

Ewa Białek

Badanie zapytań kierowanych do wyszukiwarek internetowych jako element zarządzania witryną www

Ekonomiczne Problemy Usług nr 48, 201-209

2010

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

EWA BIAŁEK

Wyższa Hanzeatycka Szkoła Zarządzania w Słupsku

BADANIE ZAPYTAŃ KIEROWANYCH DO WYSZUKIWAREK INTERNETOWYCH JAKO ELEMENT ZARZĄDZANIA WITRYNĄ WWW

Wprowadzenie

Internet to zdecentralizowana sieć sieci, składająca się z równoprawnych węzłów, przekazujących sobie wzajemne informacje zgodnie z określonymi procedurami¹. Pod koniec ubiegłego wieku, w oparciu o Internet, zbudowana została sieć WWW. Wyszukiwanie, utożsamiane przez wielu z ruchem w sieci WWW, stało się siłą napędową wszystkiego, co ma związek z Internetem. Obecnie miliony osób, zasiadając każdego dnia przed monitorami komputerów, wpisują w okna wyszukiwarek zapytania, by po sprawdzeniu kilku wyników wyszukiwania odnaleźć poszukiwane informacje. Wyszukiwarki internetowe stają się więc interfejsem do ogromnego bogactwa sieci – witryn WWW². Mimo że z założenia wyszukiwarki internetowe służą wyszukiwaniu informacji, stanowią też doskonałe źródło danych niezbędnych do zarządzania witryną internetową.

Natura wyszukiwania sprawia, że jest to jeden z ciekawszych problemów badawczych w informatyce. Twierdzi się, że prowadzone badania pozwolą w przyszłości na stworzenie komputerów potrafiących działać jak istoty ludzkie³.

¹ T. Maciejowski, *Firma w internecie*, OE, Kraków 2004.

² J. Nielsen, H. Loranger, *Optymalizacja funkcjonalności serwisów internetowych*, Helion, Warszawa 2007.

³ J. Battle, *Jak Google i konkurencja wywołali biznesową i kulturową rewolucję*, PWN, Warszawa 2007.

1. Sformułowanie problemu

Z roku na rok zmienia się nie tylko liczba użytkowników Internetu, lecz także ich zachowania. Dziś z komputerów korzysta 95% przedsiębiorstw i 59% gospodarstw domowych. Dostęp do Internetu posiada 93% firm oraz 48% gospodarstw. Internauci najczęściej korzystają z sieci w celu prowadzenia korespondencji za pomocą poczty elektronicznej (38%), wyszukiwania informacji (33%), udziału w czatach i forach dyskusyjnych (31%)⁴.

Użytkownicy Internetu, szukając czegoś konkretnego, przeważnie używają w zapytaniu więcej niż jednego słowa kluczowego, gdyż stosowanie trój-, cztero- i pięciowyrazowych wyrażań w wyszukiwarkach pomaga uzyskać dokładniejsze wyniki⁵. Podczas Usability Week 2008 Jakob Nielsen przedstawił dane pochodzące z badań zapytań użytkowników przeprowadzonych w latach 1994, 1997, 2004 i 2008. Okazuje się, że w 1994 roku wyrażenia zawierające jedno słowo stanowiły aż 91% zapytań, a w roku 2008 już tylko 20%. W tym samym czasie ilość zapytań cztero- i pięciowyrazowych znacząco wzrosła: w roku 1994: 1% – czterowyrazowe, 0% – pięciowyrazowe, w roku 2008: 16% – czterowyrazowe, 18% – pięciowyrazowe⁶. Firma iProspect opublikowała informacje, z których wynika, że w 2002 r. ponad połowę wszystkich zapytań stanowiły wyrażenia jednowyrazowe (51%), natomiast w 2004 r. największy procent stanowiły wyrażenia dwuwyrazowe (49%)⁷. Najnowsze badania Hitwise, firmy Experian⁸ pokazują wzrostową tendencję długości zapytań w ciągu ostatniego roku. Ilość długich zapytań (pięcio- i więcej wyrazowych) w okresie od stycznia 2008 r. do stycznia 2009 r. wzrosła średnio o 12%. W tym samym czasie ilość krótszych zapytań zmalała średnio o 2%.

