

# ΜΥΕ003: Ανάκτηση Πληροφορίας

*Διδάσκουσα: Ευαγγελία Πιτουρά*

Κεφάλαια 19, 20: Web. Μηχανές αναζήτησης.

*Ακαδημαϊκό Έτος 2021-2022*

# Τι θα δούμε σήμερα

- Web και μηχανές αναζήτησης
  - λίγη ιστορία
  - ο γράφος του web
  - σημασία της άγκυρας (anchor text)
- Χρήστες
- Εξατομίκευση
- Διαφημίσεις

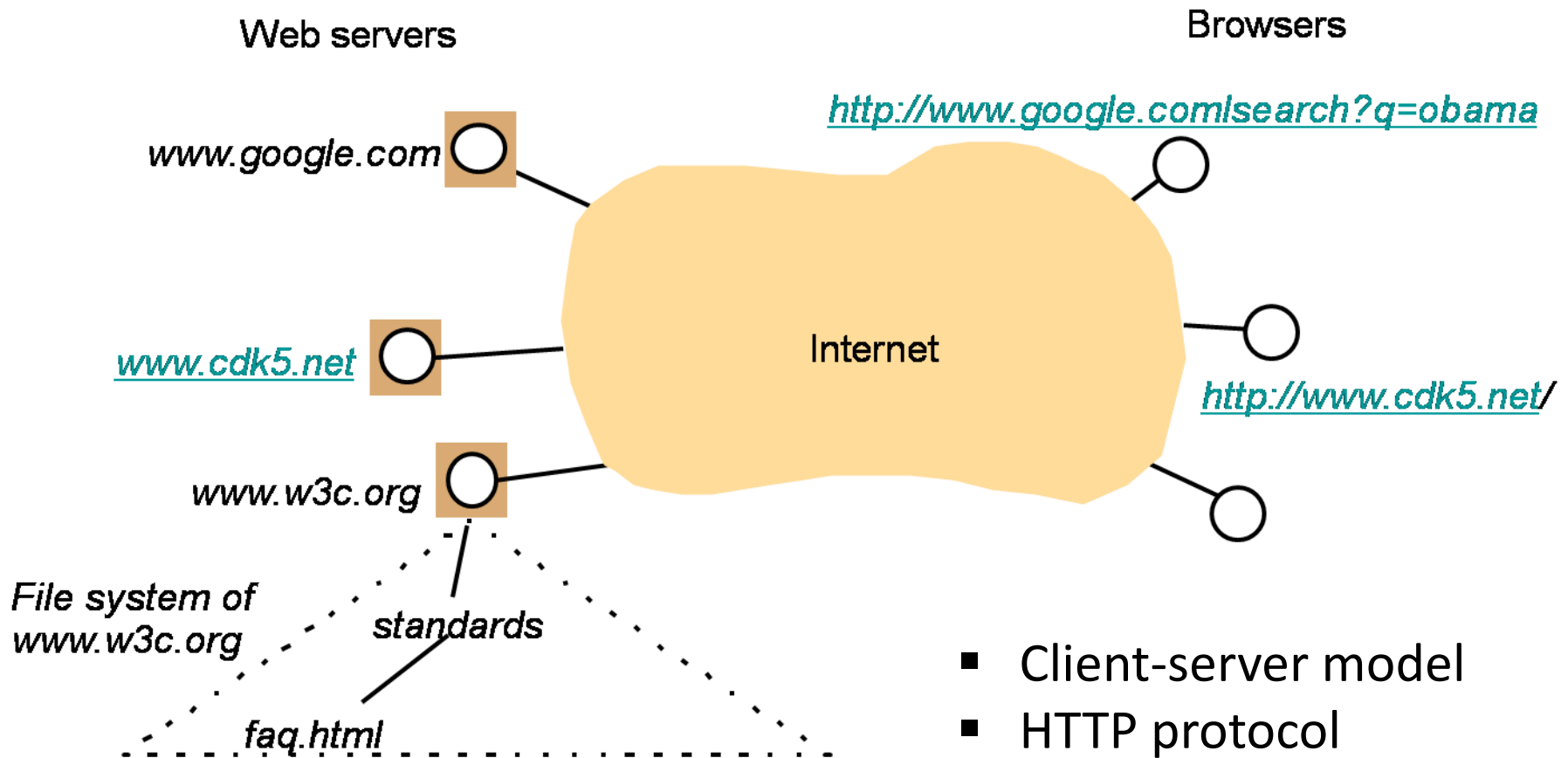
# Web: τι είναι

**Web** (World Wide Web, WWW, W3)

- μια συλλογή από **web σελίδες** (ιστοσελίδες) που είναι έγγραφα κειμένου και άλλες πηγές συνδεδεμένα με *hyperlinks* και *URLs*
- μια εφαρμογή που τρέχει πάνω από το Internet
- **63 δισεκατομμύρια** ιστοσελίδες
- **1 τρισεκατομμύριο** διαφορετικές web διευθύνσεις

<https://www.worldwidewebsite.com/>

# Web: η δομή του



- Client-server model
- HTTP protocol
- HTML
- URL/URI

# Web: Ιστορία

Στο τεύχος του **Ιουνίου 1970** του περιοδικού *Popular Science*

## **Arthur C. Clarke**

*satellites would one day "bring the accumulated knowledge of the world to your fingertips" using a console that would combine the functionality of the Xerox, telephone, television and a small computer, allowing data transfer and video conferencing around the globe.*

# Web: Ιστορία

1980, **Tim Berners-Lee** (ENQUIRE)



Νοέμβριο 1990, με τον *Robert Cailliau*, πρόταση για ένα "Hypertext project με το όνομα "WorldWideWeb" ("W3"): *"web" of "hypertext documents" to be viewed by "browsers" using a client-server architecture.*

Χριστούγεννα 1990, το πρώτο λειτουργικό σύστημα:

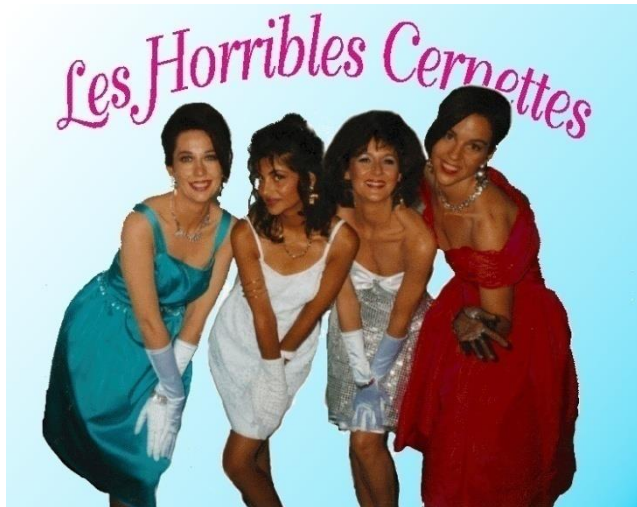
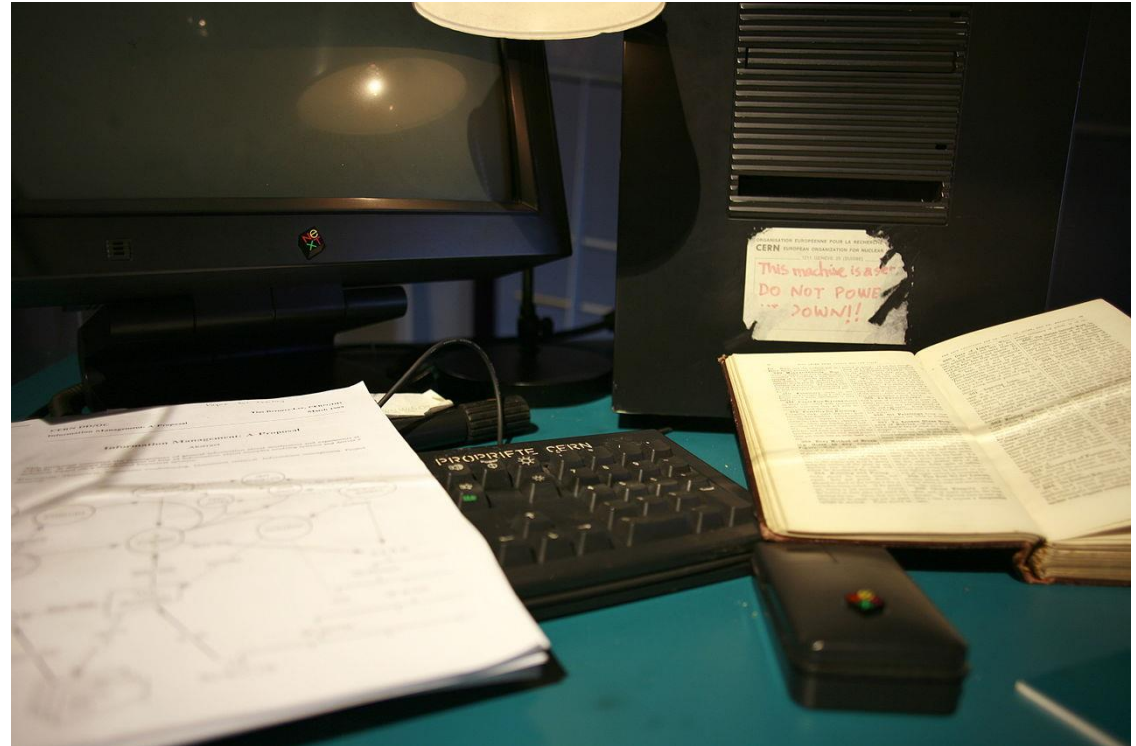
- ο πρώτος web browser (που ήταν και web editor),
- ο πρώτος web server, και
- οι πρώτες ιστοσελίδες, που περιέγραφαν το ίδιο το project.

Αύγουστο 1991, post στο alt.hypertext newsgroup – νέο service στο Ίντερνετ

# Web: Ιστορία

Ο πρώτος web server (και πρώτος web browser): A NeXT Computer

Η πρώτη φωτογραφία στο web το 1992 (CERN house band Les Horribles Cernettes)



logo by Robert Cailliau

Mosaic (1993) πρώτος graphical browser

# Web: Ιστορία



2016 ACM Turing Award

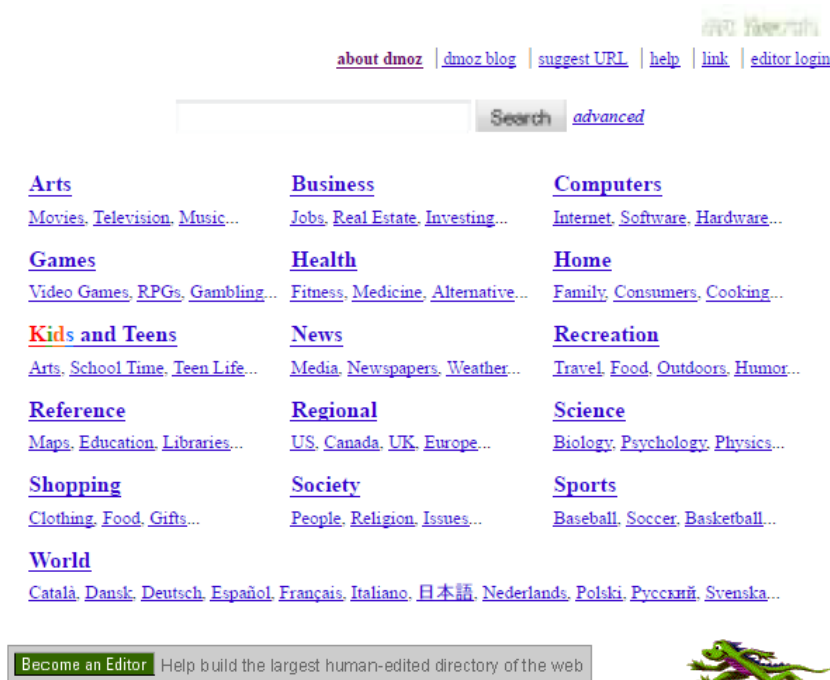
Time's magazine list of the 100 most influential people of the 20<sup>th</sup> century



# Εύρεση Πληροφορίας

Πρώτη προσπάθεια: [Ευρετήρια \(κατάλογοι\)](#) που τα διατηρούν άνθρωποι (μη αυτοματοποιημένα)

DMOZ - Open Directory Project



Copyright © 2012 Netscape

5,114,642 sites - 96,895 editors - over 1,014,858 categories



As of Mar 17, 2017, dmoz.org is no longer available.

# Εύρεση Πληροφορίας

## Welcome!

*This site includes information formerly made available via DMOZ.*

Visit [resource-zone](#) to stay in touch with the community.

*#OrganizeTheWeb*



### Arts

Movies, Television, Music...



### Business

Jobs, Real Estate, Investing...



### Computers

Internet, Software, Hardware...



### Games

Video Games, RPGs, Gambling...



### Health

Fitness, Medicine, Alternative...



### Home

Family, Consumers, Cooking...



### News

Media, Newspapers, Weather...



### Recreation

Travel, Food, Outdoors, Humor...



### Reference

Maps, Education, Libraries...



