

Ewa Białek*

Krzysztof Białek*

DŁUGI OGON WYSZUKIWANIA I JEGO WPŁYW NA OGLĄDALNOŚĆ WITRYNY WWW

Streszczenie

*W artykule przedstawiono wyniki badań przeprowadzonych metodą site/server centric w oparciu o istniejące od kilku lat w sieci internet witryny WWW i wyszukiwarkę **Google**. Pokazano, że przy projektowaniu, optymalizacji i promocji witryny należy skupić się na najpopularniejszych słowach kluczowych związanych z tą witryną, gdyż to one mają największy wpływ na jej oglądalność. Nie należy jednak zapominać o istnieniu długiego ogona, który stanowi dodatkowe źródło odwiedzin.*

1. Wprowadzenie

Długi ogon (ang. *The Long Tail*) to koncepcja sformułowana w 2004r. przez Chrisa Andersona, który opisuje pewne modele biznesowe oraz ekonomiczne, zaobserwowane w działalności takich firm jak Amazon czy Netflix. Chris Anderson w artykule dla magazynu Wired stwierdza, że posiadanie bardzo szerokiego asortymentu jest w stanie wygenerować na pojedynczych, rzadko poszukiwanych pozycjach, sumarycznie większych obrotów od tych, osiąganych na najpopularniejszych, masowo sprzedawanych towarach[1]. Koncepcja długiego ogona jest niejako zaprzeczeniem zasady włoskiego ekonomisty Vilfredo Pareto, która głosi, iż firmy uzyskują ok. 80% przychodów dzięki sprzedaży jedynie 20% oferowanego przez nie asortymentu [2].

* mgr inż., Wyższa Szkoła Zarządzania, Katedra Informatyki,
ul. Kozińskiego 6, 76-200 Słupsk, ewa@bicom.pl

* dr inż., Wyższa Szkoła Zarządzania, Katedra Informatyki,
ul. Kozińskiego 6, 76-200 Słupsk, krzysztof@bicom.pl

2. Sformułowanie problemu

Z najnowszych badań wynika, że dostęp do sieci internet posiada już 92% przedsiębiorstw i 41% gospodarstw domowych [3]. 71% internautów poproszonych o wskazanie czynności wykonywanej w przeddzień badania wybrało przeglądanie zawartości stron internetowych [4]. Strona WWW staje się więc najprostszym i najbardziej efektywnym sposobem dotarcia z ofertą firmy do potencjalnego klienta. Stworzenie witryny i umieszczenie jej na serwerze WWW to jednak nie wszystko. Spełni ona swoją rolę tylko wtedy, gdy odwiedzą ją internauci.

Badania prowadzone w internecie umożliwiają poznanie schematów postępowania, zachowań i logiki, jaką posługują się potencjalni klienci podczas przeszukiwania sieci [5]. Pokazują one, że największa część internautów trafia na witrynę dzięki wyszukiwarkom internetowym [6]. Niezbędnym jest jednak, aby witryna zajęła wysokie pozycje w rankingach tych wyszukiwarek [7].

Analizując statystyki odwiedzin stron WWW z wyszukiwarek internetowych zauważa się, że wizyty pochodzą nie tylko z zapytań (słów kluczowych), na które prowadzona jest najintensywniejsza promocja witryny. Część odwiedzin generowana jest przez wiele innych, mniej popularnych słów kluczowych [7]. Zjawisko to można nazwać **długim ogonem wyszukiwania**.

Z badań wynika, że każde z mniej popularnych zapytań osobno powoduje niewielki ruch na witrynie [7]. Czy zebrane razem są w stanie decydować o popularności witryny? Czy uzasadnione jest wykorzystanie długiego ogona wyszukiwania w celu zwiększenia oglądalności witryny WWW?

3. Opis badań

Internet ze względu na możliwości technologiczne umożliwia prowadzenie precyzyjnych badań ruchu oraz zachowań użytkowników w sieci. Badania te dzielą się na dwa podtypy: od strony serwera (*site/server centric*) i od strony użytkownika (*user centric*). W przypadku projektów *site/server centric* badanie obejmuje pełną populację użytkowników internetu, a analizie poddawane są wszystkie połączenia dokonane z wybraną witryną WWW [8]. Uznano więc, że dla przeprowadzenia badań dotyczących zagadnień związanych z długim ogonem wyszukiwania typ *site/server centric* będzie odpowiedni.

