



Zarządzanie treścią (CMS)

Wykład 14: Wyszukiwarki, a zarządzanie treścią (3)

dr inż. Mariusz Trzaska,
mtrzaska@mtrzaska.com, [http:// www.mtrzaska.com](http://www.mtrzaska.com)

Ciąg dalszy poprzedniego wykładu ...

Domena dla portalu

- Zacznij od kilku słów kluczowych.
- Następnie sprawdź ich kombinacje.
- Staraj się aby była:
 - łatwa do zapamiętania oraz wpisania,
 - krótka.
- Zadbaj o unikalność domeny - nazwy podobne do istniejących domen nie są polecane.
- Poszukiwania nazwy zacznij od domeny *.com.

Domena dla portalu (2)

● międzynarodowa, np.

- .pl,
- .us,
- .ai (Anguilla),
- .dj (Djibouti),
- .cd (Congo),
- .cv (Cape Verde),

- .fm (Micronesia),
- .io (Indian Ocean),

● nowe domeny, np.:

- .academy,
- .adult,
- .app,

- .art,
- .attorney,
- .auto, .car,
- .bar,
- .bike,
- .biz,
- .bot,
- .cheap,
- ...

https://en.wikipedia.org/wiki/List_of_Internet_top-level_domains

Domena dla portalu (3)

- Unikaj liczb oraz „-”.
- Ostrożnie z aktualnymi trendami. Domena jest wybierana na lata!
- Przydatne narzędzia:
 - <https://www.namemesh.com/>
 - <https://instantdomainsearch.com/>
 - <https://www.panabee.com/>
 - <https://namelix.com>

Domena dla portalu (4)

● TOP 20 najdrożej sprzedanych domen

1Voice.com	\$ 30 000 000,00	11Slots.com	\$ 5 500 000,00
2Sex.com	\$ 13 000 000,00	12Toys.com	\$ 5 100 000,00
3Tesla.com	\$ 11 000 000,00	13Clothes.com	\$ 4 900 000,00
4Fund.com	\$ 9 999 950,00	14IG.com	\$ 4 700 000,00
5Porn.com	\$ 9 500 000,00	15HG.com	\$ 3 770 000,00
6Porno.com	\$ 8 888 888,00	16MI.com	\$ 3 600 000,00
7HealthInsurance.com	\$ 8 133 000,00	17Ice.com	\$ 3 500 000,00
8We.com	\$ 8 000 000,00	18Whisky.com	\$ 3 100 000,00
9Diamond.com	\$ 7 500 000,00	19Sex.xxx	\$ 3 000 000,00
10Z.com	\$ 6 784 000,00	19Vodka.com	\$ 3 000 000,00
		19California.com	\$ 3 000 000,00

Źródło: <https://www.dnjournal.com/archive/domainsales/dnjournal-all-time-top-20-cash-domain-sales.htm>

Domena dla portalu (5)

● Najdrożej sprzedane domeny .PL do 2017

1.	co.pl	1 008 000,00 PLN	29.09.2011
2.	opony.pl	960 000,00 PLN	26.02.2010
3.	moneta.com	450 000,00 PLN	30.09.2013
4.	urlopy.pl	400 000,00 PLN	10.07.2008
5.	grydladzieci.pl	156 000,00 PLN	28.05.2013
6.	kino.pl	150 000,00 PLN	16.02.2009
7.	lekarze.pl	135 000,00 PLN	25.09.2009
8.	magazyny.pl	132 000,00 PLN	16.08.2010
9.	lista.pl	98 000,00 PLN	05.05.2011
10.	domain.pl	97 313,00 PLN	20.06.2008
11.	work.pl	95 000,00 PLN	01.06.2007
12.	makijaz.pl	92 500,00 PLN	06.2016
13.	l.pl	85 000,00 PLN	10.07.2012
14.	material.pl	80 000,00 PLN	22.07.2010
15.	sms.pl	73 800,00 PLN	12.02.2013