### Regional

US, Canada, UK, Europe...



### Science

Biology, Psychology, Physics...



### Shopping

Clothing, Food, Gifts...



### Society

People, Religion, Issues...



### Sports

Baseball, Soccer, Basketball...



### Kids & Teens Directory

Arts, School Time, Teen Life...



### World

Deutsch, Français, 日本語, Italiano, Español, Русский, Nederlands, Polski, Türkçe, Dansk, 简体中文, ...

<https://dmoz-odp.org/>

91,929  
Editors



1,031,722  
Categories



3,861,210  
Sites



90  
Languages



# Εύρεση Πληροφορίας

YAHOO! DIRECTORY Yahoo! | Help

Search:  the Web |  the Directory


---

Yahoo! Directory [Advanced Search](#) [Suggest a Site](#) [Email This Page](#)

<b>Arts &amp; Humanities</b> Photography, History, Literature...	<b>News &amp; Media</b> Newspapers, Radio, Weather, Blogs...
<b>Business &amp; Economy</b> B2B, Finance, Shopping, Jobs...	<b>Recreation &amp; Sports</b> Sports, Travel, Autos, Outdoors...
<b>Computer &amp; Internet</b> Hardware, Software, Web, Games...	<b>Reference</b> Phone Numbers, Dictionaries, Quotes...
<b>Education</b> Colleges, K-12, Distance Learning...	<b>Regional</b> Countries, Regions, U.S. States...
<b>Entertainment</b> Movies, TV Shows, Music, Humor...	<b>Science</b> Animals, Astronomy, Earth Science...
<b>Government</b> Elections, Military, Law, Taxes...	<b>Social Science</b> Languages, Archaeology, Psychology...
<b>Health</b> Disease, Drugs, Fitness, Nutrition...	<b>Society &amp; Culture</b> Sexuality, Religion, Food & Drink...
<b>New Additions</b> 12/3, 12/2, 12/1, 11/30, 11/29...	<b>Subscribe via RSS</b> Arts, Music, Sports, TV, more...

---

Copyright © 2012 Yahoo! Inc. All rights reserved. [Privacy Policy](#) - [About Our Ads](#) - [Terms of Service](#) - [Copyright/IP Policy](#)

 [Help us improve the Yahoo! Directory - Share your ideas](#)

officially closed on December 31, 2014

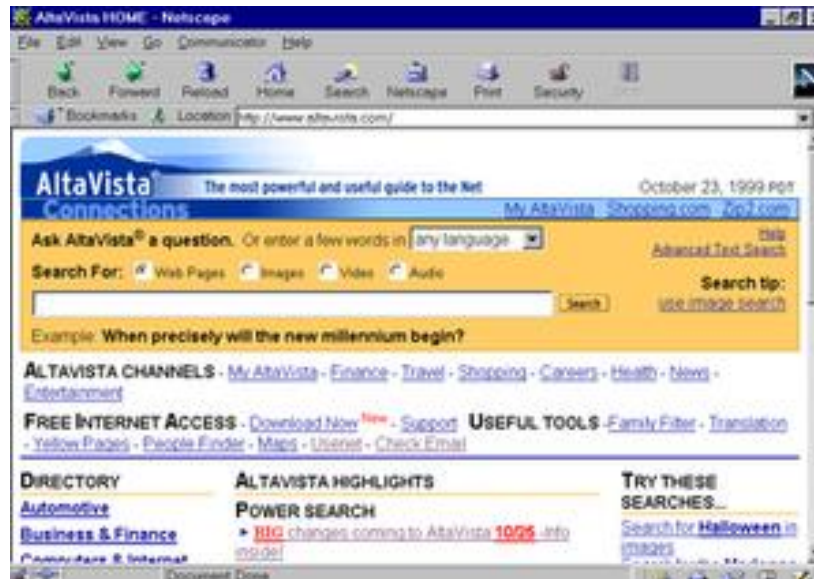
# Εύρεση Πληροφορίας

Δεύτερη προσπάθεια: Μηχανές αναζήτησης πλήρους κειμένου (full text)

Πόσο συχνά ο όρος εμφανίζεται στο κείμενο

Αρχικά οι χρήστες τις σελίδες τους για να ευρετηριοποιηθούν

Altavista



Infoseek

Type a specific question, phrase or Name.

  
the Web  [Tips](#)

To explore the Web's largest directory, click a topic below.



Excite

# Εύρεση Πληροφορίας

- Η εποχή του Google (~1998): χρήση του web ως γράφου
  - Πέρασμα από τη *συναφεια* στο *κύρος* (authoritativeness)
  - Δεν έχει μόνο σημασία μια σελίδα να είναι συναφής πρέπει να είναι και *σημαντική* στο web