Przytoczone powyżej wyniki badań dotyczą zapytań kierowanych do wyszukiwarek w języku angielskim. Obserwując polski rynek, można zauważyć, że na etapie projektowania, tworzenia i promocji witryny internetowej często zapomina się o potrzebie prowadzenia badań nad słowami kluczowymi, uwzględniającymi język poszczególnych stron tej witryny. Biorąc pod uwagę specyfikę języka polskiego, niezbędne jest przeprowadzenie badań dotyczących zapytań do wyszukiwarek internetowych zadawanych w tym języku. Należy uzyskać odpowiedzi na pytania: Ile wyrazów wpisują internauci, tworząc zapytania do wyszukiwarek, kierujące ich na polskojęzyczne witryny WWW? Jaka jest tendencja długości tych zapytań? Jaka jest ich skuteczność?

⁴ <http://www.stat.gov.pl/>, *Wykorzystanie technologii informacyjno-telekomunikacyjnych w przedsiębiorstwach, gospodarstwach domowych i przez osoby prywatne w 2008 r.* (stan na dzień 1.10.2009 r.).

⁵ S. Turow, *Pozycjonowanie w wyszukiwarkach internetowych*, Helion, Warszawa 2004.

⁶ <http://www.ittechnology.us/dane-z-badania-jakoba-nielsena/> (stan na dzień 1.10.2009 r.).

⁷ T. Frontczak, *Marketing internetowy w wyszukiwarkach*, Helion, Warszawa 2006.

⁸ <http://www.hitwise.com/us/press-center/press-releases/google-searches-jan-09/> (stan na dzień 1.10.2009 r.).

2. Opis badań

Możliwości technologiczne Internetu pozwalają na prowadzenie z jego pomocą badań zarówno ruchu jak i zachowań użytkowników w sieci. Rozróżnia się dwa typy tych badań: od strony serwera (*site/server centric*) i od strony użytkownika (*user centric*)⁹. Uznano, że dla przeprowadzenia badań związanych z długością zapytań kierowanych do wyszukiwarek internetowych odpowiednie będą badania od strony serwera. Ponieważ dla badań tego typu niezbędne jest stworzenie witryn WWW i umieszczenie ich w sieci Internet (badanie opiera się na aktywności serwera), za obszar badań wybrano pięć, ze wszystkich 75, niekomercyjnych witryn w domenie ewa.bicom.pl. Jedna z witryn jest sześćioletnia, dwie czteroletnie i dwie trzyletnie. Wszystkie witryny zbudowano w oparciu o statyczne strony WWW. Do promocji tych witryn w sieci Internet nigdy nie wykorzystywano płatnych form reklamy.

Ponieważ dla wybranych do badań witryn najlepszym źródłem odwiedzin jest wyszukiwarka Google i popularność jej wśród internautów utrzymuje się na bardzo wysokim poziomie – we wrześniu 2009 roku jej udział w liczbie wizyt dokonywanych z terenu Polski wyniósł 97%¹⁰, do realizacji badań wybrano:

- wyszukiwarkę Google,
- statystyki Google Analytics.

Dane przedstawione w badaniu obejmują dwa okresy:

- **a** – od 1 grudnia 2007r. do 31 stycznia 2008 r.,
- **b** - od 1 grudnia 2008r. do 31 stycznia 2009 r.

Przy pomocy wybranych narzędzi, dla każdej z witryn biorących udział w badaniu, uzyskano następujące dane:

- 1) Charakterystyka witryny:
 - Lp. – numer przyporządkowany witrynie, stanowiący jej identyfikator,
 - tematyka – opis witryny,
 - wiek – okres czasu, jaki upłynął od umieszczenia witryny na serwerze WWW,
 - ilość stron – liczba pojedynczych stron WWW wchodzących w skład witryny.
- 2) Odwiedziny:
 - wszystkie – całkowita liczba zdarzeń polegających na wyświetleniu strony WWW wchodzącej w skład witryny,
 - z wyników wyszukiwania – liczba odwiedzin, które pochodzą z wyników wyszukiwarek internetowych,
 - źródło Google/organic – procent odwiedzin z wyników wyszukiwania, pochodzących z organicznych (naturalnych) wyników wyszukiwarki Google.

Zestawiono je w tab. 1.