3.1. Opis obszaru badań

Dla badań typu *site/server centric* niezbędnym jest stworzenie przeznaczonych do tego witryn i umieszczenie ich w internecie (badanie opiera się na aktywności serwera). Za obszar badań wybrano więc 10 niekomercyjnych witryn WWW w domenie **ewa.bicom.pl**. Witryny te powstawały sukcesywnie od roku 2003 w celach dydaktyczno naukowych. Ze względu na różnorodność

umieszczonych na nich materiałów, grupa docelowa tych witryn jest bardzo szeroka - od dzieci w wieku przedszkolnym, poprzez młodzież do osób dorosłych.

3.2. Wybór narzędzi

Z opublikowanych danych wynika, że internetowy świat podbija dziś **wyszukiwarka Google**. Również w naszym kraju serwis wyszukiwawczy **Google** cieszy się ogromną popularnością. Poziom jego wykorzystania wciąż wzrasta. Zgodnie z wynikami badania gemiusTraffic, na początku bieżącego roku udział **Google** w liczbie wizyt z wyszukiwarek dokonywanych przez internautów z Polski na polskich witrynach równy był **89,6%**. - zaś w ostatnio sięgnął **90,7%** [9].

Uwzględniając powyższe dane do realizacji badań wybrano następujące narzędzia:

- statystyki **Webalizer Version 2.01**,
- wyszukiwarkę **Google**,
- **Google Analytics**.

3.3. Założenia

Przedstawione dane obejmują okres od **1 grudnia 2007r.** do **31 stycznia 2008r.** W tym czasie na wszystkich **75** witrynach w domenie **ewa.bicom.pl** zanotowano **703 735** odwiedzin. Taką ilość odwiedzin uznano za wystarczającą do przeprowadzenia badań.

Do przeprowadzenia eksperymentu wybrano **10** witryn: jedną najstarszą (5 lat), jedną najmłodszą (1 rok), dwie 4-letnie, dwie 2-letnie i cztery 3-letnie. Witryny 3-letnie reprezentują największą z grup, gdyż na ten okres przypada najintensywniejszy przyrost ilości witryn w domenie **ewa.bicom.pl**. Do promocji tych witryn w sieci internet nigdy nie wykorzystywano płatnych form reklamy.

Wszystkie wybrane witryny zbudowano w oparciu o **statyczne strony HTML**, które przy indeksowaniu przez wyszukiwarki traktowane są jako osobne obiekty. Każda z tych stron może stanowić zatem stronę wejściową i tym samym przyczynić się do wzrostu popularności całej witryny.

3.4. Opis eksperymentu

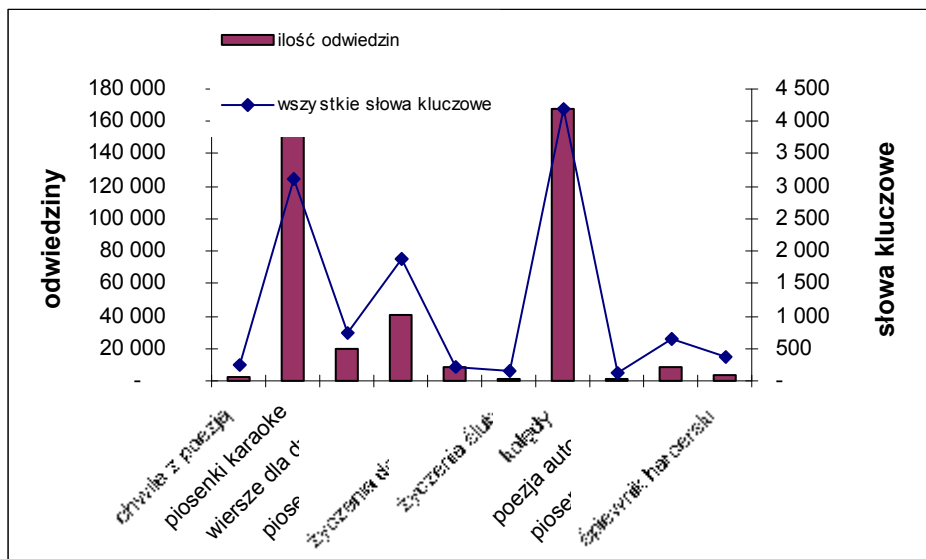
Przy pomocy narzędzi wybranych do realizacji badań, dla każdej z wybranych witryn uzyskano następujące dane:

- **Charakterystyka witryny:**
 - **wiek** – okres czasu, jaki upłynął od umieszczenia witryny na serwerze WWW,
 - **ilość stron** – liczba pojedynczych stron WWW wchodzących w skład witryny,
 - **ilość odwiedzin** – liczba zdarzeń polegających na wyświetleniu strony WWW wchodzącej w skład witryny.
- **Źródła odwiedzin:**
 - **wyszukiwarki** – procent odwiedzin, które pochodzą z wyników wyszukiwarek internetowych,
 - **bezpośrednie** – procent odwiedzin, do których doszło na skutek wpisania adresu URL bezpośrednio do przeglądarki,
 - **witryny odsyłające** – procent odwiedzin pochodzących z innych witryn internetowych.
- **Słowa kluczowe:**
 - **wszystkie** – liczba słów kluczowych, za pośrednictwem których doszło do odwiedzenia witryny,
 - **najpopularniejsze** – słowo kluczowe, przy użyciu którego wygenerowana została największa ilość odwiedzin,
 - **% odwiedzin** – procent odwiedzin wygenerowanych przy użyciu najpopularniejszego słowa kluczowego,
 - **pozycja w Google** – pozycja, jaką zajmuje witryna WWW w naturalnych wynikach wyszukiwarki **Google** dla najpopularniejszego słowa kluczowego,
 - **20%** – liczba równa 20 procentom wszystkich słów kluczowych, za pośrednictwem których doszło do odwiedzenia witryny,
 - **% odwiedzin dla 20%** - procent odwiedzin wygenerowanych przez 20 procent wszystkich słów kluczowych (posortowanych malejąco wg popularności),
 - **ilość wyników w Google** – liczba wyników, jaką zwraca wyszukiwarka **Google** dla najpopularniejszego słowa kluczowego.

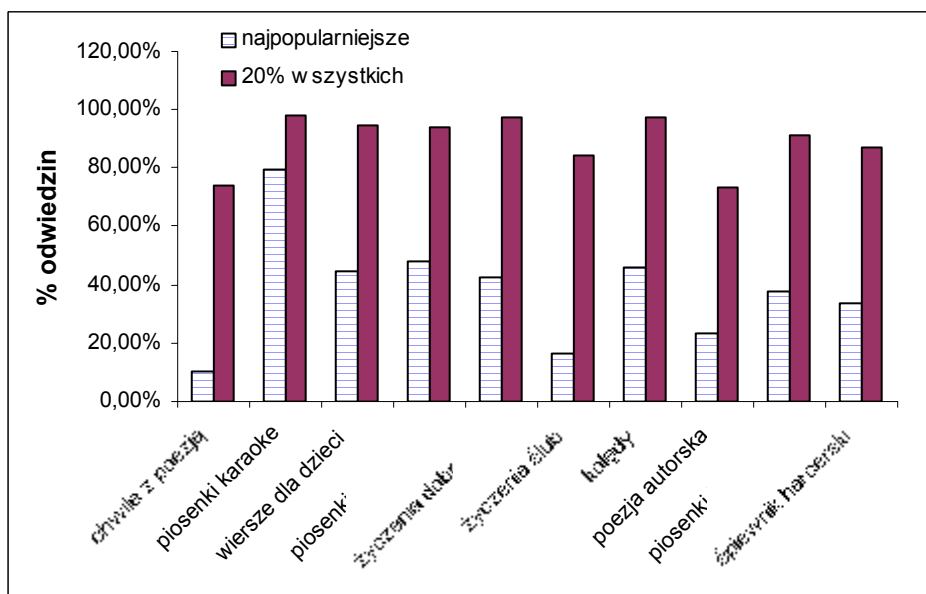
Wyniki przedstawiono w postaci tabeli (rys. 1) oraz wykresów (rys 2, rys 3, rys 4).

