



Opracowanie przedmiotowe dokumentów
z zakresu nauk ścisłych: matematyczno-przyrodniczych i technicznych.
Język haseł przedmiotowych KABA:
teoria, praktyka, przyszłość

Kazimierz Dolny, 20-22 września 2006

[Poprzedni](#) - [Spis treści](#) - [Następny](#)

Andrzej Krajka
Artur Ściborek

Institut Bibliotekoznawstwa i Informatyki Naukowej. Uniwersytet im. Marii Curie-Skłodowskiej, Lublin
akrajka@klio.umcs.lublin.pl

Ośrodek Informatyki Akademii Rolniczej w Lublinie
artur@ar.lublin.pl

Efektywność wyszukiwawcza w języku haseł przedmiotowych KABA i innych językach wyszukiwawczych ***Searching effectiveness in the SIS KABA language and other searching languages***

Słowa kluczowe

Jhp KABA, wyszukiwanie informacji, logi transakcyjne, katalogi OPAC

Keywords

Subject Indexing System KABA, information retrieval, transaction log analysis, OPAC Catalogues

Abstrakt

W artykule tym analizujemy różne sposoby wyszukiwania dokumentów, jakie stosują użytkownicy katalogów komputerowych Biblioteki Głównej Akademii Rolniczej w Lublinie. Szczególnie interesować nas będzie, jak na tle różnych typów wyszukiwań i zachowań użytkowników tych katalogów prezentują się wyszukiwania według języka haseł przedmiotowych (jhp) KABA. Jakość wyszukiwania spróbujemy ocenić na podstawie reakcji użytkowników. Podstawą tych badań jest materiał zgromadzony przez informatyków na przełomie stycznia i lutego 2006 r. rejestrujący zachowania użytkowników katalogów komputerowych biblioteki podczas komunikowania się z serwerem Biblioteki Głównej AR.

Abstract

In this paper we would like to present the different ways of information retrievals, which are used by Main Library of the Agricultural University of Lublin users. Especially we are interested in the problem, how on the background of different retrieval methods seems the SIS KABA retrieval. The efficiency of searching are evaluated on the base of further reactions of users. The research has been based on the rich material given for two weeks of January and February 2006.

1. Historia jhp KABA w Bibliotece Głównej AR

1.1. VTLS w Bibliotece Głównej AR

Pracę w zintegrowanym systemie bibliotecznym VTLS w BG AR rozpoczęto w marcu 1996 r. Każdą pozycję opisywano z autopsji, a nie na podstawie katalogu kartkowego; zapewniło to wysoką jakość danych, podnosząc jednocześnie efektywność wyszukiwania w katalogu komputerowym. Szybkość wprowadzania ilustruje Tab. 1.

Tab. 1 – Stan ilościowy katalogu komputerowego w latach 1996-2004

Stan katalogu komputerowego Biblioteki Głównej AR w latach 1996-2004				
Rok	Liczba rekordów bibliograficznych	Liczba rekordów egzemplarzy	Liczba rekordów zasobów wydaw. ciągłych	Liczba rekordów haseł wzorcowych
1996	4 307	28 533	181	5 715
1997	18 387	57 595	719	22 608

1998	22 364	66 262	1 314	27 904
1999	26 665	74166	1689	34 495
2000	29 897	81472	1914	40 611
2001	33 665	87 683	2 214	46 147
2002	37 876	92 826	2 349	52 220
2003	42 030	98 124	2 487	55 278
2004	45 058	104 495	2 530	61079

Równoległe z opracowaniem zbiorów trwały przygotowania do wdrożenia kolejnego modułu VTLS - udostępniania. We wrześniu 1998 r. uruchomiono obsługę wypożyczalni. Drugi etap wdrażania modułu udostępniania przyniósł kolejne udogodnienia. W lutym 2000 r. wprowadzono automatyczne zamawianie książek i czasopism. Z katalogiem można było połączyć się za pomocą telnetu lub poprzez bramkę WWW-VTLS Web Gateway, a także bezpośrednio ze strony domowej biblioteki.

Analiza połączeń wykazała, że użytkownicy bardzo szybko docenili to udogodnienie. Przydatne okazały się także inne usprawnienia, podnoszące jakość korzystania z katalogu komputerowego, na przykład możliwość zapoznania się czytelnika ze stanem konta oraz informacja o statusie książki, pozwalająca zorientować się, czy dana pozycja jest w Bibliotece dostępna czy też wypożyczona. Ale największą zaletą z punktu widzenia potrzeb użytkowników była możliwość wieloaspektowego wyszukiwania informacji według następujących kryteriów: autor, tytuł, hasło przedmiotowe, słowa kluczowe, instytucja sprawcza, tytuł serii, tytuł czasopisma oraz sygnatura. Zarówno studenci, jak i pracownicy naukowcy uczelni, doceniając walory katalogu komputerowego, w tym także dużą oszczędność czasu, bez większych oporów zamienili katalogi kartkowe na katalog dostępny on-line.

Katalogi kartkowe były sukcesywnie zamykane. Jako pierwszy zamknięto w 1997 r. katalog przedmiotowy wydawnictw zwartych. Od tej pory wszystkie książki opracowywano w jhp KABA. Ze względu na czytelników dłużej prowadzono katalogi alfabetyczne wydawnictw zwartych i wydawnictw ciągłych.