16.	deser.pl	73 792,00 PLN	30.06.2008
17.	ss.pl	72 105,00 PLN	25.09.2008
18.	baza.pl	72 000,00 PLN	30.04.2009
19.	adwokaci.pl	70 000,00 PLN	04.10.2008
20.	eventy.pl	70 000,00 PLN	01.07.2009
21.	CasinoOnline.pl	66 000,00 PLN	03.2016
22.	historia.pl	61 000,00 PLN	24.02.2009
23.	telefony.pl	59 600,00 PLN	20.04.2009
24.	choinki.pl	53 100,00 PLN	02.01.2013
25.	garnki.pl	52 800,00 PLN	16.05.2013
26.	lekarze.pl	52 188,00 PLN	17.11.2008
27.	lampa.pl	50 200,00 PLN	04.05.2012
28.	meble.pl	50 000,00 PLN	01.10.2003
29.	oczyszczalnie.pl	50 000,00 PLN	28.09.2009
30.	wedkarski.pl	50 000,00 PLN	25.10.2010

Źródło: <https://sferadomen.pl/top-100-sprzedanych/>

Domena dla portalu (6)

- Liczba domen PL w latach 2001 - 2019

	Liczba domen .pl	Zmiana
2001	134 721	51,44%
2002	159 509	18,40%
2003	195 842	22,78%
2004	276 687	41,28%
2005	443 249	60,20%
2006	557 091	25,68%
2007	762 656	36,90%
2008	1 323 454	73,53%
2009	1 629 929	23,16%
2010	1 994 120	22,34%
2011	2 278 722	14,27%
2012	2 402 107	5,41%
2013	2 461 499	2,47%
2014	2 524 538	2,56%
2015	2 681 752	6,23%
2016	2 703 165	0,80%
2017	2 576 063	-4,70%
2018	2 596 748	0,80%
2019	2 513 491	-3,21%

Źródło: <https://danieldryzek.pl/2020/05/29/co-90-sekund-rejestrowana-jest-nowa-domena-pl/>

Umiejscowienie słów kluczowych

- Wydaje się, że aktualnie największe znaczenie ma tytuł strony: tag `title`.
 - Słowa kluczowe na początku, nazwa firmy na końcu.
 - Długość (razem ze spacjami): maksymalnie 65 znaków.
 - Warto wykorzystywać dłuższe frazy, a nie tylko pojedyncze słowa.
 - Czasami można stosować znaki rozdzielające: |, >, -, :

Umiejscowienie słów kluczowych (2)

● Tytuł strony – *c.d.*

- Jeżeli to możliwe to tytuł powinien być zachętą do działania, np.: „Kup kamerę cyfrową”.
- Zadbaj o spójność tytułów w portalu.

● Meta tag: `description`

- Wykorzystywany przez wyszukiwarke na stronach SERP;
- Upewnij się, że odpowiada rzeczywistości – nie koloryzuj!
- Najlepiej zmieścić się w 160 znakach (Google) lub 200 (Bing).

Umieszczenie słów kluczowych (3)

- Meta tag: `description` - *c.d.*
 - Potraktuj go jak reklamę: krótką i treściwą.
 - Testuj różne wersje dopóki nie będziesz zadowolony z wyniku.
 - Wykorzystaj pasujące słowa kluczowe.
- Tagi nagłówków: H1, H2, H3
 - H1 jako pojedynczy nagłówek całej strony. Raczej nie używamy go do opisanego logo/nazwy portalu.
 - H2, H3 to tytuły sekcji na stronie. Mogą występować w kilku miejscach.

Umieszczenie słów kluczowych (4)

● Treść dokumentu

- Kiedyś umieszczanie słów kluczowych w tekście było podstawą SEO.
- Aktualnie przywiązuje się do tego mniejszą wagę.
- Jest to spowodowane uodpornieniem się wyszukiwarek na liczne próby oszustw.

● Nazwy plików graficznych oraz atrybut `alt`.

● Tagi `` oraz ``. Niektórzy uważają, że warto (a raczej nie zaszkodzi).

