

# Wyszukiwanie i Przetwarzanie Informacji WWW

## Ekonomiczne aspekty WWW: (1) reklamy wyszukiwarkowe

Marcin Sydow

PJWSTK

# Plan tego wykładu

- Wprowadzenie
- Podstawy Reklam WWW
- Podsumowanie Wykładu

# Pierwotne Założenia WWW

Pierwotnie WWW: niewielka sieć akademicka w CERN (1991)

W założeniu bardzo szlachetne pobudki:

- “demokratyczne” medium dla wszystkich
- otwartość, brak jakiegokolwiek wrogości
- brak jakichkolwiek elementów komercyjnych

# WWW dzisiaj?

Niewiele dziś zostało z tych szlachetnych założeń:

Co dzisiaj obserwujemy:

- uzależnienie wszystkich od największych wyszukiwarek i pozycji w nich
- wszechobecne współzawodnictwo i działania o znamionach wrogości (spam, etc.) - (ang. adversarial behaviour)
- wielkie zaangażowanie ekonomiczne
- ogromną ilość reklam

WWW staje się **kolejnym komercyjnym medium** rządzone **swoistymi prawami ekonomicznymi**

Właściwie utrzymane zostało tylko jedno założenie: dostępność dla wszystkich (prawie) za darmo (- miejmy nadzieję, że to się nie zmieni)

# WWW - łańcuch zależności

- 1 Web zależy od:
  - **“wydawców”**, którzy publikują swoje dokumenty <sup>1</sup>
  - **użytkowników**, którzy je oglądają
- 2 Użytkownicy zależą od **wyszukiwarek** - bez których dzisiejszy Web nie mógłby istnieć
- 3 Wyszukiwarki - zależą od **reklam** (główne źródło ich dochodu)

Proza życia: Centralną siłę napędową WWW tworzą dzisiaj **Reklamy**

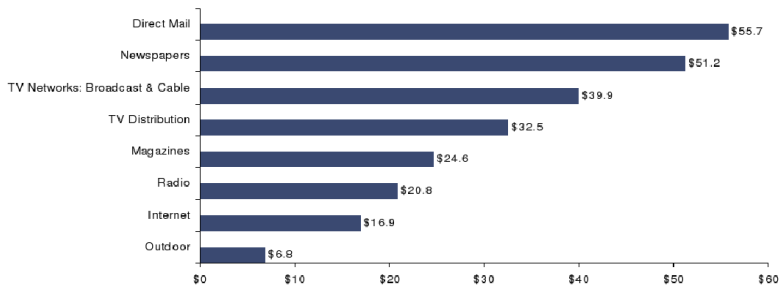
---

<sup>1</sup>nie chodzi tu tylko o ludzi, mogą to być korporacje, roboty, etc.

# Ogólny Rynek Reklamy (USA, 2006)

- Internet advertising revenues accounted for approximately 5.9 percent of total U.S. ad spending\* in 2006, up from approximately 4.7 percent in 2005.

## U.S. Advertising Market-Media Comparisons—2006 (\$ Billions)

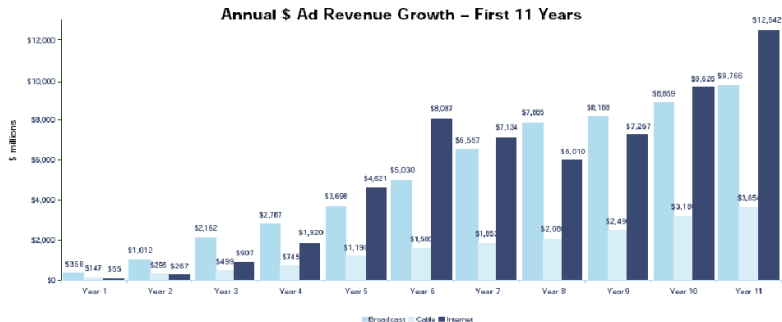


\*The total U.S. advertising market is estimated at approximately \$285 billion, and includes other segments not charted here.