1.2. NUKAT

Omówienie działalności Biblioteki w zakresie komputeryzacji nie byłoby pełne bez wspomnienia o bardzo ważnym dla środowiska bibliotekarskiego przedsięwzięciu, jakim jest Narodowy Uniwersalny Katalog Centralny NUKAT. Koncepcja NUKAT zakładała budowę katalogu metodą współkatalogowania; jego podstawę stanowiły istniejące bazy CKHW i CKTCz. Nadzór nad całością przedsięwzięcia sprawował powołany w 1998 r. Zespół Koordynacyjny NUKAT. W Zespole tym biblioteki lubelskie reprezentował Andrzej Padziński, przedstawiciel Biblioteki Głównej AR. W 2001 r. dotychczasowe Centrum Formatów i Kartotek Haseł Wzorcowych zostało przekształcone w Centrum NUKAT i przejęło zadanie prowadzenia katalogu centralnego.

Rok 2003 przyniósł kolejne zmiany w Bibliotece. Kupiony osiem lat wcześniej sprzęt był już przestarzały i nie nadawał się do rozbudowy. Należało wymienić serwer i oprogramowanie na nowsze. Z podobnymi problemami borykały się także inne biblioteki pracujące w VTLS. Razem podjęto decyzję o wymianie dotychczas stosowanego oprogramowania na Virtue. Wspólne negocjacje z firmą komputerową okazały się bardzo korzystne, ponieważ umożliwiły uzyskanie znacznych upustów dla tak dużego przedsięwzięcia. Zmiana oprogramowania i przeniesienie danych ze starego systemu do nowego zostały przeprowadzone równocześnie w 20 bibliotekach w kraju. Czytelnicy nie odczuli z tego powodu większych utrudnień, gdyż migrację przeprowadzono w okresie wakacyjnym. Virtue została dobrze przyjęta. Zarówno studenci, jak i pracownicy nie mieli większych problemów z przestawieniem się na nowy system, ponieważ jest on bardziej przyjazny dla użytkownika i zwiększa możliwość efektywnego wyszukiwania. Nie bez znaczenia jest także i to, że oprogramowanie wykorzystywane w katalogu lokalnym jest identyczne z oprogramowaniem katalogu centralnego.

Zarysowana pokrótce historia komputeryzacji biblioteki wyraźnie pokazuje, że konsekwencją wdrożenia zintegrowanego systemu bibliotecznego VTLS były głębokie zmiany w jej działalności. Bibliotekarze stanęli przed trudnym zadaniem. Zdobycie nowej wiedzy i umiejętności sprawnego posługiwania się komputerem stało się koniecznością w każdym oddziale i na każdym stanowisku. Rezygnacja z wieloletnich przyzwyczajeń i nawyków nie dla wszystkich okazała się łatwa do zaakceptowania, ważne były pozytywne nastawienie i psychiczna gotowość na tak radykalne zmiany.

2. Opis danych

2.1. Zbiór wyszukiwawczy

Księgozbiór Biblioteki Głównej Akademii Rolniczej w Lublinie obejmuje swym zakresem: nauki rolnicze, ogrodnictwo, medycynę weterynaryjną, zootechnikę, technikę rolniczą, maszynoznawstwo, przemysł spożywczy, technologię żywności, ekonomikę rolnictwa, agrobiznes, biologię, ekologię, ochronę środowiska oraz wybiórczo: matematykę, medycynę, chemię i podstawy techniki. Zbiory liczą ok. 358 tys. woluminów, w tym 220 tys. wol. książek, 125 tys. wol. czasopism, 13 tys. wol. zbiorów specjalnych.

2.2. Metody wyszukiwania informacji w katalogach OPAC

Komputerowy katalog OPAC dostępny na stronie internetowej <http://katalog.ar.lublin.pl/> posiada następujące typy interfejsów: indeksy, słowa kluczowe, zaawansowane, słowa w hasłach, konto czytelnika i szukaj słowa. W każdym z tych wyszukiwań można również przejść do schowka, w którym można przechowywać opisy egzemplarzy, obejrzeć historię wyszukiwań, szukać słowa kluczowego w specjalnego typu dokumentach (wszędzie, książki, czasopisma, dysertacje) oraz używać jednego z trzech filtrów: język, data wydania i lokalizacja. W przypadku wyświetlania wyszukiwanych pozycji możliwe jest obejrzanie rekordów bibliograficznych lub obejrzanie opisów w formacie MARC. W przypadku indeksów możemy wyszukiwać

według następujących indeksów:

- a. Autorski
- b. Tytułowy
- c. Tytuł czasopisma
- d. Hasło przedmiotowe
- e. Hasło przedmiotowe ang./franc.
- f. Sygnatura

W przypadku wyszukiwań wg słów kluczowych możemy wybrać:

- a. Autor
- b. Tytuł
- c. Hasło przedmiotowe
- d. ISBN
- e. ISSN
- f. Wszędzie

Wyszukiwania zaawansowane można stosować z następującymi opcjami:

- aw: wszędzie
- a: autor
- t: tytuł
- zwbib: wcześniejszy lub późniejszy tytuł czasopisma
- tse: tytuł serii
- s: hasło przedmiotowe
- n: nazwa (osoby, instytucji, konferencji)
- u: uwagi
- w: wydawca
- mw: miejsce wydania
- sd: data wydania
- nr: ISBN, ISSN
- lang: język (kod języka)

Natomiast w wyszukiwaniach słowa w hasłach można poszukiwać zadanego słowa wg kategorii

- a. Tytuł
- b. Hasło przedmiotowe
- c. Autor
- d. Author

Widać stąd, że możliwości użytkownika są bardzo urozmaicone. Do tego samego indeksu (np. haseł przedmiotowych) można dochodzić z wielu różnych ekranów – z ekranu „indeks”, wybierając hasło przedmiotowe lub hasło przedmiotowe ang./franc., z ekranu „słowa kluczowe”, wybierając kategorię hasło przedmiotowe, w ekranie „wyszukiwanie zaawansowane”, pisząc „s:...” czy też z ekranu „słowa w hasłach”, wybierając opcję hasło przedmiotowe. Celowo w tym krótkim opisie pominieliśmy możliwości oglądania rekordów bibliograficznych czy też własnego konta jako mniej nas interesujące.