⁹ P. Sztabiński, F. Sztabiński, Z. Sawiński, *Nowe metody, nowe podejścia badawcze w naukach społecznych*, Wyd. IFiS PAN, Warszawa 2004.

¹⁰ Biuletyn ranking.pl z dnia 30.09.2009.

Tabela 1. Charakterystyka witryn w domenie ewa.bicom.pl, stan z dnia 1.10.2009 r.

Lp.	Charakterystyka witryny			Odwiedziny					
	tematyka	wiek	ilość stron	wszystkie		z wyników wyszukiwania		źródło Google/organic	
				a	b	a	b	a	b
1	chwile z poezją	6 lat	102	2 653	1 766	905	975	91%	92%
2	życzenia ślubne	4 lata	40	1 379	2 041	952	1 638	65%	77%
3	życzenia dobranoc	4 lata	30	8 570	11 388	7 715	10 719	86%	92%
4	piosenki biesiadne	3 lata	78	8 917	11 192	6 993	9 580	76%	80%
5	poezja Kobryńczuk	3 lata	37	555	466	282	348	51%	75%

Źródło: opracowanie własne.

W tab. 2 i tab. 3 przedstawiono wyniki dotyczące zapytań, które skierowały internautów na poszczególne witryny biorące udział w badaniu:

Zapytania:

- Lp. – numer przyporządkowany witrynie, stanowiący jej identyfikator,
- wszystkie (liczba): – liczba wszystkich zapytań, za pośrednictwem których doszło do odwiedzenia witryny,
- wszystkie (liczba odwiedzin): – liczba odwiedzin witryny, do których doszło za pośrednictwem wszystkich zapytań (równa liczbie odwiedzin z wyników wyszukiwania – tabela 1),
- jeden wyraz (ilość %) – % wszystkich zapytań, za pośrednictwem których doszło do odwiedzenia witryny przy użyciu zapytań jednowyrazowych,
- jeden wyraz (% odwiedzin) – % wszystkich odwiedzin, do których doszło przy użyciu zapytań jednowyrazowych.

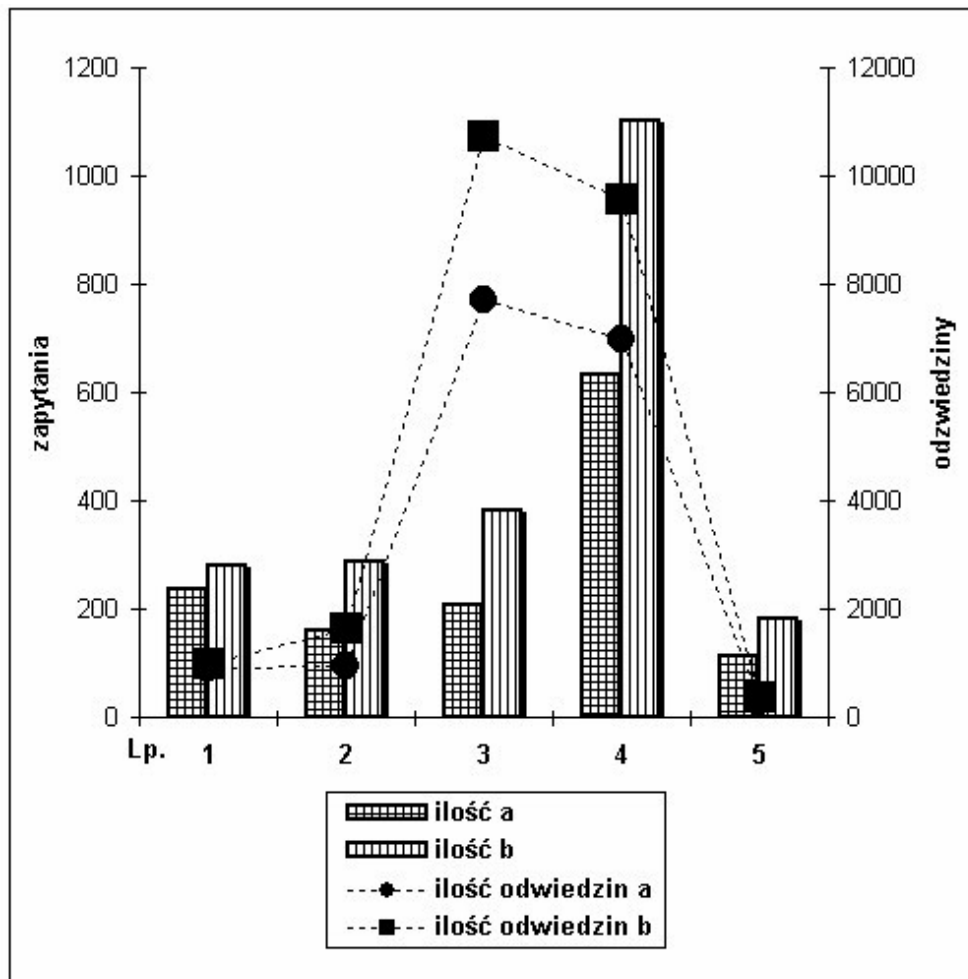
Dla zapytań dwu-, trój-, cztero- itd. wyrazowych przyjmuje się analogiczny opis, jak dla wyrażeń jednowyrazowych.