Larry Page, Sergey Brin



# Εύρεση Πληροφορίας

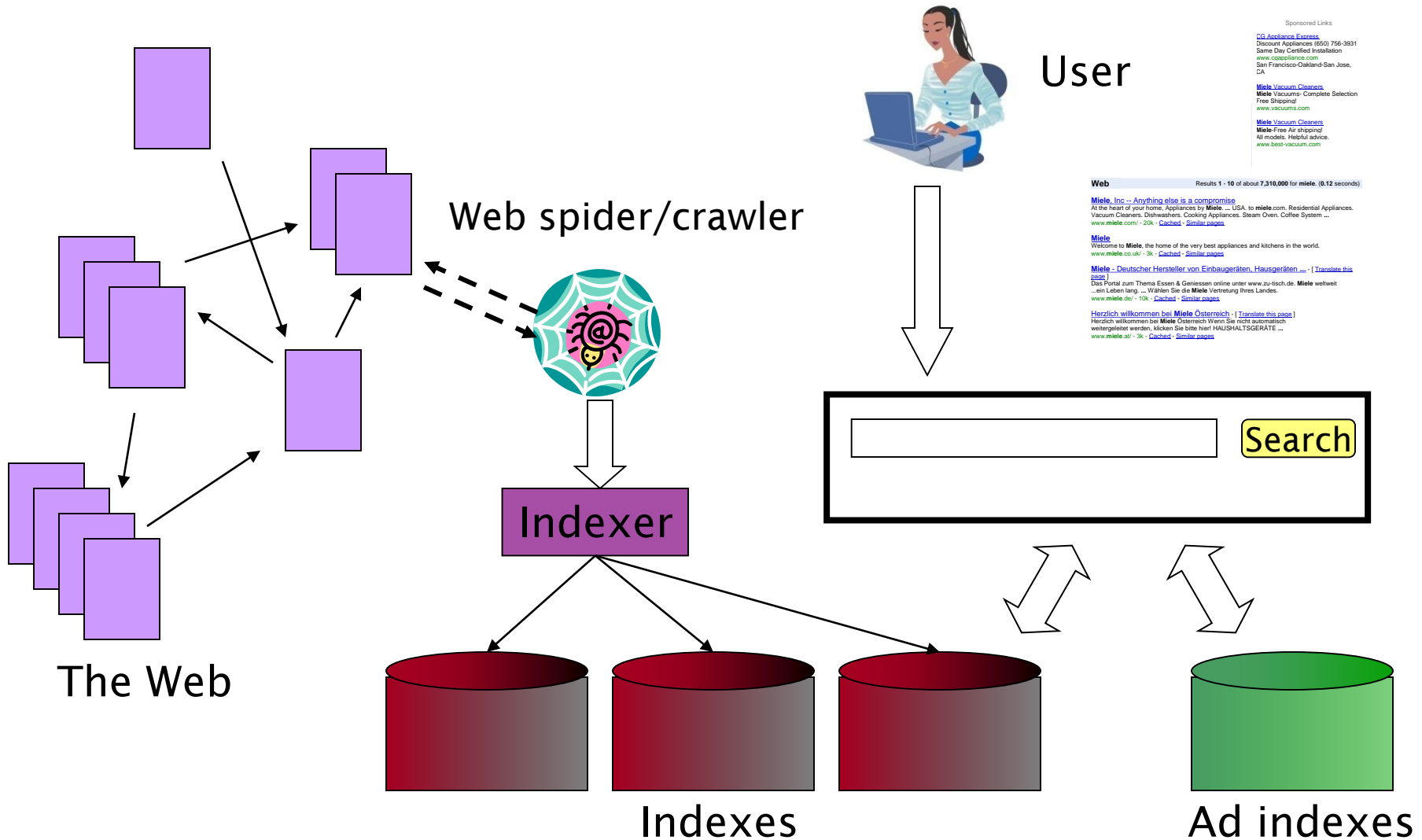


από τη λέξη **googol**, ο αριθμός 1 ακολουθούμενος από 100 μηδενικά

αρχικά στο website του Stanford University

[The anatomy of a large-scale hypertextual web search engine](#)

# Αναζήτηση στο Web: βασικά σημεία



# Δυναμικές και στατικές σελίδες

**Στατικές:** σελίδες που το περιεχόμενό τους δεν αλλάζει από τη μία αίτηση στην άλλη

**Δυναμικές** σελίδες: Hidden web – Deep web

- ✓ Παράδειγμα: προσωπική ιστοσελίδα vs σελίδα με την κατάσταση των πτήσεων σε ένα αεροδρόμιο

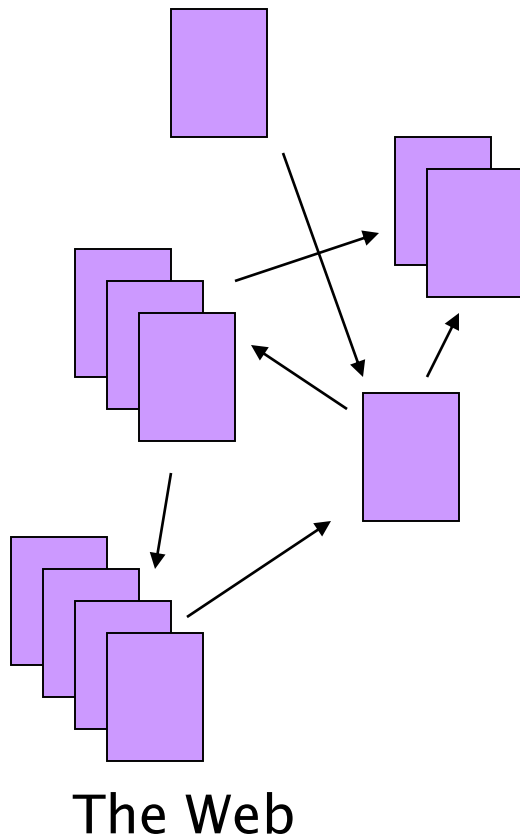
URL: συνήθως όχι κάποιο αρχείο αλλά κάποιο πρόγραμμα στον server

Input part of the GET, e.g., `http://www.google.com/search?q=obama`



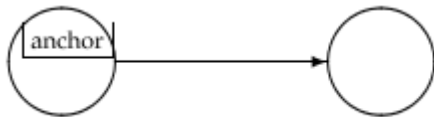


# Η συλλογή εγγράφων του Web



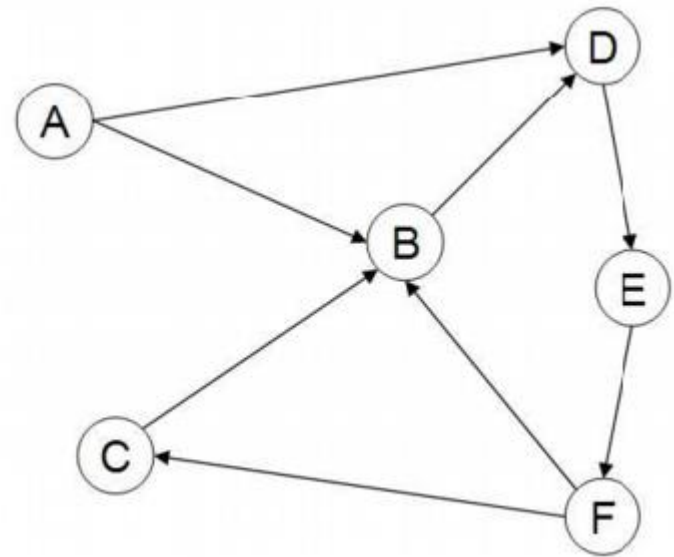
- No design/co-ordination
- *Distributed* content creation, linking, democratization of publishing
- Content includes *truth, lies*, obsolete information, contradictions ...
- *Unstructured* (text, html, ...), semi-structured (XML, annotated photos), structured (Databases)...
- *Scale* much larger than previous text collections
- *Growth* – slowed down from initial “volume doubling every few months” but still expanding
- Content can be *dynamically generated*

# Ο γράφος του Web



*Anchor text* `<a></a>`

- Κατευθυνόμενος, (In-links/Out-links)
- In-degree (8-15)
- Out-degree



# Ο γράφος του Web

- Η κατανομή των εισερχόμενων ακμών δεν ακολουθεί την κατανομή Poisson (την κατανομή που θα είχαμε αν κάθε σελίδα επέλεγε ποιες σελίδες θα κάνει link τυχαία (uniformly at random)).
- Κατανομή **Power law**,  
Το πλήθος των σελίδων με in-degree  $i$  είναι ανάλογο του  $1/i^\alpha$   
 $\alpha$  είναι συνήθως 2,1