2.3. Logi transakcyjne katalogu OPAC

Analiza rejestrów katalogu bibliotecznego OPAC to jedna z najważniejszych metod używanych przez badaczy-praktyków wyszukiwania informacji, która umożliwia analizę komunikacji użytkownika z systemem wyszukiującym. Choć metoda ta używana była w systemach bibliecznych i systemach wyszukiwania informacji [1], stała się ona jednym z najbardziej odpowiednich sposobów zbierania danych od wielu użytkowników końcowych. Rejestry z serwera często zawierają adresy IP (lub nazwy domen) wyszukiwanych, datę i godzinę wyszukiwania, komendy wyszukiwania oraz wyniki wyszukiwania. Analiza rejestrów jest użyteczna przy tworzeniu ogólnych statystyk dotyczących sposobów wyszukiwania i zachowań użytkowników.

Dane pobrane zostały z rejestru zapytań za okres czternastu dni, tj. od 25 stycznia 2006 r. do 7 lutego 2006 r., co dało w efekcie 11 692 rekordów od użytkowników korzystających z terminali komputerowych zlokalizowanych w budynku biblioteki głównej, jak również z dowolnego miejsca na świecie. Przykładowy fragment tego materiału wyglądał następująco:

Rys. 2. Przykład logów transakcyjnych

Adres IP	Identyfikator sesji	Czas	Typ wyszukiwania	Pytanie	Liczba rekordów
212.182.105.14	2006012514552426903	01:25:2006:14:56:04	[SCAN]	4\;encyklopedia łowiectwo	10
212.182.27.165	2006012515053427848	01:25:2006:15:05:49	[SCAN]	21\;Copepoda	10
212.182.27.165	2006012515053427848	01:25:2006:15:06:07	[BIBREQ]	vtls000031168	1
212.182.27.165	2006012515061527938	01:25:2006:15:06:43	[KW]	1003\;Copepoda	0
212.182.27.165	2006012515061527938	01:25:2006:15:06:53	[KW]	1035\;Copepoda	5

Komputer o adresie IP 212.182.105.14 skontaktował się 25.01.2006 r. o godzinie 14:56:04 i użytkownik w ekranie „indeks” ([SCAN]) za pomocą „tytułu” (4:\) zadał pytanie *encyklopedia łowiectwa*, w odpowiedzi na które otrzymał 10 rekordów wyświetlonych na ekranie. Jednak żaden z tych rekordów nie zainteresował użytkownika, bowiem było to jego ostatnie pytanie w sesji 2006012514552426903 (użytkownik nie przeszedł do zamawiania rekordów bibliograficznych). W ramach drugiej sesji komputer o adresie IP 212.182.27.165

wybrała w indeksie ([SCAN]) „hasło przedmiotowe” (21:\) „Copepoda” i po wybraniu jednego spośród 10 wyświetlonych rekordów w odpowiedzi zamówił pozycję o numerze vtls000031168. Wyszukiwanie to zakończyło się sukcesem w tym sensie, że użytkownik obejrzał i zamówił jedną pozycję bibliograficzną. W ramach następnej sesji komputer o tym samym numerze IP 212.182.27.165 wyszukiwał wg słów kluczowych ([KW]) w polu autor (1003\:) słowa „Copepoda” i po otrzymaniu w odpowiedzi 0 wyników (ewidentna pomyłka użytkownika) powtórzył pytanie również wg słów kluczowych, ale już teraz wybierając opcję „wszędzie”. Tak więc typ wyszukiwania to omawiany w rozdziale 2.2 interfejs użytkownika, a pytanie opisuje rodzaj operacji wykonanej przez użytkownika na tej stronie.