Kanibalizacja słów kluczowych

- Polega na stosowaniu tych samych słów kluczowych na wielu stronach i/lub tytułach stron.
- W efekcie wyszukiwarka musi „wybrać” którą (jedną) stronę uwzględnić w SERP.
- Innymi słowy, taki portal konkuruje sam ze sobą.
- Aby tego uniknąć trzeba zadbać o właściwą i konsekwentnie realizowaną strukturę informacyjną portalu.

Problem zduplikowanych treści

- Spamerzy pobierają treści z różnych źródeł i (po niewielkich przeróbkach) publikują je na swoich stronach licząc na:
 - Przyciągnięcie odwiedzających,
 - Pozyskanie linków prowadzących do ich witryn.
- W efekcie wyszukiwarka musi wybrać, którą witrynę zwrócić w SERP.

Problem zduplikowanych treści (2)

- Czasami ten problem występuje również w ramach tego samego portalu:
 - Wyszukiwarka indeksując wielokrotnie te same treści, może stracić za dużo czasu i nie przetworzy czegoś wartościowego;
 - Linki do zduplikowanej zawartości, mogą powodować „rozdrobienie” wartości naszych treści.
 - W celu przeciwdziałania można skorzystać pliku *robots.txt* lub meta tagów: `name=„robots”`
`content=„noindex, follow”`.

Problem zduplikowanych treści (3)

- Aby wykryć zduplikowane treści można skorzystać z narzędzia: <http://copyscape.com/>
- W takiej sytuacji można zgłosić naruszenie DMCA (*Digital Millennium Copyright Act*) i powiadomić o tym obsługę wyszukiwarek.
- Ewentualnie wytoczyć proces cywilny.

Cloaking

- Polega na dostarczaniu dwóch (lub więcej) wersji witryny w zależności od odbiorcy:
 - Roboty wyszukiwarek,
 - Użytkownicy.
- Rozpoznawanie może odbywać się np. na podstawie adresów IP i/lub *UserAgent*.
- Technika ta może być wykorzystywana w celu oszukiwania wyszukiwarek. W takiej sytuacji należy liczyć się z surową karą.

Cloaking (2)

- Można z niej korzystać także w celu poprawy współpracy z wyszukiwarką (niemniej trzeba uważać):
 - Testy A/B dla użytkowników oraz jedna (prosta) wersja dla robota,
 - Treści wymagające rejestracji,
 - System nawigacji niewspierający robotów (np. JavaScript),
 - Zduplikowane treści,
 - Treści uzależnione np. od lokalizacji geograficznej.

Wybranie właściwego CMS'a

- Możliwość modyfikacji tytułu strony.
- Przyjazne URL'e ze słowami kluczowymi.
- Modyfikacja meta tagów.
- Stosowanie własnych atrybutów/tagów HTML.
- Zmiana opisów linków (*anchor*).
- Elastyczny system kategorii z możliwością opisów.
- Kontrola stronicowania.
- Przekierowanie 301.

Wybranie właściwego CMS'a (2)

- Wsparcie dla RSS.
- Zarządzanie obrazkami (w tym atrybutem `alt`).
- Cache'owanie treści.
- URL bez parametrów (np. sesji).
- Tagowanie (taxonomie).
- Nawigacja okruszkowa (*breadcrumb*).
- Generator mapy witryny (*sitemap.xml*).
- Walidacja HTML5.

Portale wielonarodowe

- Domena dla określonego kraju (pl, com.pl).
- Hosting na terenie kraju.
- Fizyczny (geograficzny) adres na wielu stronach witryny.
- Ustawienie kraju w [Google Webmaster Central](#).
- Linki docelowe z różnych krajowych witryn.
- Wykorzystanie języka narodowego.

Portale wielonarodowe (2)

- Jeżeli główna strona ma silną markę to można utworzyć wersje narodowe korzystając z:
 - Pod-domen, np. *pl.mojafirma.com/*. Każda poddomena (język) ma własną „wartość”,
 - Wirtualnych/logicznych katalogów, np. *mojafirma.com/pl/*. „Wartość” podkatalogów (języków) zależy od ogólnej „wartości” domeny.