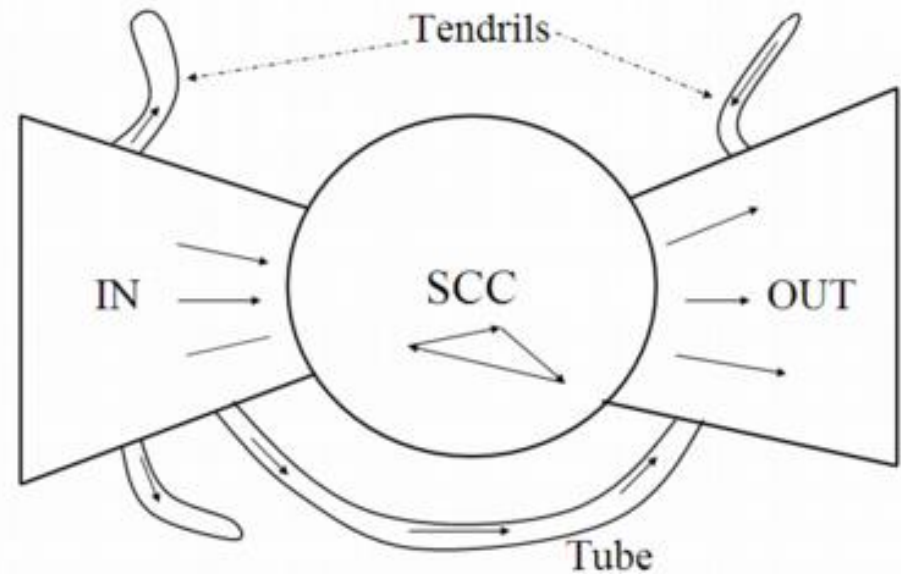
*Που αλλού είδαμε παρόμοια κατανομή;*

# Ο γράφος του Web

## Bow-tie shape

Τρεις κατηγορίες: **IN**, **OUT**, **SCC**

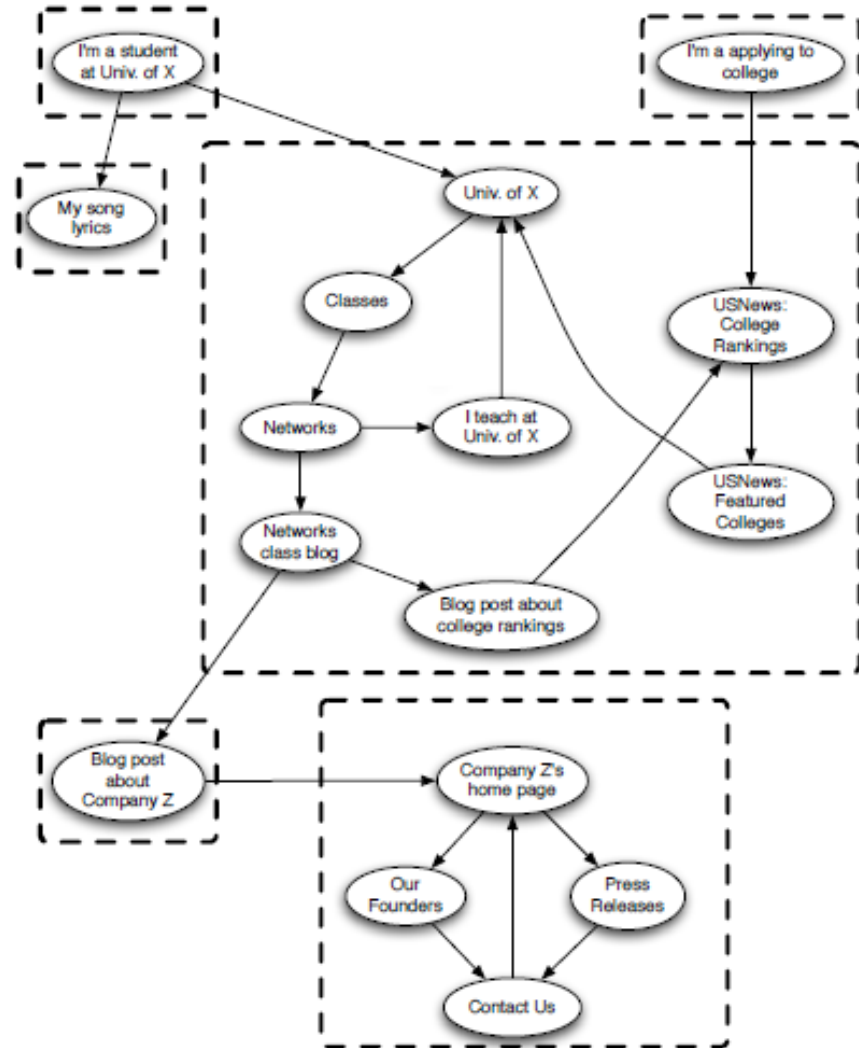
Περιέχει μια μεγάλη ισχυρά συνδεδεμένη συνιστώσα (Strongly Connected Component (**SCC**))



**IN**: Σελίδες που οδηγούν στο SCC αλλά όχι το ανάποδο

**OUT**: Σελίδες στις οποίες μπορούμε να φτάσουμε από το SCC αλλά δεν οδηγούν σε αυτό

# Ο γράφος του Web



From the book *Networks, Crowds, and Markets: Reasoning about Highly Connected World*. By David Easley and Jon Kleinberg. C University Press, 2010. Complete preprint on-line at <http://www.cs.cornell.edu/home/kleinber/networks-book/>