3. Metodologia badań

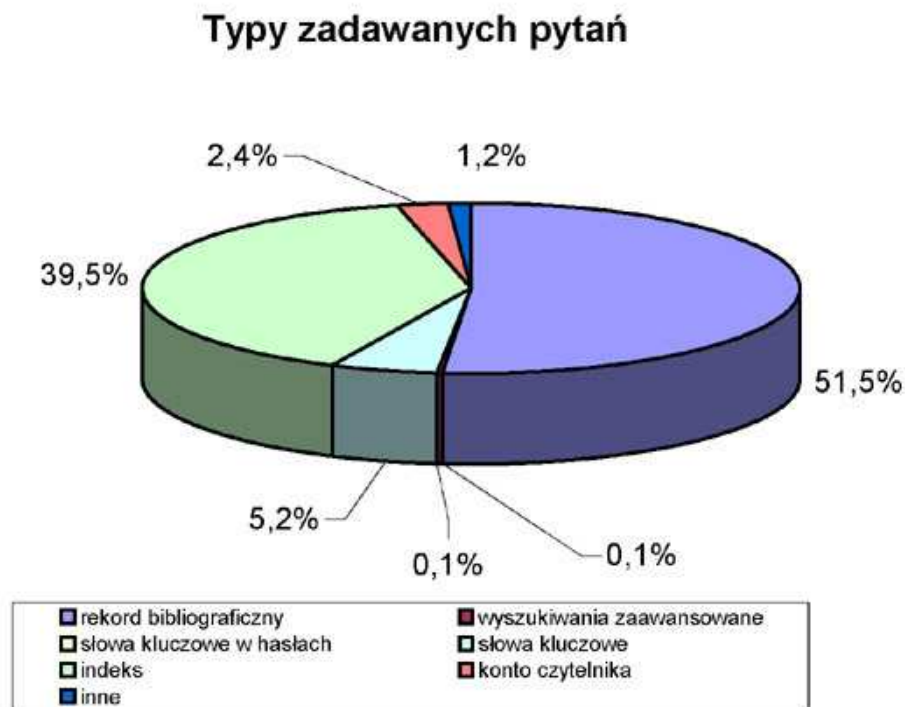
Materiał opracowano statystycznie, używając naprzemiennie języka Visual Basic for Access i Visual Basic for Excel. Pierwszy program w tabeli zawierającej dane dorzucił jeszcze trzy pola. Jedno, techniczne, był to identyfikator transakcji, w drugim wyliczono ilość rekordów bibliograficznych, jakie przejrano bezpośrednio po danej operacji wyszukiwania, przed następną operacją wyszukiwania lub przed końcem sesji. Tutaj operację przeglądania konta lub zrzutu ekranu są traktowane tak, jakby ich w ogóle nie było. Trzecim wyliczonym polem, jest długość czasu, przez jaki użytkownik obserwował wyniki wyszukiwania aż do wykonania jakiegokolwiek innej transakcji lub końca sesji. Następny program zsumował odpowiednie pola VTLS i CZAS dla wszystkich jednakowych typów i utworzył tabelę o polach: TYP jest to typ wyszukiwania (np. [BIBREQ]-rekord bibliograficzny, [SCAN]-indeks, [KW] – słowa kluczowe, [KIH]- słowa w hasłach przedmiotowych), pole TYPKWE mówi o typie kwerendy np. 1003\: to autor, 21\: to hasło przedmiotowe, ILOSC to ilość tego typu wyszukiwania, VTLS – to oglądana łącznie ilość rekordów bibliograficznych, CZASILOSC to suma czasów oglądania ekranu tego typu, a CZAS to ilość tych transakcji, dla których możliwy był pomiar czasu. Wyliczane pola EVTLS i ECZAS to przeciętna oglądana liczba rekordów bibliograficznych i przeciętny czas. Tak wyliczona tabela stanowiła podstawę do tworzenia większości wykresów i tabel.

4. Wyniki

4.1. Analiza z punktu widzenia interfejsu użytkownika

Jeśli chodzi o ilość typów przeglądanych ekranów przez użytkownika, to dominuje przeglądanie rekordów bibliograficznych (51,5%+/-0,9%) – jest to oglądanie efektów wyszukiwania. Następną grupą akcji użytkownika to wyszukiwania w indeksach (39,5%+/-0,9%). Wyszukiwanie wg słów kluczowych to tylko 5,2%+/-0,4%, przeglądanie konta czytelnika to 2,4%+/-0,3% , a pozostałe typy operacji są już bardzo rzadkie (zob. Rys.3).

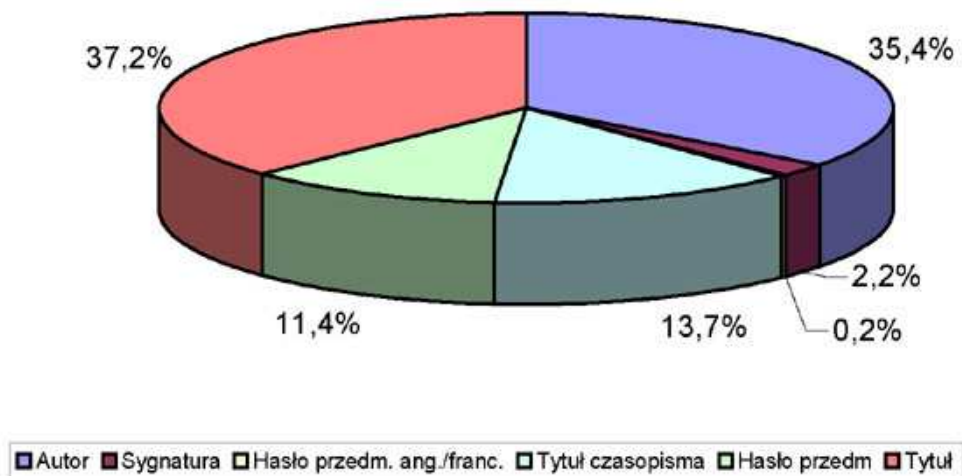
Rys. 3. Typy interfejsu użytkownika



Wyszukiwanie wg indeksu jest wyszukiwaniem podstawowym (wyszukiwania zaawansowane czy wg słów kluczowych są trudniejsze), zresztą ekran z indeksem pojawia się jako pierwszy. Ekran słowa kluczowe w hasłach (13 na 11692) czy zaawansowane (11 na 11692) to tylko bardzo niewielki ułamek wszystkich wyszukiwań. Wyniki te świadczą o małej wiedzy dotyczącej możliwości wyszukiwania przez użytkowników katalogów komputerowych.

Omówmy teraz typy wyszukiwań wg. interfejsu – indeks (Rys.4). Najczęściej, w ponad 37%(?1,4%) wyszukiwano wg. tytułu, 35,4%(+/-1,4%) wyszukiwań był to autor (proponowane przez komputer), 13,7%(+/-1%) tytuł czasopisma, 11,4%(+/-1%) hasła przedmiotowe jhp KABA, 2,2%(+/-0,4%) sygnatura a tylko 0,2%(+/-0,1%) hasła przedmiotowe ang./franc. W najczęściej używanym rodzaju wyszukiwań, wg indeksu, wyszukiwanie wg haseł przedmiotowych jest dość rzadkie, dominuje autor albo tytuł dokumentu. Dominują więc wyszukiwania znanych (podanych na zajęciach?) pozycji, a nie przeglądanie, co na dany temat napisano. Ciekawe, że ilość wyszukiwań wg Indeksu - haseł przedmiotowych (525 wyszukiwań) jest zbliżona do wszystkich wyszukiwań wg słów kluczowych (611) i dużo większa od wszystkich wyszukiwań wg słów kluczowych w hasłach (15). Użytkownik, aby znaleźć interesujące go dokumenty na zadany temat, wyszukiwał więc wg słów kluczowych czy według Indeksu - haseł przedmiotowych, a kategoria słowa kluczowe w hasłach wydaje się niezrozumiała dla użytkowników.