Sources: IAB Internet Ad Revenue Report; PricewaterhouseCoopers Global Entertainment and Media Outlook

# Wzrost Rynku Reklam Internetowych

Pierwsze 11 lat funkcjonowania reklam Internetowych (1995 - 2005) (w porównaniu do pierwszych 11 lat funkcjonowania reklam w telewizji (od 1949) i w telewizji kablowej (od 1980) (wartości w “dzisiejszych” dolarach)



Sources: IAB Internet Ad Revenue Report; PricewaterhouseCoopers LLP; Universal McCann

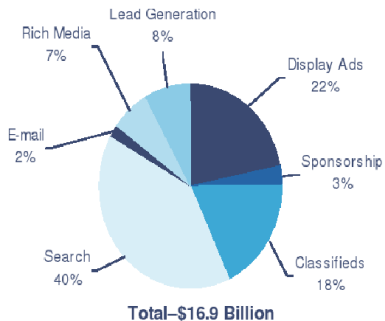
# Rodzaje Reklam Internetowych

- reklamy wyświetlane statycznie (stałe pokazywanie baneru, logo, etc.)  
- ang. Display Advertising
- reklamy pocztowe (banery, kampanie wysyłkowe, etc.)
- **reklamy wyszukiwarkowe**
- przekierowania do reklamodawcy, etc. (ang. Lead Generation/Referrals) - reklamodawca płaci wyspecjalizowanej firmie za znajdowanie potencjalnych klientów (drogą Internetową) i ich przekierowywanie oraz udostępnianie specjalistycznej wiedzy dotyczącej danych demograficznych, geograficznych, etc.
- reklamy “klasyfikowane” (ang. classifieds) - reklamodawca płaci za umieszczanie reklam w specjalistycznych listach (np. “yellow pages”, sprzedaż nieruchomości, etc.)

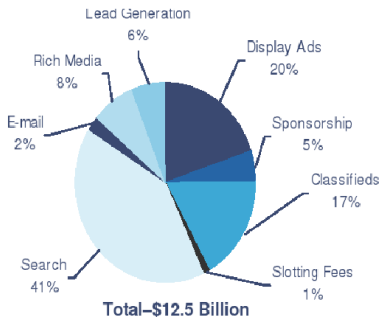


# Rozkład Reklam Internetowych w ostatnich latach

% of 2006 Full-Year Revenues



% of 2005 Full-Year Revenues



Jak widać, **reklamy wyszukiwarkowe** stanowią **dominujący składnik** rynku reklam internetowych (obecnie 40%) (źródło: IAB Ad Revenue Report, PricewaterhouseCoopers LLP, USA, 2005-2006)

# Rynek Reklam Internetowych Obecnie (USA, 2006)

Type of Advertising	2006 (Ttl = \$16,879M)		2005 (Ttl = \$12,542M)	
	\$	% share of market	\$	% share of market
Display Advertising	3,685	22%	2,508	20%
Sponsorship	496	3%	627	5%
Slotting Fees	0	0%	125	1%
Rich Media (including Broadband Video)	1,192	7%	1,004	8%
<b>All Display</b>	<b>5,373</b>	<b>32%</b>	<b>4,264</b>	<b>34%</b>
Keyword Search	6,799	40%	5,142	41%
Classifieds	3,059	18%	2,132	17%
E-mail	338	2%	251	2%
Lead Generation	1,310	8%	753	6%
<b>TOTALS:</b>	<b>16,879</b>	<b>100%</b>	<b>12,542</b>	<b>100%</b>

Zyski z reklam wyszukiwarkowych w USA w 2006 wyniosły aż **\$6.8B**

# Reklamy a wyszukiwanie

Aż **97%** zysku z wyszukiwarek stanowią zyski z reklam (A.Broder, Yahoo, 2006)

Z tych powodów, tematy wyszukiwarek i reklam są ze sobą **mocno powiązane**.

Szczególnie interesujące jest, że w przypadku reklam wyszukiwarkowych można bardzo dokładnie mierzyć rynek (w tym: skuteczność, zyski, etc.)

# Rodzaje Reklam Wyszukiwarkowych

- linki sponsorowane (np. Google AdWords, Netsprint Link Sponsorowany i Boksy Reklamowe)
- reklamy kontekstowe (np. Google AdSense, Netsprint AdKontekst)
- płatne pozycjonowanie w wynikach wyszukiwania (Goto, Overture - obecnie nie stosowane w największych wyszukiwarkach)
- płatne włączenie do indeksu (ang. paid inclusion)
- pozycjonowanie witryn (ang. site optimization/search engine optimization(SEO)) - specjalne dostosowywanie witryny, **zgodnie z regułami** wyszukiwarek, tak aby wyżej pojawiały się w rankingach