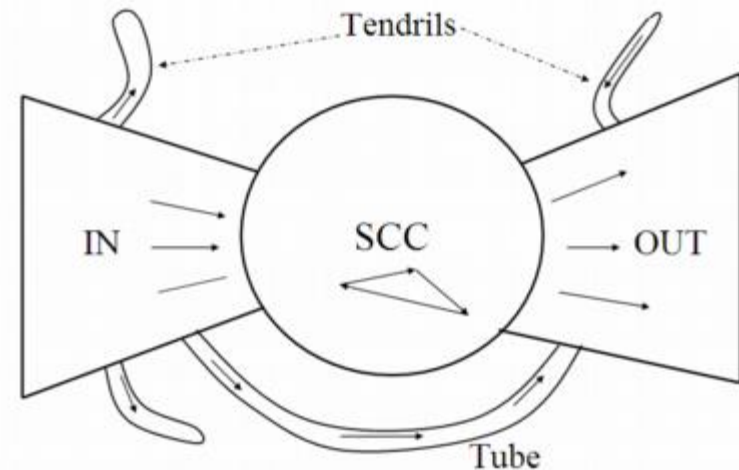
# Ο γράφος του Web

IN, OUT παρόμοιο μέγεθος, SCC μεγαλύτερο

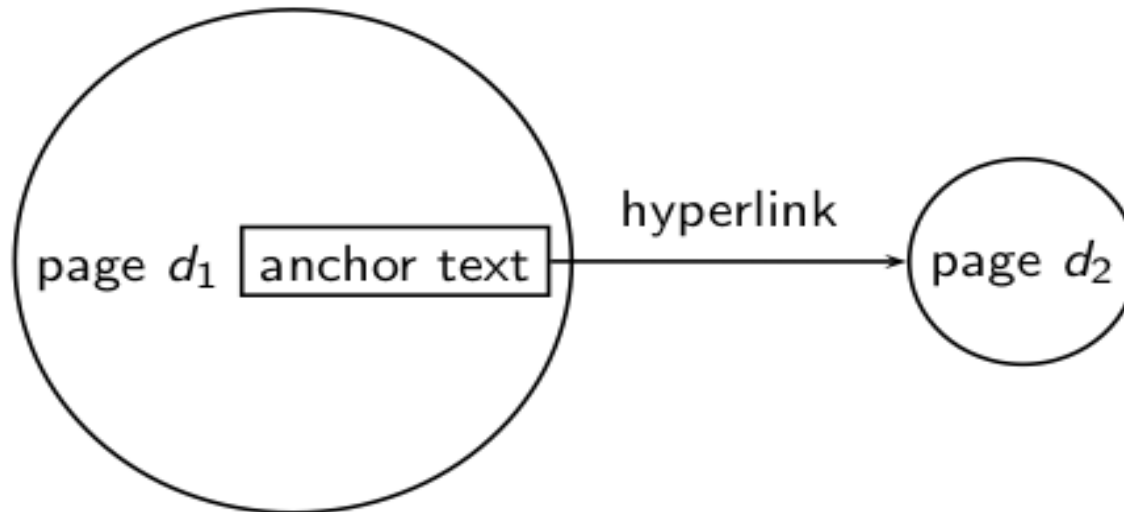
Υπόλοιπες σελίδες:

- Tubes: μικρά σύνολα σελίδων έξω από το SCC που δείχνουν απευθείας από το IN στο OUT,
- Tendrils: είτε δεν οδηγούν πουθενά από το IN είτε δείχνουν από πουθενά στο OUT.

Μικρές μη συνδεδεμένες  
συνιστώσες



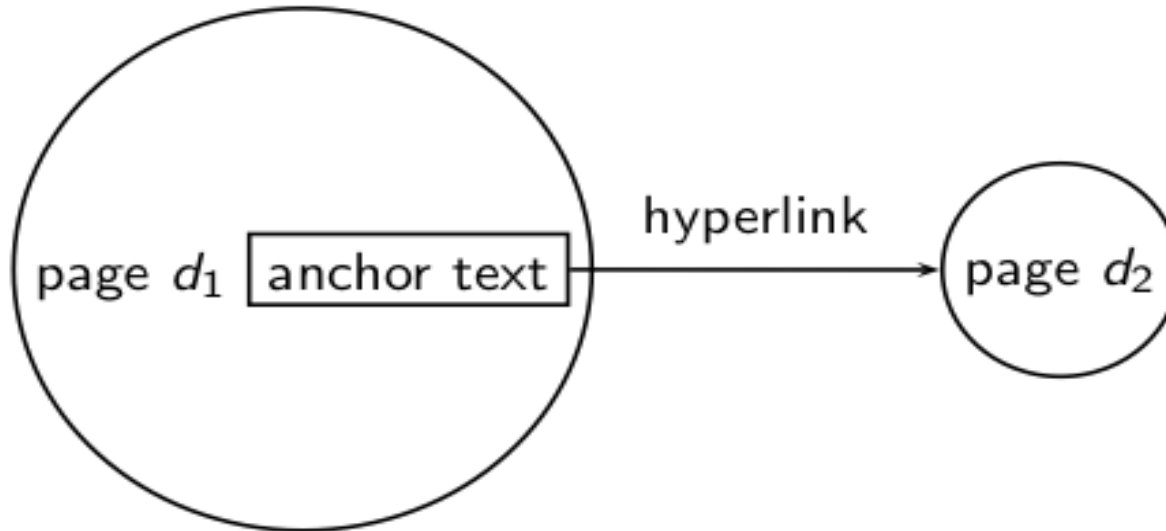
# Κείμενο Άγκυρας



**Anchor text (κείμενο άγκυρας)** κείμενο που περιβάλλει τον σύνδεσμο

- Παράδειγμα: “You can find cheap cars [- Παράδειγμα: “You can find \[- Anchor text: “You can find cheap cars here”\]\(http://...>\)](http://...>)

# Σημασία των συνδέσεων



- **1<sup>η</sup> Υπόθεση:** A hyperlink is a quality signal. (PageRank)
  - Η σύνδεση  $d_1 \rightarrow d_2$  υποδηλώνει ότι ο συγγραφέας του  $d_1$  θεωρεί το  $d_2$  καλής ποιότητας και συναφές.
- **2<sup>η</sup> Υπόθεση:** Το κείμενο της άγκυρας περιγράφει το περιεχόμενο του  $d_2$ .



# Κείμενο Άγκυρας

Χρήση μόνο **[text of  $d_2$ ]** ή **[text of  $d_2$ ] + [anchor text  $\rightarrow d_2$ ]**

- Αναζήτηση του **[text of  $d_2$ ] + [anchor text  $\rightarrow d_2$ ]** συχνά πιο αποτελεσματική από την αναζήτηση μόνο του **[text of  $d_2$ ]**
- Παράδειγμα: Ερώτημα *IBM*
  - Matches IBM's copyright page
  - Matches many spam pages
  - Matches IBM wikipedia article
  - *May not match IBM home page!* if IBM home page is mostly graphics

# Κείμενο Άγκυρας

- Αναζήτηση με χρήση του [anchor text  $\rightarrow d_2$ ] καλύτερη για το ερώτημα IBM
  - Η σελίδα με τις περισσότερες εμφανίσεις του όρου *IBM* στο anchor text είναι η [www.ibm.com](http://www.ibm.com)

[www.nytimes.com](http://www.nytimes.com): "IBM acquires Webify"

A million pieces of anchor text with "ibm" send a strong signal

[www.slashdot.org](http://www.slashdot.org): "New IBM optical chip"