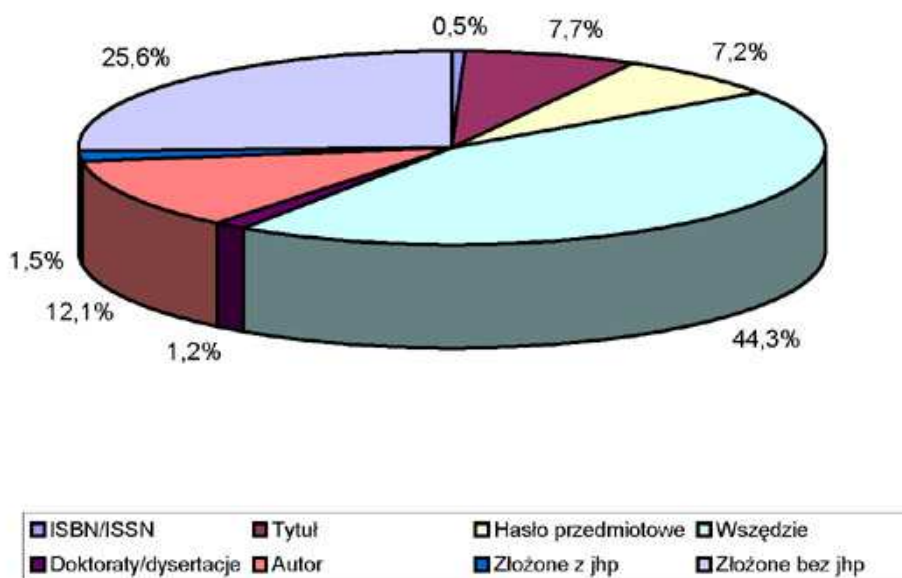
Wyszukiwania wg indeksu



Natomiast w przypadku zdecydowanie rzadszych wyszukiwań wg. słów kluczowych (Rys. 5) aż w 44,3%(+/-4%) wyszukiwano te słowa kluczowe wszędzie, w 12,1%(+/-2,6%) w indeksie autorskim, w 7,7%(+/-2,1%) w tytule, 7,2%(+/-2%) w hasłach przedmiotowych. Duża ilość występowała też wyszukiwań złożonych (łącznie 25,6%(+/-3,5%) przy czym aż w 24,8%(+/-3,5%) były to wyszukiwania w których żadna składowa nie dotyczyła jhp a 1,5%(+/-1%) były to wyszukiwania złożone w których co najmniej w jednym składniku wyszukiwano wg. jhp.

Rys. 5. Rodzaje wyszukiwań wg interfejsu – słowa kluczowe

Wyszukiwanie wg. słów kluczowych



Słowa kluczowe wyszukiwane się więc albo wszędzie, albo w kilku indeksach, lecz bez haseł przedmiotowych (ze spójnikami I, LUB, NIE lub BLISKO).

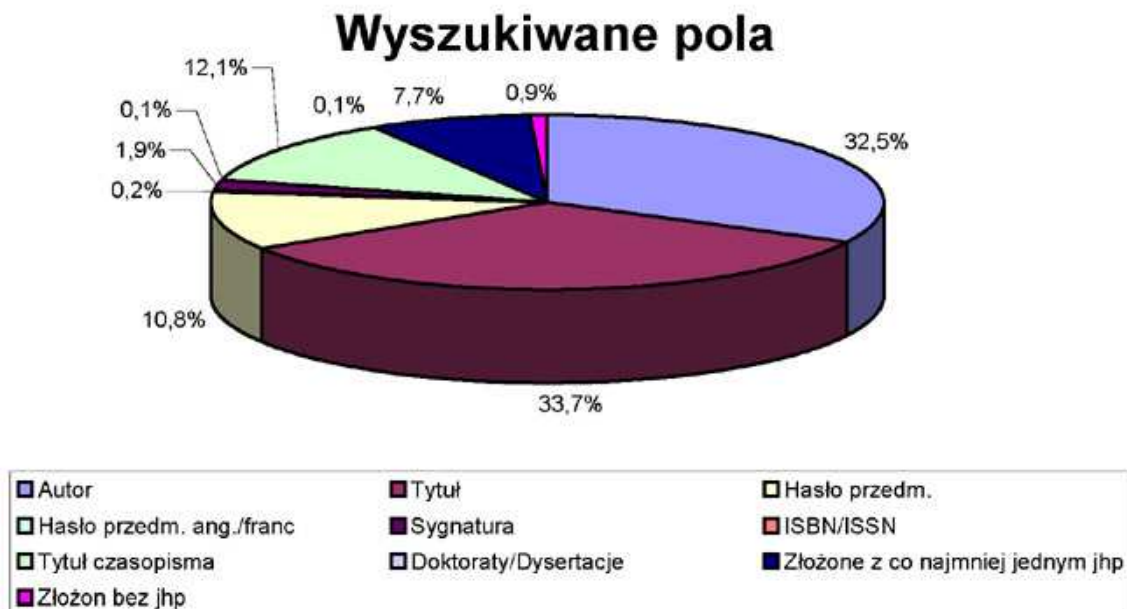
Spośród innych typów wyszukiwań 13 wyszukiwań (0,11%) były to wyszukiwania zaawansowane (w siedmiu sytuacjach wyszukiwanie wszędzie, w trzech nr ISBN/ISSN, natomiast pozostałe wyszukiwania dotyczyły autora i autora z tytułem) oraz 15 wyszukiwań dotyczyło słów w hasłach (jedno to hasło przedmiotowe, a pozostałe to proste i złożone wyszukiwania słów w tytule).

