Ministrów: Kultury i Sztuki, Szkolnictwa Wyższego i Techniki oraz Sekretarza Naukowego PAN i Prezesa GUKPiW z 30 kwietnia 1984 roku. Oprócz barier formalnych istnieją też inne np. finansowe czy związane z udostępnianiem, wypożyczaniem, które wprowadzają również same biblioteki.

Po wysłuchaniu tych referatów nastąpiła wymiana informacji o trybie rozwiązywania pewnych problemów związanych z dostępnością do publikacji w poszczególnych bibliotekach jak i zagorzała dyskusja, jakie podjąć działania szczególnie w zakresie uprzywilejowania informacji i udostępniania powszechnego tak druków zagranicznych

jak i tych pochodzących z tzw. drugiego obiegu. Główne jej ustalenia zawarte zostały we wnioskach końcowych konferencji.

Drugi dzień obrad poświęcono zagadnieniom gromadzenia zbiorów, informacji i udostępniania. Rozpoczął się od wystąpienia Artura Jazdona i Emilii Kihl-Szymańskiej, przedstawiającego "Problemy gromadzenia piśmiennictwa zagranicznego w bibliotekach szkół wyższych". Ogólnie sytuacja w tym zakresie jest bibliotekarzom znana, lecz to wystąpienie omawiające wyniki ankiety przeprowadzonej wśród 44 bibliotek uczelnianych różnego typu pozwoliło na uzyskanie pewnego statystycznego obrazu zjawiska i nakreślenie tendencji jego rozwoju na przestrzeni ostatnich lat. Ogólnie referenci na podstawie danych ocenili je jako dość niekorzystne, wskazując problemy konieczne do rozwiązania, co znalazło także swój wydzźwięk we wnioskach końcowych.

Stanowi realizacji i problemom związanym z postępowaniem specjalizacji w zakresie gromadzenia zbiorów bibliotecznych poświęcony był referat Anny Romańskiej. Po przedstawieniu historii atarań, głównych zadań i obecnie realizowanego w tym zakresie obszaru działań, prelegentka przedstawiła 4 wnioski, których realizacja winna przyczynić się do przyspieszenia prac nad specjalizacją zbiorów. Przedstawiając je hasłowo wynotujemy: 1. intensyfikacji wymagają prowadzone prace metodyczne, 2. konieczne jest zwiększenie środków z budżetu centralnego na biblioteki centralne przy uświadomieniu im ich obowiązków na rzecz bibliotek sieci, 3. kontynuowanie prac nadal musi pozostawać w gestii wszystkich sygnatariuszy wcześniejszych dokumentów, 4. za główny cel uważać należy - i realizować - dążenie do wzbogacenia zbiorów jako podstawy do dalszej działalności (wypożyczeń, informacji, reprografii).

Andrzej Gwizdź przedstawił problem powszechnej dostępności do dokumentów i wydawnictw urzędowych. Referent zdefiniował tego typu wydawnictwa i stwierdził, że nie zawierają one dziś wszystkich przepisów, pogarsza się też sytuacja w zakresie tempa ich publikowania. Te ograniczenia ich dostępności są potęgowane przez trudności w dotarciu do większości druków urzędowych oprócz "Monitora Polskiego" i "Dziennika Ustaw" oraz trudności w docieraniu do określonych przepisów co utrudnia brak odpowiednich indeksów. Podsumowując, referent określił dostępność do dokumentów urzędowych w Polsce jako dużo gorszą niż w wielu państwach Europy Zachodniej, choć stwierdził, że obserwowane zmiany polityczne, rosnąca świadomość potrzeb w tym zakresie ogółu społeczeństwa rokują wprowadzenie pozytywnych zmian.

Kolejna referentka, Jakobina Kowalczyk, przedstawiła referat pt.: "Informacje o zbiorach bibliotecznych". Po udzieleniu odpowiedzi na pytanie ile i jakich mamy bibliotek w Polsce, autorka przedstawiła szczegółowo, jaką obecnie dysponujemy informacją o zasobach bibliotecznych w Polsce i jak na jej podstawie zorientować się można o zawartości bibliotek. Konkluzję wywodów stanowiło przedstawienie wniosku o konieczności rozwoju katalogów centralnych, co uzależnione jest z kolei od stopnia współpracy bibliotek i ich wyposażenia technicznego, rozwoju prac metodycznych. Wszystko to w rezultacie winno ułatwić działania nad stworzeniem krajowego systemu udostępnień, podjęcie których to prac autorka również postulowała.