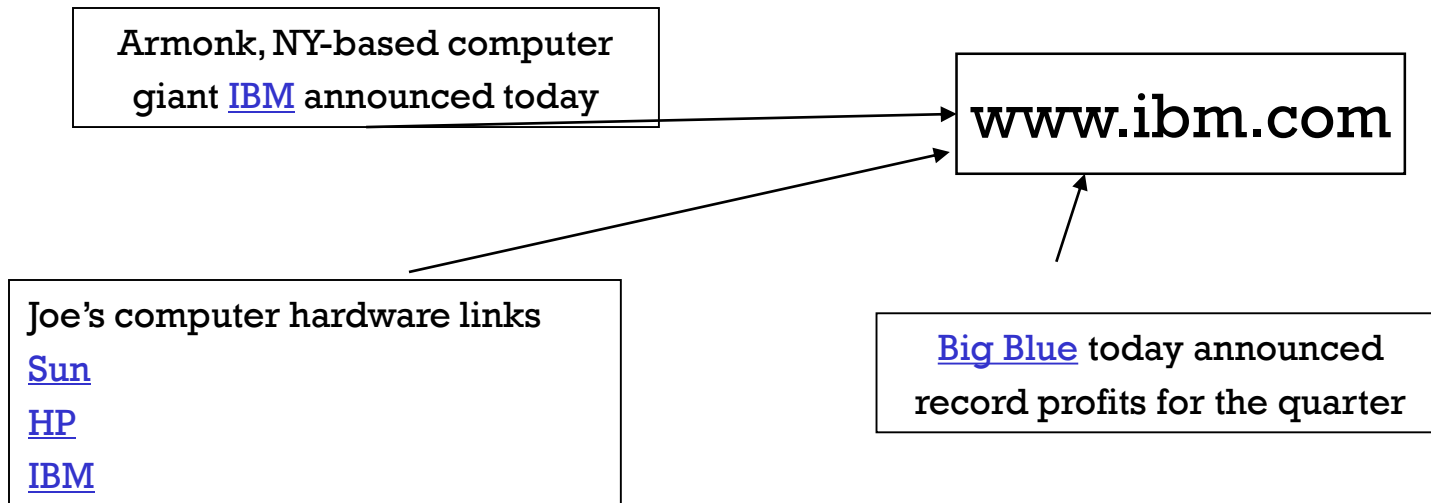
[www.stanford.edu](http://www.stanford.edu): "IBM faculty award recipients"

[www.ibm.com](http://www.ibm.com)

# Κείμενο Άγκυρας στο Ευρετήριο

Άρα: Το κείμενο στην άγκυρα αποτελεί καλύτερη περιγραφή του περιεχομένου της σελίδας από ό,τι το περιεχόμενό της

- Όταν κατασκευάζουμε το ευρετήριο για ένα έγγραφο  $D$ , συμπεριλαμβάνουμε (με κάποιο βάρος) και το κείμενο της άγκυρας των συνδέσεων που δείχνουν στο  $D$ .



- Χρήση idf για κοινούς όρους όπως Click, Here
- Επίσης, extended anchor text

# Google Bombs

**Google bomb:** μια αναζήτηση με «κακά» αποτελέσματα εξαιτίας κακόβουλης χρήσης κειμένου άγκυρας

Η Google εισήγαγε μια νέα συνάρτηση βαρών το Ιανουάριο του 2007

- *Βάρος στο κείμενο άγκυρας ανάλογα με την εγκυρότητα/κύρος της σελίδα που το περιέχει*
  - Miserable failure (Bush 2004)
- **Ακόμα κάποια υπόλοιπα:** [dangerous cult] στο Google, Bing, Yahoo
  - the Church of Scientology
- **Defused Google bombs:** [dumb motherf...], [who is a failure?], [evil empire] [cheerful achievement]

# Κείμενο Άγκυρας

- Άλλες εφαρμογές
  - Απόδοση βαρών/φιλτράρισμα στο γράφο
  - Για δημιουργία περιλήψεων μιας σελίδας

# Υπόθεση 2: annotation of target

The image shows two screenshots of the Tohoku University website. The top screenshot shows the navigation menu with the 'English' link circled in red. A green arrow points from this link to the bottom screenshot, which shows the 'English' link selected in the language dropdown menu. The bottom screenshot also shows a 'Click Here' button on a 'New! Video Channel' banner.

東北大学 TOHOKU UNIVERSITY

中文 | 한국어 | **English** | 日本語

お問い合わせ

大学概要 学部・大学院・研究所 教育・学生支援 国際交流 研究・産学連携

東北大学 TOHOKU UNIVERSITY

Chinese | Korean | English | Japanese

Search

Inquiry Access Sitemap

About Tohoku University Faculties, Schools and Institutes Campus Life International Exchange Research and Cooperation Disclosure and Public Information Entrance Exam Information

Prospective Students

General Public

Corporations

Alumni

Current Students

Faculty and Staff (Internal use)

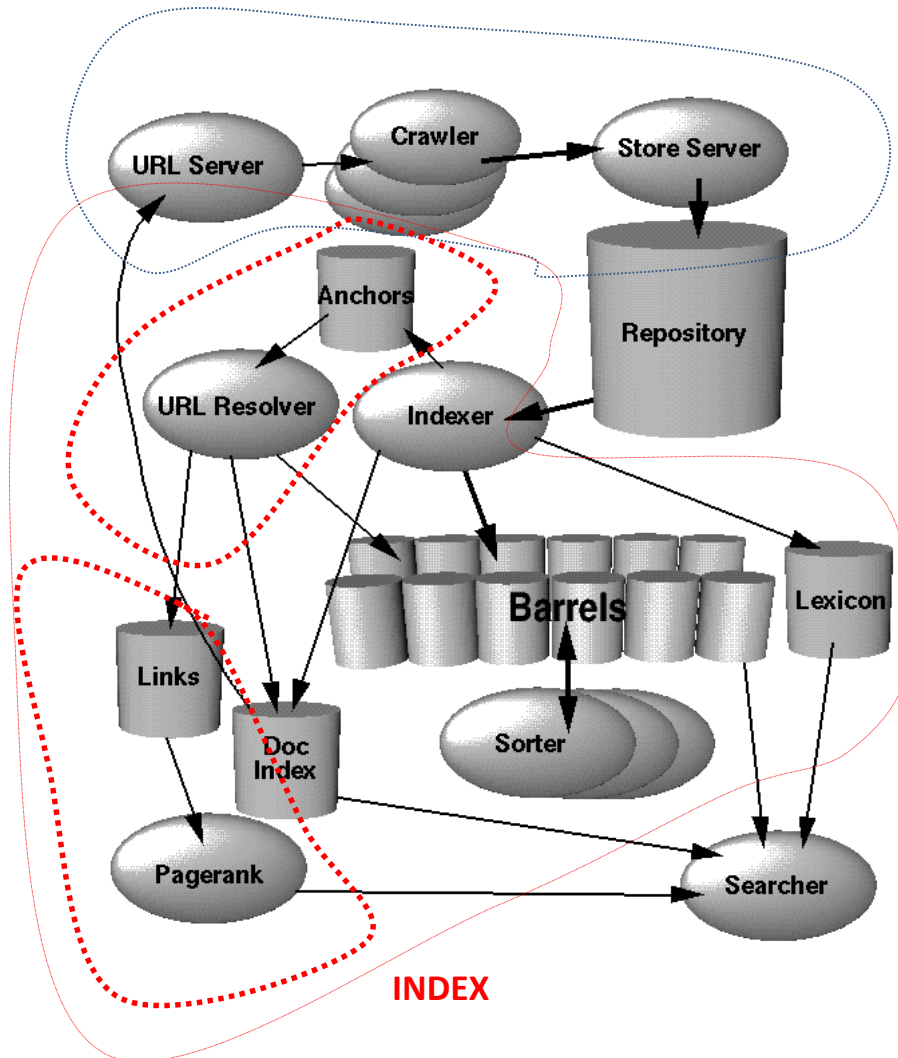
東北大学入学式(平成23年5月)