Tabela 2. Wyniki badań wszystkich zapytań dla witryn 1-5

Lp.		1		2		3		4		5		
Okres		a	b	a	b	a	b	a	b	a	b	
Zapytania	wszystkie	ilość	235	280	161	289	206	380	634	1 102	111	183
		ilość odwiedzin	905	975	952	1 638	7 715	10 719	6 993	9 580	282	348

Źródło: opracowanie własne.

Tab. 2 oraz rys. 1 pokazują liczbę oraz skuteczność (liczbę odwiedzin) wszystkich zapytań w wybranych okresach a i b. Za skuteczne uważa się to zapytanie, które powoduje generowanie odwiedzin na witrynie WWW.



Rys. 1. Liczba oraz skuteczność wszystkich zapytań

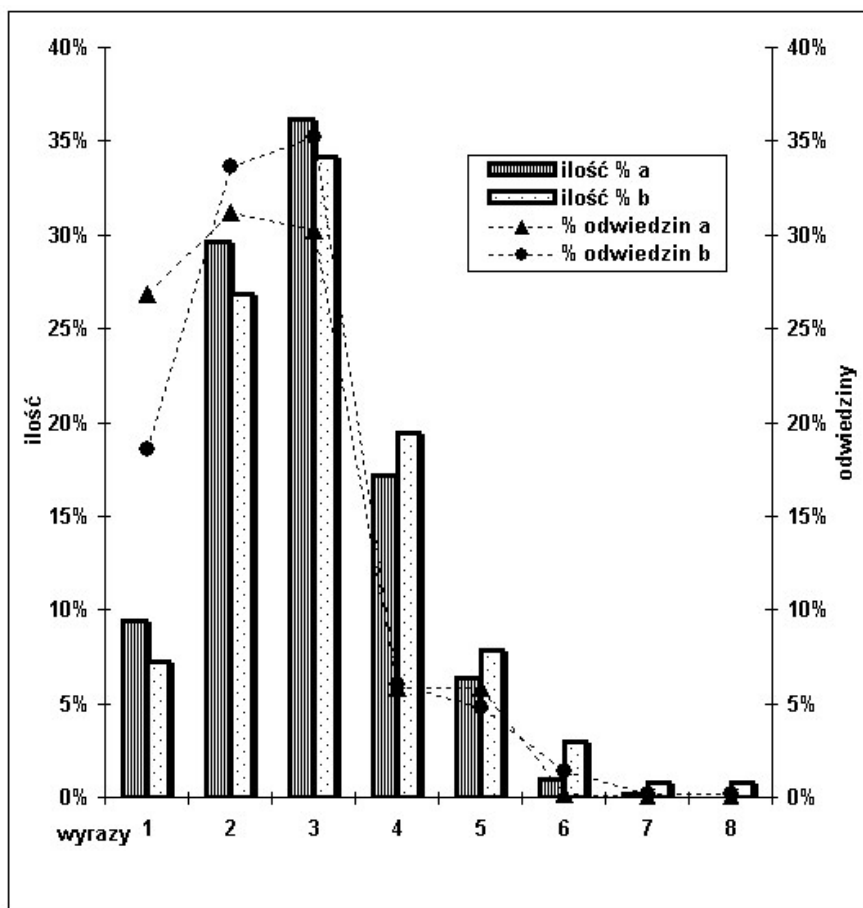
Źródło: opracowanie własne.