# Wyzwania

Zagadnienia związane z reklamami wyszukiwarkowymi (i ogólnie Internetowymi) sprawiają powstawanie **bardzo ciekawych problemów** z zakresu:

- wyszukiwania (Information Retrieval) - dopasowywanie reklam
- ekonomii (modele wyceny słów kluczowych, miejsc reklamowych, etc.)
- modelowania (maksymalizacja dochodów)
- analizy danych (np. wykrywanie nadużyć, etc.)
- inżynierii (duże obciążenia obliczeniowe)

# Perspektywa Ekonomiczna: Grupy Interesów

## 4 Grupy aktorów:

- Wydawcy (tworzący dokumenty i portale)
- Reklamodawcy
- Użytkownicy
- Wyszukiwarki

Pomiędzy tymi grupami tworzą się złożone i subtelne zależności. Należy tak **zaprojektować mechanizm<sup>2</sup>** żeby **służył jak najlepiej wszystkim grupom**

---

<sup>2</sup>ang. Mechanism Design

## Cele grup

Wydawcy - zyski z wyświetlanych reklam, atrakcyjność strony

Użytkownicy - minimalizacja “niezadowolenia” z oglądania niechcianych reklam

Reklamodawcy - maksymalizacja skuteczności kampanii (dotarcie do odpowiednich grup celowych), przy ograniczonym budżecie

Wyszukiwarki - maksymalizacja obrotów, zysków z pośrednictwa, jakości dopasowania

# Reklama oparta na słowach kluczowych

Reklamodawca płaci za wyświetlenie reklamy w odpowiedzi na **odpowiednie słowa kluczowe**

Podobne do “zwykłego” wyszukiwania z pewnymi różnicami:

Baza reklam: słowa kluczowe, tytuł, opis, URL

Baza reklam jest mniejsza od indeksu, reklamy to “małe” strony, ranking zależy nie tylko od podobieństwa, ale także od **licytowanej kwoty**



# Reklama kontekstowa

W tym przypadku są różnice w stosunku do “klasycznego” wyszukiwania:

Baza reklam: słowa kluczowe, tytuł, opis, URL (jak przedtem)

“Zapytanie” reklamowe: **strona** + **kontekst** (geograficzny, personalny, czasowy, etc.)

“Zapytanie” jest stosunkowo **duże**

# Zagadnienia Ekonomiczno-Matematyczne

Reklamodawcy:

- jakie słowa kluczowe kupić?
- ile płacić za poszczególne słowa?

Wyszukiwarki:

- jak pogodzić sprzeczne interesy reklamodawców? (koszt - słowa - miejsce)
- jak zmaksymalizować zyski z pośrednictwa przy ograniczonym budżecie reklamodawców?

# Reklamy a Zagadnienia Wyszukiwawcze

Wyszukiwarki:

Jak **dobrze dopasować reklamy**

- do słów zapytania?
- do stron?

Są to podobne zagadnienia do klasycznych zagadnień wyszukiwarkowych - ale są różnice

## Obserwacje dotyczące reklamy opartej na słowach kluczowych

- Stosunkowo niewiele drogich słów
- Bardzo dużo tanich słów
- Wiele różnych słów dla jednego pojęcia (synonimy, etc.)
- Reklama jest potrzebna, gdy strona nie pojawia się w 10-20 najlepszych wynikach
- Reprezentacja “intencji zapytania” za pomocą odpowiednich (i jak najtańszych!) słów jest kluczem do udanej kampanii

Wyszukiwarki dostarczają wielu narzędzi do pomocy w konstruowaniu kampanii (dobrze słów kluczowych)

# Metody rozliczania reklam

- CPM (ang. cost per mille) - cena za 1000 odsłon
- CPC (ang. cost per click) - tylko jeśli link był kliknięty
- CPT/CPA (ang. cost per transaction/action) - tylko jeśli doszło do transakcji

## Pricing Models

	FY 2006	FY 2005
<b>CPM or Impression</b>	<b>48%</b> (\$8,102)	<b>46%</b> (\$5,769)
<b>Performance Deals</b>	<b>47%</b> (\$7,933)	<b>41%</b> (\$5,142)
<b>Hybrid</b>	<b>5%</b> (\$844)	<b>13%</b> (\$1,630)

# Problem Wyceny

Zagadnienia wyceny są ogromnie skomplikowane ze względu na sprzeczność interesów i cel jakim jest maksymalizacja zysków.