Lp.	Charakterystyka witryny				Źródła odwiedzin			Słowa kluczowe						
	tematyka	wiek	ilość stron	ilość odwiedzin	wyszukiwarki	bezpośrednie	odsyłające	wszystkie	najpopularniejsze	% odwiedzin	pozycja w Google	20%	% odwiedzin dla 20%	ilość wyników w Google
1	chwile z poezją	5 lat	100	2 653	42,78%	34,11%	23,11%	235	ewa bicom	10,61%	1	47	73,92%	198 000
2	piosenki karaoke	4 lata	160	164 921	85,93%	10,57%	3,50%	3 117	karaoke	79,21%	2	623	97,73%	76 700 000
3	wiersze dla dzieci	4 lata	100	19 874	63,77%	7,01%	29,22%	726	wiersze dla dzieci	44,55%	1	145	94,82%	711 000
4	piosenki dla dzieci	3 lata	170	40 792	75,99%	12,19%	11,82%	1 870	piosenki dla dzieci	47,82%	2	374	94,01%	557 000
5	życzenia dobranoc	3 lata	25	8 570	90,02%	7,57%	2,40%	206	kołysanki	42,55%	7	41	97,23%	577 000
6	życzenia ślub	3 lata	30	1 379	69,04%	14,72%	16,24%	161	życzenia na ślub	16,70%	4	32	84,14%	378 000
7	kolędy	3 lata	70	167 783	82,00%	5,69%	12,31%	4 198	kolędy	46,20%	2	840	97,17%	6 190 000
8	poezja autorska	2 lata	70	983	36,01%	34,59%	29,40%	114	eligiusz dymowski	23,45%	1	23	73,16%	9 860
9	piosenki biesiadne	2 lata	70	8 917	78,42%	16,75%	4,82%	634	piosenki biesiadne	37,52%	3	127	91,09%	45 000
10	śpiewnik harcerski	1 rok	60	3 611	68,90%	24,90%	6,20%	361	piosenki harcerskie	33,28%	3	72	86,82%	38 400

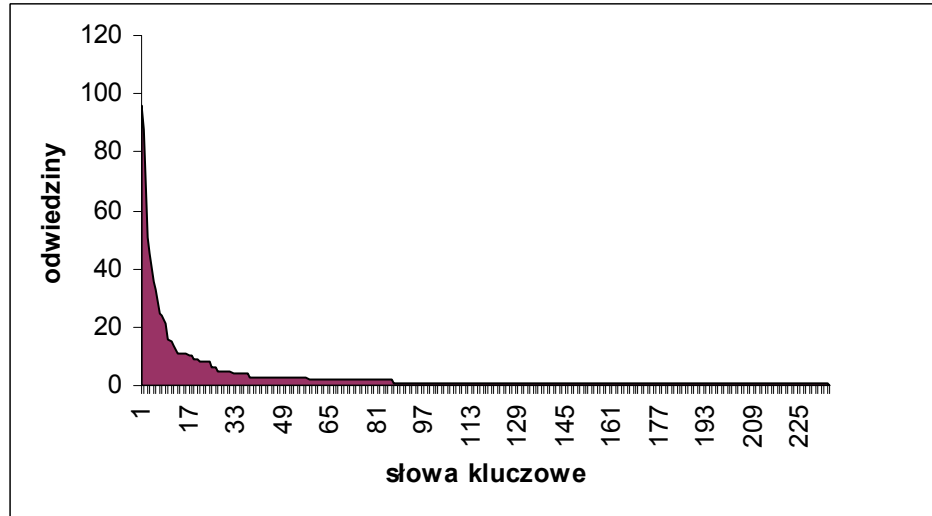
Rysunek 1 Wyniki badań przeprowadzonych dla 10 witryn w domenie ewa.bicom.pl
źródło: opracowanie własne



Rysunek 2 Zależność ilości odwiedzin od ilości wszystkich słów kluczowych, które je wygenerowały
źródło: opracowanie własne



Rysunek 3 Zestawienie ilości procentowych odwiedzin dla najpopularniejszego słowa kluczowego oraz dla 20% wszystkich słów kluczowych poszczególnych witryn
źródło: opracowanie własne