Tabela 3. Wyniki badań zapytań według liczby użytych w nich wyrazów dla witryn 1-5

Lp.		1.		2.		3.		4.		5.		Średnia		
Okres		a	b	a	b	a	b	a	b	a	b	a	b	
Zapytania	jeden wyraz	ilość %	11%	7%	7%	6%	19%	14%	6%	4%	4%	5%	9%	7%
		% odwiedzin	24%	21%	16%	5%	64%	51%	23%	13%	7%	3%	27%	19%
	dwa wyrazy	ilość %	33%	28%	32%	31%	26%	26%	26%	24%	31%	25%	30%	27%
		% odwiedzin	41%	42%	19%	30%	3%	3%	59%	65%	34%	28%	31%	34%
	trzy wyrazy	ilość %	34%	33%	42%	45%	35%	33%	38%	40%	32%	20%	36%	34%
		% odwiedzin	22%	24%	61%	61%	32%	45%	13%	16%	23%	30%	30%	35%
	cztery wyrazy	ilość %	16%	23%	15%	13%	16%	21%	21%	23%	18%	17%	17%	19%
		% odwiedzin	11%	10%	3%	3%	1%	1%	4%	4%	10%	12%	6%	6%
	pięć wyrazów	ilość %	4%	7%	3%	4%	3%	4%	7%	6%	15%	18%	6%	8%
		% odwiedzin	1%	2%	1%	1%	0%	0%	1%	1%	26%	20%	6%	5%
	sześć wyrazów	ilość %	2%	2%	1%	1%	1%	1%	1%	2%	0%	9%	1%	3%
		% odwiedzin	1%	1%	0%	0%	0%	0%	0%	1%	0%	5%	0%	1%
	siedem wyrazów	ilość %	0%	0%	0%	0%	0%	0%	1%	1%	0%	3%	0%	1%
		% odwiedzin	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%
	osiem wyrazów	ilość %	0%	0%	0%	0%	0%	1%	0%	0%	0%	3%	0%	1%
		% odwiedzin	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%
	pozostałe	ilość %	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
		% odwiedzin	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Źródło: opracowanie własne.

W tab. 3 pokazano liczbę oraz skuteczność zapytań o określonej długości w okresach a i b. Wartości średnie przedstawiono w postaci wykresu (rys. 2).



Rys. 2. Liczba oraz skuteczność zapytań określonej długości

Źródło: opracowanie własne.

3. Opis wyników i wnioski końcowe

Z danych przedstawionych w tab. 1 wynika, że:

- głównym źródłem odwiedzin witryn znajdujących się w domenie ewa.bicom.pl są wyszukiwarki internetowe, wśród nich Google, której udział w kierowaniu internautów na tę witrynę w ciągu badanego okresu wzrósł,
- na większości witryn zauważa się wzrost liczby wszystkich odwiedzin,
- spadek liczby odwiedzin na pozostałych witrynach mógł zostać spowodowany faktem, że w badanym okresie przepustowość łącza dla opisywanych witryn nie wzrosła (zarówno w okresie **a**, jak i w okresie **b** w godzinach największej oglądalności było ono obciążone niemal w 100%).

Analizując wyniki z tab. 2, przedstawione na wykresie (rys. 1) zauważa się, że:

- wzrosła liczba wszystkich zapytań kierujących internautów na poszczególne witryny przedstawione w badaniu,
- wzrosła również liczba wszystkich odwiedzin, których źródłem były wyszukiwarki internetowe, co jest dowodem prawidłowego przebiegu procesu promocji tych witryn w Internecie.

Wyniki przedstawione w tab. 3, umieszczone na wykresie 2 (rys. 2) pokazują, że:

- największą grupę w obu badanych okresach stanowiły wyrażenia trójwyrazowe,
- najdłuższe z zanotowanych wyrażeń miały osiem słów,
- zmalała liczba zapytań jedno-, dwu- i trójwyrazowych,
- wzrosła liczba zapytań długich: cztero-, pięcio-, sześćo-, siedmio- i ośmiowyrazowych,
- zauważa się widoczny spadek liczby odwiedzin generowanych przez wyrażenia jednowyrazowe,
- mimo spadku liczby wyrażeń dwu- i trójwyrazowych ich udział w kierowaniu internautów na witryny wzrósł, co świadczy o bardzo dużej ich skuteczności.