Dominują więc wyszukiwania według indeksów autorskich i tytułowych. Język haseł przedmiotowych występuje dość rzadko. W wyszukiwaniach tematycznych użytkownicy sięgają do wyszukiwań wg indeksu - hasła przedmiotowe, wg słów kluczowych - wszędzie czy też używa się słowa kluczowe - złożone jednak bez haseł przedmiotowych. Wbrew obawom wydaje się, że użytkownicy swobodnie wykorzystują hasła przedmiotowe.

4.2. Pola, według których wyszukiwano

Abstrahując od wyboru przez użytkownika, czy to indeksu, czy też słów kluczowych, a koncentrując się na tym, jakich pól rekordu MARC dotyczyły wyszukiwania, można otrzymać następujący wykres:

Rys. 6. Typy wyszukiwań pól (niezależnie od interfejsu)

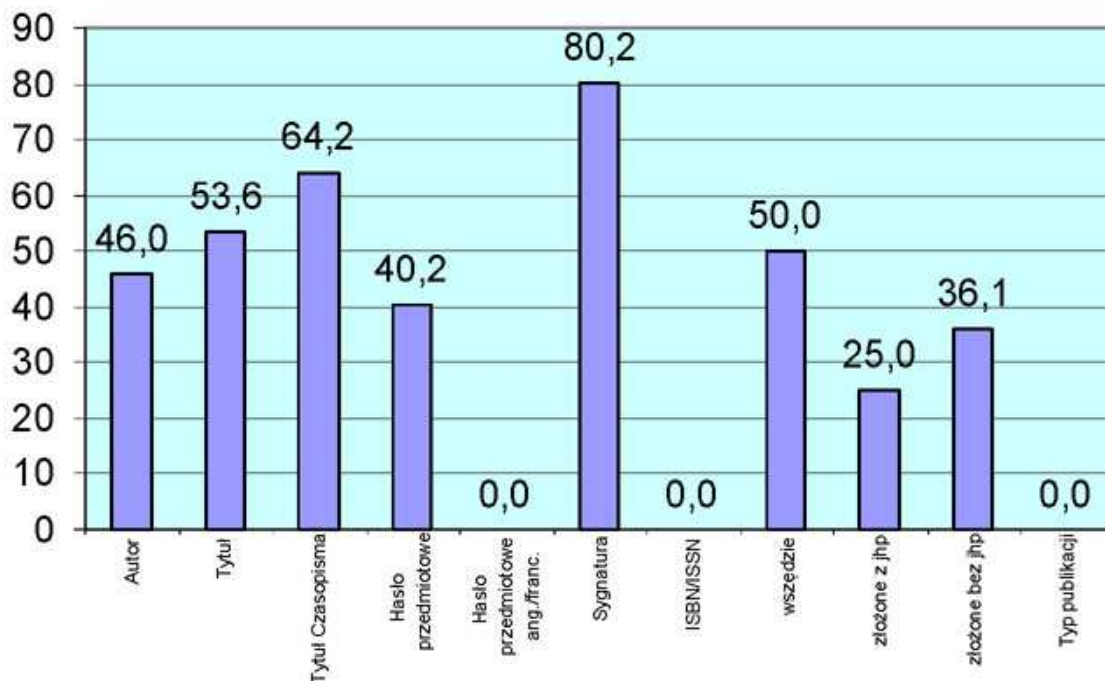


Przeważają więc spośród wyszukiwań pytania o TYTUŁ 33,7%(+/-1,3%) i o AUTORA 32,5%(+/-1,3%), następnie o TYTUŁ CZASOPISMA 12,1%(+/-0,9%), HASŁO PRZEDMIOTOWE 10,8%(+/-0,8%) oraz pytania złożone do wielu lub kilku pól ale również zawierające pytania o HASŁA PRZEDMIOTOWE, w szczególności dość liczna grupa pytań o wszystko 7,7%(+/-0,7%). Następną grupą są pytania o SYGNATURĘ 1,9%(+/-0,4%), złożone pytania do wielu pól bez jhp KABA 0,9%(+/-0,2%) oraz w granicach 0,1%(+/-0,09%) HASŁO PRZEDMIOTOWE ANG./FRANC., DOKTORATY i DYSERTACJE oraz o numer ISBN/ISSN. Hasła przedmiotowe dość rzadko interesują więc użytkowników.