Linki

- Linki prowadzące do danej witryny są jednym z najważniejszych czynników określających jej jakość.
 - Liczba,
 - Jakość witryny źródłowej,
 - Tekst opisujący.

Algorytm PageRank

- W uproszczeniu:
 - Każda witryna ma swoją „jakość”,
 - Publikując linki do witryn, niejako głosujemy na nie.
 - „Siła” tego głosu zależy właśnie od jakości naszej witryny.
 - Jest rozkładana na wszystkie linki („głosy”). Czyli im więcej linków tym mniejsza jednostkowa siła głosu.
- Sprawy się komplikują gdy uwzględnimy linki powrotne i dowolne kombinacje pomiędzy witrynami.
- Algorytm PR podlega nieustającym zmianom.

Wartość linku

● Tekst opisowy (*anchor*)

- Wykorzystywany do określenia czego dotyczy docelowa witryna.
- Może okazać się bardzo wartościowy gdy wskazywany portal korzysta z technologii, która nie jest przetwarzana przez wyszukiwarki (np. JavaScript).

Wartość linku (2)

- Tekst opisowy (*anchor*) – c.d.
 - *Google Bombing* – tworzenie wielu linków do tej samej witryny z tym samym opisem. W efekcie witryna dla zapytania zawierającego słowa kluczowe z opisu znajduje się bardzo wysoko w SERP. Google w 2009 roku wprowadził odpowiednie zmiany do algorytmu.
- Odpowiedniość (*relevance*). Linki tworzone w ramach stron o tej samej tematyce są dużo bardziej wartościowe.

Wartość linku (3)

● Autorytet (authority)

- *Hubs* – witryny posiadające linki do ważnych witryn;
- *Authorities* – witryny, które są wskazywane przez strony o określonej tematyce;
- Początkowo wprowadzony jako algorytm Teoma w Ask.com. Aktualnie zaimplementowany w większości wyszukiwarek.

Wartość linku (4)

● Zaufanie (*Trust*)

- Bierze pod uwagę, czy linkujące strony/domeny są wiarygodne (np. trudne do zinfiltrowania przez spamerów).
- Uwzględniane jest zaufanie do witryn wskazujących.
- Startujemy z pewną pulą, ręcznie wybranych witryn i sprawdzamy jak daleko (*clicks*) jest oceniana witryna (*TrustRank*).

Wartość linku (5)

● Źródła.

- Linki w ramach tej samej witryny mają oczywiście mniejszą wartość.
- Analogicznie jest gdy występuje dużo wzajemnych linków pomiędzy kilkoma witrynami.
- Uwzględniane są też informacje dotyczące własności/zależności z *whois*.

● Zróżnicowanie źródeł – warto aby mieć linki np. z: blogów, forów, portali, serwisów społecznościowych, itd.

Wartość linku (6)

- Aspekt czasu
 - Kiedy link się pojawił?
 - Kiedy zniknął?
 - Jak długo był widoczny?
 - Jak szybko pojawiały się nowe linki?
- Sąsiedztwo innych linków na tej samej stronie.
- Lokalizacja na stronie (*prominence*).
- Tekst otaczający link (w tym nagłówek).
- Tematyka treści strony oraz witryny.

Wartość linku (7)

● Rodzaj domeny TLD

- .edu, .gov, .mil, .com;
- Raczej [nie jest uwzględniany bezpośrednio.](#)

● Samodzielne szacowanie:

- *PageRank* („wartość” bezwzględna) strony publikującej link (lub strony domowej witryny),
- Autorytet witryny („wartość” w kontekście określonej tematyki),
- *PageRank* oraz Autorytet strony docelowej,

Wartość linku (8)

- Samodzielne szacowanie – c.d.:
 - Liczba linków „wychodzących” na stronie źródłowej,
 - Odpowiedniość (zbliżona tematyka) strony i witryny źródłowej.