Ostatni referat "Wypożyczania międzybiblioteczne: stan i perspektywy" przedstawił Stefan Kubów. Po bardzo klarownym naszkicowaniu szerokiego tła (zasad organizacyjnych i działania wypożyczeń międzybibliotecznych w wybranych krajach świata) referent ekon-

centrował się na scharakteryzowaniu sytuacji panującej w Polsce. Na dość negatywną jej ocenę wpływają jego zdaniem głównie: brak postępów nad opracowaniem katalogów centralnych, brak narzędzi łączności, zauważalna niechęć do udostępniania zbiorów, nieznamość języków, niestaranność opisów dokumentów, brak zainteresowania tematyką wypożyczeń teoretyków bibliotekoznawstwa i pomijanie tego zagadnienia w zasadniczych dokumentach (np. Raport o bibliotekach). Likwidacja tych niedostatków, rozwijanie wśród użytkowników przeświadczenia o tym, że mogą do wszystkiego dotrzeć i ze wszystkiego korzystać, pozwoli stworzyć skutecznie działający system wypożyczeń międzybibliotecznych.

I w tym dniu rozgorzała dyskusja, z której należy odnotować szczególnie głos prof. Jadwigi Kołodziejskiej, która apelowała o obserwowanie dokonujących się w Polsce zmian. Jeśli bibliotekarze chcą, aby to, co gromadzą w wielkim trudzie było wykorzystywane, a więc było faktycznie dostępne, musi zaistnieć chęć korzystania, powszechnego docierania do bibliotek i ich zbiorów. Dokona się to wówczas, gdy zmiany społeczno-gospodarcze zaowocują tym, że wiedza, wykształcenie, kwalifikacje będą w cenie, co przywiedzie do bibliotek nowe zastępy aktywnie poszukujących informacji użytkowników.

W wyniku dyskusji i postulatów wysuwanych w wystąpieniach programowych przyjęto szereg wniosków, z których za najważniejsze uważać należy wnioski o:

- uchylenie nowelizacji Ustawy o Kontroli Publikacji i Widowisk z 1983 r., wspomnianych zarządzeń z 1984 r. a także liberalizację przepisów kodeksu karnego w tych miejscach, które ograniczają udostępnianie;
- przyznanie debitu wszystkim wydawnictwom zagranicznym oraz zwrócenie bibliotekom materiałów zatrzymanych w ostatnich latach;

- zniesienie wszelkich prohibitów, tajnych katalogów i kartotek w bibliotekach i wprowadzenie zasady, iż o wszelkich ograniczeniach udostępniania decydować mają same biblioteki w zależności od posiadanych możliwości lokalowych, technicznych itp.;
- objęcie pełną rejestracją w polskiej bieżącej bibliografii narodowej druków z tzw. drugiego obiegu;
- podjęcie działań nad usprawnieniem systemu prenumeraty czasopism;
- poddanie analizie dotychczasowego stanu w zakresie specjalizacji zbiorów i wyznaczenie prac w tym zakresie na najbliższe 10-15 lat;
- podjęcie prac nad krajowym systemem wypożyczeń.

Te i inne wnioski, ostatecznie zredagowane przez komisję wnioskową, zostaną przesłane do wszystkich bibliotek biorących udział w konferencji oraz innych instytucji i urzędów celem ich realizacji.

Organizatorzy narady zamierzają pełne teksty wystąpień opublikować w jednym z kolejnych tomów "Rocznika Biblioteki Narodowej". Za to jak i za bardzo sprawną organizację narady, która naświetliła wiele problemów i powinna przyczynić się poprzez realizację przyjętych ustaleń do zwiększenia faktycznej, powszechnej dostępności do publikacji w Polsce, należą się słowa uznania.

Artur Jazdon

5.05.1989 r.





## SPIS TREŚCI

1. E. Artowicz Interpretacje terminu "relewancja" w informacji naukowej .....	3
2. J. Woźniak O tak zwanych potrzebach informacyjnych .....	39
3. A. W. Sokołow Potrzeby informacyjne .....	61
4. L. V. Sacharnyj Struktury kojarzeniowe w języku naturalnym i języku informacyjno-wyszukiwawczym (na przykładzie języka haseł przedmiotowych) .....	75
5. E. Artowicz, B. Szomański System Informacji o Metodach i Technikach Organizatorskich w Sterowaniu Jakością .....	103

### Materiały i przyczynki

1. D. Głazek Możliwość zwiększenia efektywności wyszukiwania informacji w języku swobodnych słów kluczowych .....	125
2. P. Żak Międzynarodowy format wymiany danych bibliograficznych .....	139
3. J. Sadowska Uwagi na temat formatu POLMARC .....	155