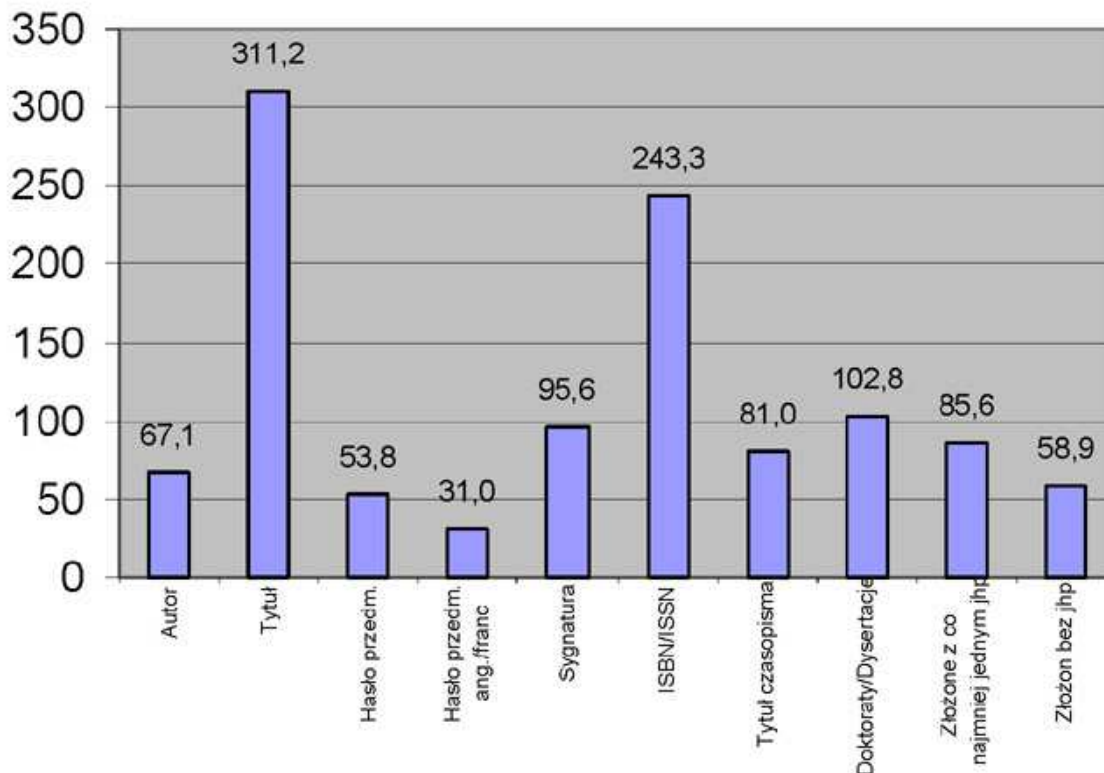
Efektywność wyszukiwania próbowaliśmy mierzyć pośrednio w następujący sposób: porównaliśmy, w ilu procentach bezpośrednio po różnych typach wyszukiwań użytkownik sięgał do rekordów bibliograficznych (przed następnym wyszukiwaniem), to znaczy dokładnie oglądał opis co najmniej jednej pozycji. Oczywiście taka efektywność nie uwzględnia sytuacji, gdy efektem wyszukiwania miała być bibliografia a nie dokument, jednak wyszukiwania bibliograficzne w tej bibliotece są rzadkie. Największą efektywność (Rys. 7) wyszukiwań daje więc wyszukiwanie wg. sygnatury. Nie jest to chyba zaskoczeniem, podając sygnaturę trafiamy bezpośrednio w poszukiwane dokumenty i potem pozostaje nam tylko zamówić te dokumenty. W dalszej kolejności najefektywniejszy jest tytuł czasopisma, mniej więcej na jednakowym poziomie efektywność wyszukiwania według tytułu i słowa kluczowe wszędzie. Efektywność wyszukiwania według Autora jest tylko nieznacznie większa niż według Hasła przedmiotowego i nieznacznie przekracza efektywność wyszukiwań złożonych bez haseł z jhp KABA. Pozostałe wyniki nie są zbyt wiarygodne ze względu na małą ilość tych typów wyszukiwań. Podsumowując efektywność wyszukiwań według haseł przedmiotowych KABA jest niska. Obejrzone hasła przedmiotowe raczej skłaniają do dalszego szukania a nie do zamawiania pozycji.

Rys. 7. Procent wyszukiwań, po których użytkownik przechodził do oglądania rekordów bibliograficznych

Procentowa zawartość wyszukiwań zakończonych sukcesem



Efektywność wyszukiwania - czas



Długi czas oglądania ekranu przez użytkownika (Rys. 8) świadczy o tym, że dużo jest pozycji interesujących tematycznie (w przeciwnym razie użytkownik odrzuciłby szybko przeglądane pozycje), ale nie ma pozycji bardzo dobrych dla użytkownika (wtedy również użytkownik szybko przeszedłby do rekordu bibliograficznego). Długi czas wyszukiwania nie mówi więc o efektywności wyszukiwania, ale raczej o trudnościach interpretacyjnych użytkownika. Analiza czasu wskazuje, że najczęściej kłopotu użytkownikom sprawia interpretacja wyników wyszukiwania za pomocą tytułu i za pomocą numeru ISBN/ISSN. Tutaj wyniki wyszukiwań hasłami przedmiotowymi należą do najszybciej interpretowalnych

4.3. Przykłady wyszukiwań wg jhp KABA

Przyjrzyjmy się teraz przykładom wyszukiwań wg jhp KABA. Podamy najpierw wyszukiwania wg haseł przedmiotowych obcojęzycznych.

Rys. 9. Wyszukiwania według haseł przedmiotowych ang./franc.

Hasło przedm. ang/franc	Ilość
project IT	2
przypadki robinsona cruzoe	1
heavy metals	1
chelate	1
chłodnictwo	1
orafi	1
Festulolium	1

Oczywiście wyszukiwanie „przypadków robinsona cruzoe” czy „chłodnictwa” za pomocą obcojęzycznych haseł przedmiotowych to poważny błąd.