Pozyskiwanie linków

- Publikujący znalazł interesujące treści na naszej stronie
 - Chce aby jego odwiedzający wiedzieli o tym,
 - Jak rozpropagować informacje o naszych treściach?
 - Udostępnianie fragmentów/treści portalu,
 - Media społeczne, np. Facebook, Instagram, Twitter, LinkedIn, Pinterest, Reddit,
 - Blogi,
 - Katalogi (*Directories*).
 - Najbardziej wartościowe z punktu widzenia wyszukiwarek.

Pozyskiwanie linków (2)

- Emocjonalna reakcja publikującego spowodowana np. zabawnym rysunkiem czy kontrowersyjną treścią.
 - *Link baiting*,
 - Trzeba zdecydować czy jest spójne np. z wizerunkiem organizacji.
- Zależności biznesowe – np.
 - Sieć sprzedaży (warto umieścić odpowiedni zapis w umowie o współpracy),
 - Podwykonawcy.

Pozyskiwanie linków (3)

- Bezpośrednia zachęta/prośba do utworzenia linku do naszego portalu
 - Skierowana do witryny (spersonalizowana), która może być zainteresowana naszą tematyką,
 - Wypełnienie formularza kontaktowego lub wysłanie maila,
 - Podanie konkretnych powodów/korzyści z utworzenia takiego linku.
- Samodzielne utworzenie linku w mediach społecznościowych, np. Facebook, Instagram, Twitter, LinkedIn, Pinterest, Reddit.

Pozyskiwanie linków (4)

- Wymiana pomiędzy witrynami
 - *Reciprocal linking*.
 - Ograniczona wartość dla wyszukiwarek.
- Linki płatne
 - Z punktu widzenia wyszukiwarek są bezwartościowe.
 - W niektórych sytuacjach takie zachowania mogą być karane niższymi wynikami w SERP.
- Komunikacja tradycyjna, np. oferta zorganizowania (bezpłatnego) seminarium (które będzie reklamowane w witrynie – link).

Pozyskiwanie linków – wątpliwe sposoby

- Korzystanie z nich może skutkować **ukaraniem** przez wyszukiwarki (aż do usunięcia z indeksu/SERP).
- Kupowanie linków
 - Łatwe,
 - Google sugeruje aby reklamy korzystały z `NoFollow` (brak wartości dla wyszukiwarek).

Pozyskiwanie linków – wątpliwe sposoby (2)

● Kupowanie linków – c.d.

○ Sposoby

- Bezpośrednie zakupy,
- Brokerzy linków,
- Dotacje dla organizacji (w zamian za linki do sponsorów).

● Farmy linków

- Postrzegane jako spam,
- (Ewentualnie) skuteczne krótkoterminowo,

Pozyskiwanie linków – wątpliwe sposoby (3)

- Automaty (bot'y) tworzące linki;
 - Publikują wpisy w:
 - Komentarzach,
 - Forach dyskusyjnych
 - Można im przeciwdziałać za pomocą captch'a, rejestracji, itp.
 - Efekty linku można niwelować korzystając z `NoFollow`.
- Widżety w JavaScript, które zawierają link w tagu `NoScript` (tylko ta część jest przetwarzana przez wyszukiwarki).

Strategia pozyskiwania linków

- Potencjalni kandydaci do linkowania do naszej witryny:
 - Niekonkurencyjne witryny z tego samego rynku (np. informacje branżowe),
 - Media,
 - Blogerzy,
 - Uczelnie,
 - Strony rządowe,
 - Hobbyści,
 - Strony posiadające linki do naszej konkurencji (sukces jak osiągniemy skuteczność 10%).

Strategia pozyskiwania linków (2)

- Skąd biorą linki nasi konkurenci?
 - Narzędzia,
 - Zwróćmy uwagę na najważniejsze z nich (wysoki *PageRank*, *mozRank/mozTrust*)
- Ocena własnych zasobów
 - Które warto promować,
 - Co nowego możemy stworzyć?
- Niekończący się proces.

Strategia pozyskiwania linków (3)

- Porównanie skuteczności różnych podejść, np.
 - Utworzenie nowych treści kosztuje X PLN,
 - Może taniej wyjdzie np. kupić reklamy?