Słowa kluczowe, występujące na witrynach WWW, z których swoje zapytania budują internauci, są podstawą rankingów wyszukiwarek internetowych¹¹. Wyniki badań przedstawione powyżej powinny więc być brane pod uwagę w procesie projektowania, tworzenia, promocji, a przede wszystkim zarządzania witryną WWW. Jakie będą z tego wymierne korzyści i jak zostaną wykorzystane, zależy od kreatywności osób zajmujących się tymi zagadnieniami.

Podsumowanie

W artykule przedstawiono wyniki badań dotyczących zapytań do wyszukiwarek, które kierują użytkowników internetu na polskojęzyczne witryny WWW. Szczególną uwagę zwrócono na tendencje wzrostowe liczby słów tworzących te zapytania oraz liczby wizyt na witrynie pochodzących z długich zapytań. Jako obszar badań wybrano pięć witryn internetowych, których głównym źródłem odwiedzin są wyszukiwarki. Dane do badań w postaci słów kluczowych, liczby i źródeł odwiedzin na opisywanych witrynach pochodzą ze statystyk Google.

Zwrócono uwagę, że głównym elementem uwzględnianym przez wyszukiwarki do tworzenia swoich rankingów są słowa kluczowe, znajdujące się na witrynie internetowej. Fakt ten powoduje, iż prowadzenie badań związanych ze słowami kluczowymi jest niezbędnym elementem zarządzania witryną WWW.

¹¹ E. Białek, *Teoretyczne i praktyczne aspekty rozwoju regionalnego, Kryteria oceny relewantności wyników wyszukiwania*, WHSZ Słupsk, Słupsk 2007.

Literatura

- Battle J., *Jak Google i konkurencja wywołali biznesową i kulturową rewolucję*, PWN, Warszawa 2007.
- Białek E., *Teoretyczne i praktyczne aspekty rozwoju regionalnego, Kryteria oceny relewantności wyników wyszukiwania*, WHSZ Słupsk, Słupsk 2007.
- Frontczak T., *Marketing internetowy w wyszukiwarkach*, Helion, Warszawa 2006.
- Maciejowski T., *Firma w internecie*, OE, Kraków 2004.
- Nielsen J., Loranger H., *Optymalizacja funkcjonalności serwisów internetowych*, Helion, Warszawa 2007.
- Turow S., *Pozycjonowanie w wyszukiwarkach internetowych*, Helion, Warszawa 2004.
- Sztabiński P., Sztabiński F., Sawiński Z., *Nowe metody, nowe podejścia badawcze w naukach społecznych*, Wyd. IFiS PAN, 2004.
- Biuletyn ranking.pl z dnia 30.09.2009.
- <http://www.hitwise.com/us/press-center/press-releases/google-searches-jan-09/> (stan na dzień 1.10.2009 r.).
- <http://www.ittechnology.us/dane-z-badania-jakoba-nielsena/> (stan na dzień 1.10.2009 r.).
- <http://www.stat.gov.pl/>, Wykorzystanie technologii informacyjno-telekomunikacyjnych w przedsiębiorstwach, gospodarstwach domowych i przez osoby prywatne w 2008 r. (stan na dzień 1.10.2009 r.).

STUDY ON THE INQUIRIES REFERRED TO INTERNET SEARCH ENGINES AS AN ELEMENT OF WWW PORTAL MENAGEMENT

Summary

In the article results concerning the study on the inquiries referred to www searching engines redirecting internet users to Polish language www portal sites are presented. Particular attention has been paid to the growing tendencies of the number of words used to formulate such inquires and the number of entries to a site when the inquiry is rather long. As the study area five portals have been selected the entries in which are motivated mainly by the intention of using search engines. The data for the study, as keywords, the number and entry sources concerning the described portals originate from the Google statistics.

It has been emphasised that the main element taken into consideration by search engines when they form their own rankings are keywords occurring in the portal. This makes the study related to keywords an indispensable element of www portal management.

Translated by Ewa Bialek