Wyszukiwanie zwykłymi hasłami przedmiotowymi (podajemy hasła w oryginalnej pisowni i tylko o częstotliwości większej niż 2):

Rys. 10. Najczęstsze wyszukiwania według haseł przedmiotowych

Hasło przedmiotowe	Ilość	Poprawność hasła
analiza przedsiębiorstwa	10	Niepoprawne ("analiza przemysłowa")
przeciwutleniacze	9	Poprawne
rośliny strączkowe	8	Poprawne
byliny	6	Poprawne

chłodnictwo	6	Poprawne
mięso	6	Poprawne
mieszanie	5	Poprawne
Strączki	5	Niepoprawne ("rośliny strączkowe" lub "strączkowce")
Hodowla	5	Poprawne chociaż bardzo ogólne
geografia ekonomiczna	5	Poprawne
Wędkarstwo	5	Poprawne
Bezrobocie	4	Poprawne
kwaśne deszcze	4	Poprawne (bardzo szczegółowe tylko 1 pozycja)
rośliny jednoroczne	4	Poprawne
Suszarnictwo	4	Niepoprawne ("suszarnie" lub "suszenie")
zarządzanie finansami	4	Mały błąd ("zarządzanie finansowe")
Entomologia	3	Poprawne
Ekologia	3	Poprawne
produkcja owoców	3	Niepoprawne ("owoce - produkcja i handel -...")
Fasola	3	Poprawne
Etyka	3	Poprawne
Lizyna	3	Niepoprawne ("aminokwasy") Zbyt szczegółowe.
Chelaty	3	Poprawne ale brak pozycji w bibliotece
Akumulatory	3	Poprawne ale brak pozycji w bibliotece
Autocad	3	Poprawne
rośliny ozdobne	3	Poprawne
zatrucia pokarmowe	3	Mały błąd ("zatrucie pokarmowe")
Giełda	3	Poprawne
inżynieria chemiczna	3	Poprawne
inżynieria genetyczna	3	Poprawne
Stres	3	Poprawne
Szczeniąt	3	Mały błąd ("szczenięta")
Białka roślin	3	Mały błąd ("białka roślinne")

Można zauważyć w tym wykazie pewne wątpliwości użytkowników co do polskich diakrytyków („zarządzanie finansami”). Najczęściej używane hasło „analiza przedsiębiorstwa” nie występuje w ogóle w jhp KABA - powinno się wyszukiwać hasłem „analiza przemysłowa” (z hasłem „przedsiębiorstwo – zyski” nie ma związanych żadnych dokumentów). Następne hasła przedmiotowe na liście najczęściej używanych: „przeciwutleniacze”, „rośliny strączkowe”, „byliny”, „chłodnictwo”, „mięso” i „mieszanie” są bogato reprezentowane w jhp KABA i mają odsyłacze do wielu dokumentów. W ostatniej kolumnie zaznaczyliśmy poprawność haseł podawanych przez użytkowników (gdy były niepoprawne w nawiasie podawaliśmy najbardziej zbliżony odpowiednik). Razem z krotnościami w tym wykazie mamy 17,7% poważnych błędów, 9,2% drobnych błędów i 73,1% wyszukiwań bezbłędnych. Liczba błędów jest więc spora. Korzystanie z jhp KABA okazuje się więc trudne dla użytkowników.

5. Wnioski

Użytkownik najczęściej wyszukuje według indeksu (pierwsza strona, która mu się pojawia) i najczęściej szuka według tytułu lub autora. Te dwa typy wyszukiwań są zdecydowanie najczęstsze, o wiele rzadziej pojawia się wyszukiwanie według tytułu czasopisma i haseł przedmiotowych. W drugim co do częstości występowania typie interfejsu – słowa kluczowe, zdecydowanie najczęściej wybierana jest opcja – „wszędzie”. Hasła jhp KABA były wyszukiwane pojedynczo w 43 przypadkach w wyszukiwaniach wg słów kluczowych, a ponadto stanowiły w dziewięciu przypadkach część pytań złożonych, w 525 przypadkach wyszukiwania wg indeksu – hasła przedmiotowe i w jednym przypadku wyszukiwanie słów w hasłach, natomiast hasła ang./franc. zostały osiem razy użyte w indeksie na 11 692 wykonane wszystkie transakcje czy też na 5 254 operacji samego wyszukiwania. Łącznie stanowi to 5,01% wszystkich transakcji i 10,8% operacji wyszukiwania.

Język haseł przedmiotowych wydaje się dość trudny, tylko niewielka ilość użytkowników posługuje się nim swobodnie. Rzadko też wyszukiwania według haseł przedmiotowych od razu prowadzą do zamawiania dokumentów, zazwyczaj następuje dalsze wyszukiwanie. Hasła przedmiotowe są więc najczęściej etapem pośrednim. Szczególnie dobrze widać problemy w posługiwaniu się hasłami przedmiotowymi na podanych przykładach - ponad 17,7% użycia haseł przedmiotowych to są błędy. Zwykły użytkownik ma kłopoty w posługiwaniu się hasłami przedmiotowymi, często popełnia błędy i raczej unika tego typu wyszukiwań, ale z drugiej strony jest bardzo niewielka liczba użytkowników prawie wyłącznie używających jhp KABA.

Przypisy:

[1] PETERS, T. The history and development of transaction log analysis. *Library Hi Tech* 1993, 42(11), s. 41-66.

Efektywność wyszukiwawcza w języku haseł przedmiotowych KABA i innych językach wyszukiwawczych / Andrzej Krajka, Artur Sciborek // W: Opracowanie przedmiotowe dokumentów z zakresu nauk ścisłych: matematyczno-przyrodniczych i technicznych. Język haseł przedmiotowych KABA: teoria, praktyka, przyszłość. Kazimierz Dolny, 20-22 września 2006 roku. - [Warszawa] : Stowarzyszenie Bibliotekarzy Polskich, K[omisja] W[ydawnictw] E[lektronicznych], Redakcja "Elektronicznej Biblioteki", 2006. - (EBIB Materiały konferencyjne nr 15). - ISBN 83-921757-6-X. -Tryb dostępu : http://www.ebib.info/publikacje/matkonf/kaba/krajka_sciborek.php