### Recenzje i omówienia

1. Z doświadczeń krakowskiej Pracowni OIN PAN nad tworzeniem słowników słów kluczowych nauk społecznych - A. Fitowa ...	167
2. W poszukiwaniu modelu narodowych systemów informacji - E. Artowicz .....	180

### Kronika

Konferencja "Powszechna Dostępność do Publikacji" Warszawa 24-25 kwietnia 1989 r. - Artur Jazdon .....	185
--	-----

## CONTENTS

1. E. Artowicz Interpretations of the term "relevance" in information science .....	3
2. J. Woźniak On so called information needs .....	39
3. A.W. Sokolov Information needs .....	61
4. L.V. Sacharnyj The associative structures in natural and information retrieval language (on the exemple of the subject headings language) .....	75
5. E. Artowicz, B. Szomański The Information System on Organizational Methods and Techniques in Quality Control (SIMTOJ)	103

### Materials and Contributions

1. D. Głażek Possibility of increasing the efficiency of information retrieval by means of the free key words .....	125
2. P. Żak The international format for the exchange of bibliographic data .....	139
3. J. Sadowska Remarks of the POLMARC format .....	155

### Reviews and Surveys

1. About some experience of the Department of Scientific Information Centre of the Polish Academy of Sciences in Kraków from works on preparing the dictionaries of key words for social sciences - A. Fitowa .....	167
2. Looking for a model of national information system - E. Artowicz .....	180
Chronicles .....	185

## СО Д Е Р Ж А Н И Е

1. Э. Артович Интерпретации термина "релевантность в научной информации.....	3
2. Я. Вовшик О так называемых информационных потребностях..	39
3. А.В. Соколов Информационные потребности.....	61
4. Л.В. Сахарный Ассоциативные структуры в естественном языке и в предметизационном илп.....	75
5. Э. Артович, Б. Шоманьски Информационная система об организаторских методах и техниках для стимулирования качеством ( SIMTOJ ) .....	103

### М а т е р и а л ы и п р и м е ч а н и я

1. Д.Глазек Возможность повышения эффективности поиска информации на языке свободных ключевых слов .....	125
2. П. Жак Международный формат для обмена библиографическими данными .....	139
3. Я.Садовска Замечания на формат POLMARC .....	155

### Р е ц е н з и и и о б з о р ы

1. Из опыта краковского Отдела ЦНИ ПАН по созданию словарей ключевых слов общественных наук - А. Фитова .....	167
2. Поиск модели национальных информационных систем - Э.Артович	180
Х р о н и к а .....	185

## Wskazówki dla Autorów

Redakcja „Zagadnień Informatyki Naukowej” uprzejmie prosi Autorów o przestrzeganie następujących zasad przy nadsyłaniu materiałów.

Artykuł nie powinien przekraczać 30 stron maszynopisu formatu A4 wraz z przypisami i ewentualnymi tablicami (podwójny odstęp między wierszami, ok. 30 wierszy na stronie, margines 3,5 cm lewy, 1 cm prawy).

Maszynopis artykułu należy dostarczyć w 2 egzemplarzach. Ilustracje (tablice, wykresy) powinny być umieszczone na osobnych ponumerowanych stronach z zaznaczeniem ich miejsca w tekście.

Przypisy i bibliografię należy umieszczać na końcu tekstu i na osobnych stronach. Opis bibliograficzny powinien składać się z: nazwiska i imienia autora, tytułu publikacji, miejsca wydania, nazwy wydawnictwa i roku wydania oraz stron. W wypadku powoływania się na artykuł opis powinien zawierać: nazwisko i imię autora, tytuł artykułu, nazwę czasopisma, rok, tom (numer), strony lub stronę zawierającą cytowany fragment.

Każdy artykuł powinien być zaopatrzony w streszczenie autor-skie ok.  $\frac{1}{2}$  strony maszynopisu.

Autorzy proszeni są o podawanie do wiadomości redakcji (oprócz imienia i nazwiska) również tytułu naukowego, nazwy i adresu instytucji, w której pracują, prywatnego adresu zamieszkania, numeru telefonu i ewentualnie, jeśli honorarium autorskie ma być przekazane na konto, numer konta PKO.

O przyjęciu do druku decyduje Rada Redakcyjna.

Maszynopisów artykułów nie zamówionych nie przyjętych do druku redakcja nie zwraca.

**Cena zł 1300,-**