Reklamy Google'a – najdroższe słowa kluczowe (opłata za kliknięcie):

1. Business Services: \$58.64
2. Bail Bonds: \$58.48
3. Casino: \$55.48
4. Lawyer: \$54.86
5. Asset Management: \$49.86
6. Insurance: \$48.41
7. Cash Services & Payday Loans: \$48.18
8. Cleanup & Restoration Services: \$47.61
9. Degree: \$47.36
10. Medical Coding Services: \$46.84

Źródło: <https://www.socialmediatoday.com/news/these-are-the-10-most-expensive-google-keywords/512893/>

Śledzenie efektów SEO

- SEO to długotrwały proces nie zawsze oparty na jasnych przesłankach.
- Ustalenie warunków początkowych (abyśmy mieli z czym porównywać)
 - Ustalenie słów kluczowych przyciągających odwiedzających,
 - Podział ruchu wynikającego z organicznych wyników na podstawie słów kluczowych,
 - Najpopularniejsze sekcje/treści witryny,

Śledzenie efektów SEO (2)

- Ustalenie warunków początkowych – c.d.
 - Niezbyt popularne treści,
 - Aktywność robotów wyszukiwujących,
 - Liczba zaindeksowanych stron.
- Miary sukcesu
 - Poprawa wyników sprzedaży,
 - Zwiększona oglądalność reklam,

Śledzenie efektów SEO (2)

- Miary sukcesu – c.d.
 - Zwiększenie liczby
 - subskrypcji RSS, *newsletter'a*,
 - rejestracji,
 - pobrań,
 - prób kontaktów,
 - linków do naszej witryny.
 - Poprawa ruchu w witrynie,
 - Pozyskanie dotacji.

Śledzenie efektów SEO (3)

● Narzędzia

- Analityczne udostępniane przez firmę hostingową (analiza logów),
- [Google Analytics](#),
- <https://www.bing.com/toolbox/webmaster>
- <https://www.adobe.com/experience-cloud.html> (płatne),
- <http://www.unica.com> (płatne),
- <http://www.webtrends.com/> (płatne),
- ...

Śledzenie efektów SEO (4)

- Ważniejsze informacje:

- Ruch w zależności od wyszukiwarki,
- Ruch w zależności od słów kluczowych,
- Strony źródłowe (*referring sites*),
- Najpopularniejsze strony/treści,

- Najczęstsze błędy - podejmowanie decyzji na podstawie:

- Zbyt małej próbki danych,
- Danych z niereprezentatywnego okresu czasowego (np. grudzień w sklepie internetowym).

Śledzenie efektów SEO (5)

- Wyszukiwania marki. Jeżeli większość ruchu z wyszukiwarek pochodzi z zapytań dotyczących naszej marki to prawdopodobnie mamy problem
 - Niewykorzystanie ogólnych słów kluczowych,
 - Brak zapytań dotyczących tzw. długiego ogona (ważne!).

Wartościowe zasoby dotyczące SEO

- [Search Engine Land](#),
- [Search Engine Watch](#),
- [SEOMoz](#),
- [Web Site Magazine](#),
- [Google Webmaster Central Blog](#),
- [Microsoft Bing Blogs](#),
- [Google Webmaster Help](#),
- [Webmaster World](#),
- [Search Engine Roundtable](#).

Słowniczek

- *Cloaking* (w SEO) – udostępnianie różnych wersji witryny w zależności od odbiorcy.
- PPC (*Pay Per Click*) – wyniki sponsorowane, reklamy.
- ROI (*Return On Investment*) – zwrot z inwestycji
- SERP (*Search Engine Result Page*) – strona wyników wyszukiwania.
- TLD (*Top Level Domain*) – domena najwyższego rzędu, np. com.

Podsumowanie

- SEO jest bardzo istotnym aspektem współczesnego zarządzania treścią.
- Bardzo często, od jego prawidłowej realizacji zależy sukces naszego przedsięwzięcia.
- Jest to niekończący się proces, który wymaga rozległej, interdyscyplinarnej wiedzy.
- Może wykorzystywać techniki spamerskie chociaż jest to zwykle nieopłacalne długoterminowo.