Problemy wyceny:

- słów kluczowych
- miejsc wyświetlenia
- kosztów CPC i CPM, etc.

są obecnie rozwiązywane za pomocą **mechanizmów aukcyjnych**

# Ranking Reklam

Reklama może ukazać się w jednej z **wielu dostępnych “przegródek”** obok wyników wyszukiwania. Wyższe przegródki są bardziej atrakcyjne - są częściej klikane (analogia do rankingu wyników wyszukiwania).

W jaki sposób zdecydować o kolejności wyświetlanych reklam?

Można liczyć **“efektywny CPM”**:  $ECPM = CPC \cdot CTR$  (CTR - ang. click-through rate)

# Mechanizmy aukcyjne

Obecnie w najbardziej zaawansowanych produktach reklamowych wycenę opiera się na wynikach **teorii gier**.

Wszystkich uczestników traktuje się jako **racjonalnych graczy**, których celem jest **maksymalizacja funkcji użyteczności** i którzy stosują pewne **strategie**

Do ustalenia cen i miejsc wyświetleń reklam stosuje się **wyrafinowane mechanizmy aukcyjne**



## Cechy rozsądnego mechanizmu

W przypadku braku konkurencji - cena powinna być minimalna

Ustalić “sprawiedliwą” cenę w warunkach współzawodnictwa i ograniczonych zasobów (miejsc reklamowych) - poprzez znalezienie równowagi między graczami

Motywować graczy do “uczciwej” licytacji - bez zaniżania kwot, które są w stanie zapłacić

# Rodzaje aukcji

## Rodzaje aukcji:

- angielska - początkowa, niska cena (minimalna do zaakceptowania dla sprzedającego) - uczestnicy stopniowo przebijają cenę aż do jej zatrzymania (problem tzw. "pierścieni" - grup uczestników którzy nie przebijają się nawzajem)
- holenderska (opadająca cena) - szybkość (pierwsza akcja któregośkolwiek uczestnika powoduje dojście do transakcji)
- "Zapiecztowana" - gracze ustalają niejawnie swoje ceny - wygrywa najwyższa

Powyższe typy są **teoretycznie równoważne** (dają taki sam **oczekiwany zysk** dla sprzedawcy - ("The Revenue Equivalence Theorem").

Experymenty pokazują, że w praktyce aukcja holenderska daje nieco niższą cenę od pozostałych.

## Problem “przekleństwa zwycięzcy”

Powyższe aukcje mają pewien mankament zwany “przekleństwem zwycięzcy” (ang. the winners curse) - ich mechanizm powoduje tendencję do przepłacania, żeby pokonać konkurentów.

W efekcie gracze starają się **zaniżyć licytację** - nie podają swych **prawdziwych wartości**.

Pewnym Rozwiązaniem tego problemu jest **Aukcja Vickrey'a**<sup>3</sup>. Rodzaj “zapieczętowanej” aukcji, w której zwycięzca płaci **drugą co do wielkości cenę**. Udowodniono, że w takiej aukcji **nie ma motywacji do zaniżania licytacji**.

Odmiany: aukcja typu VCG oraz “uogólniona aukcja drugiej ceny” (ang. generalized second-price auction) - **stosowana w Google i Yahoo!**

---

<sup>3</sup>William Vickrey 1914-1996 - noblista z ekonomii w 1996

# Ochrona Interesów

Ze względu na znaczące zaangażowanie ekonomiczne podmiotów oraz konkurencyjność, mogą pojawiać się zjawiska negatywne, np:

- “wyklikiwanie” kampanii reklamowych konkurencji
- generowanie sztucznych strumieni zapytań lub kliknięć mających na celu manipulowanie cenami pewnych słów kluczowych, CPC, et.

Wyszukiwarki muszą stosować **zaawansowane metody automatycznego wykrywania nadużyć** (ang. fraud detection) aby chronić interesy swoich klientów. Wiąże się to ze stosowaniem zaawansowanych technik Web Mining i Sztucznej Inteligencji.

# Na zaliczenie tego wykładu:

- 1 Mechanizmy ekonomiczne rządzące dzisiejszym WWW
- 2 Rodzaje reklam wyszukiwarkowych
- 3 Główne wyzwania dla wyszukiwarek (w kontekście reklam)
- 4 Podstawowe pojęcia z zakresu reklam WWW
- 5 Sposoby wyceny reklam

Dziękuję za uwagę