Rysunek 4 Długi ogon wyszukiwania dla witryny nr 1
źródło: opracowanie własne

4. Opis wyników i wnioski końcowe

W wyniku przeprowadzonego eksperymentu stwierdza się, że:

- Głównym źródłem odwiedzin na wszystkich witrynach są wyszukiwarki internetowe. Jest to efekt skutecznego pozycjonowania tych witryn w wyszukiwarkach. O skuteczności pozycjonowania świadczą wysokie (1-7) pozycje w wyszukiwarce **Google** najpopularniejszego słowa kluczowego witryny oraz duża ilość wszystkich słów kluczowych prowadzących na tą witrynę.

Najpopularniejsze słowo każdej z witryn jest również popularnym słowem dla **Google**, gdyż wyszukiwarka zwraca dużą ilość wyników dla każdego z tych słów.

- Ilość odwiedzin na poszczególnych witrynach jest zróżnicowana. Jest ona wprost proporcjonalna do ilości słów kluczowych generujących te odwiedziny (rys.2).
- Na poszczególnych witrynach zauważa się zróżnicowany udział procentowy odwiedzin generowanych przez najpopularniejsze słowo kluczowe.

Procentowy udział odwiedzin wygenerowanych przez 20% słów kluczowych (posortowanych malejąco według popularności) dla wszystkich witryn jest podobny (rys. 3).

Udział poszczególnych słów kluczowych (z pozostałych 80%) w generowaniu odwiedzin jest niewielki, stąd widoczny na rys 4 długi ogon.

Jak wynika z przytoczonych powyżej danych na ilość odwiedzin, które są wyznacznikiem popularności witryny, największy wpływ miało 20% najpopularniejszych słów kluczowych związanych tą witryną. Udział pozostałych 80% słów kluczowych w zwiększeniu oglądalności witryny był nieznaczny. Okazuje się więc, nie sprawdziła się tu koncepcja, którą przedstawił Chris Anderson. Zjawisko nazwane **długim ogonem wyszukiwania** przy przyjętych w badaniu założeniach okazało się **bliższe zasadzie Pareto**.

W procesie tworzenia witryny WWW, jej optymalizacji oraz promocji szczególną uwagę należy zwrócić zatem na grupę najpopularniejszych słów kluczowych, przy pomocy których potencjalny klient trafia na tą witrynę. Nie można jednak zapominać o istnieniu długiego ogona wyszukiwania, który dla dobrze zoptymalizowanych witryn będzie rósł sam wraz ze wzrostem ich popularności. Powinno się jednak na bieżąco analizować statystyki odwiedzin, gdyż słowo z długiego ogona wyszukiwania może stać się jednym z popularniejszych i wtedy na nie należałoby skierować uwagę.

Literatura

- [1] Anderson Ch., *The Long Tail: Why the Future of Business is Selling Less of More*, Wyd. Hyperion, Nowy Jork 2006.
Dostępne na <http://www.wired.com/wired/archive/12.10/tail.html>
- [2] Brynjolfsson, Erik, Hu, Yu Jeffrey and Simester, Duncan, *Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales* (November 2007).
Dostępne na: SSRN: <http://ssrn.com/abstract=953587>
- [3] http://www.stat.gov.pl/gus/spoleczenstwo_informacyjne_PLK_HTML.htm
- [3] http://pliki.gemius.pl/Raporty/2008/11_12_2007_Konsumpcja_mediow.pdf
- [5] Frontczak T., *Marketing internetowy w wyszukiwarkach*, Wyd. HELION, Gliwice 2006.
- [6] Thurow S., *Pozycjonowanie w wyszukiwarkach internetowych*, Wyd. HELION, Gliwice 2004.
- [7] Białek E., *Ocena skuteczności pozycji linku w rankingach wyszukiwarek*, Wyd. Uczelniane PK, Koszalin 2007.
- [8] Sztabiński P., Sztabiński F., Sawiński Z., *Nowe metody, nowe podejścia badawcze w naukach społecznych*, Wyd. IFiS PAN, 2004.
- [9] Biuletyn ranking.pl z dnia 10.04.2008.