New! Video Channel

Click Here

# Search Engine Anatomy\*

## WEB CRAWLING



**BigFiles:** virtual files spanning multiple file systems  
**Repository:** full HTML of every web page  
**DocIndex:** info about files

### Indexer

**Hit list:** list of occurrences of a particular word in a particular document including position, font, and capitalization (fancy (title, anchor, etc), plain)

*a list of tokens per document*

Distributes the hits to **Barrels**

partially sorted **Forward index** (ranges of wordids)

**Sorter**

inverted index

**Anchores** -

# Τι θα δούμε σήμερα

- Web και μηχανές αναζήτησης
  - λίγη ιστορία
  - ο γράφος του web
  - σημασία της άγκυρας (anchor text)
- Χρήστες
- Εξατομίκευση
- Διαφημίσεις



# ΟΙ ΧΡΗΣΤΕΣ

# Ανάγκες Χρηστών

- Ποιοι είναι οι χρήστες;
- Μέσος αριθμός λέξεων ανά αναζήτηση 2-3
- Σπάνια χρησιμοποιούν τελεστές

# Ανάγκες Χρηστών

Need [Brod02, RL04]

- **Informational** (πληροφοριακά ερωτήματα) – θέλουν να μάθουν (learn) για κάτι (~40% / 65%)
    - Συνήθως, όχι μια μοναδική ιστοσελίδα, συνδυασμός πληροφορίας από πολλές ιστοσελίδες
- Low hemoglobin**
- **Navigational** (ερωτήματα πλοήγησης) – θέλουν να πάνε (go) σε μια συγκεκριμένη ιστοσελίδα (~25% / 15%)
    - Μια μοναδική ιστοσελίδα, το καλύτερο μέτρο = ακρίβεια στο 1 (δεν ενδιαφέρονται γενικά για ιστοσελίδες που περιέχουν τους όρους United Airlines)

**United Airlines**

# Ανάγκες Χρηστών

**Transactional** (ερωτήματα συναλλαγής) – θέλουν **να κάνουν (do)** κάτι (σχετιζόμενο με το web) (~35% / 20%)

- Προσπελάσουν μια υπηρεσία (Access a service)
- Να κατεβάσουν ένα αρχείο (Downloads)
- Να αγοράσουν κάτι
- Να κάνουν κράτηση

Seattle weather

Mars surface images

Canon S410

– **Γκρι περιοχές** (Gray areas)

- Find a good hub
- Exploratory search “see what’s there”

Car rental Brasil

# Typing Queries: παραδείγματα

Calculation: `5+4`

Unit conversion: `1 kg in pounds`

Currency conversion: `1 euro in kronor`

Tracking number: `8167 2278 6764`

Flight info: `LH 454`

Area code: `650`

Map: `columbus oh`

Stock price: `msft`

Albums/movies etc: `coldplay`

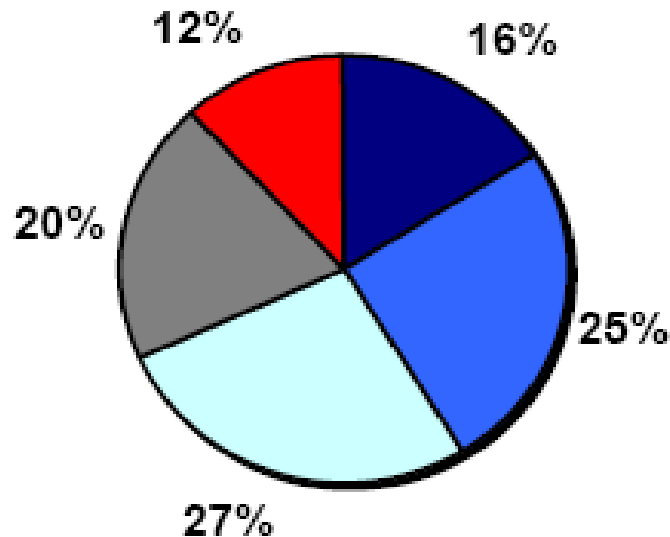
# Ανάγκες Χρηστών

Επηρεάζει (ανάμεσα σε άλλα)

- την καταλληλότητα του ερωτήματος για την παρουσίαση *διαφημίσεων*
- τον *αλγόριθμο/αξιολόγηση*, για παράδειγμα για ερωτήματα πλοήγησης (navigational) ένα αποτέλεσμα ίσως αρκεί, για τα άλλα (και κυρίως πληροφοριακά) ενδιαφερόμαστε για την περιεκτικότητα/ανάκληση

# Πόσα αποτελέσματα βλέπουν οι χρήστες

“When you perform a search on a search engine and don't find what you are looking for, at what point do you typically either revise your search, or move on to another search engine? (Select one)”



- After reviewing the first few entries
- After reviewing the first page
- After reviewing the first 2 pages
- After reviewing the first 3 pages
- After reviewing more than 3 pages

(Source: [iprospect.com](http://iprospect.com) WhitePaper\_2006\_SearchEngineUserBehavior.pdf)



## Percentage of Traffic by Google Results Position



Data Driven by Chitika Insights June 2013

Data is representative of Google search traffic within the U.S. and Canada only



# Τι ψάχνουν;

Δημοφιλή ερωτήματα

- <http://www.google.com/trends/hottrends>

Και ανά χώρα

Τα ερωτήματα ακολουθούν επίσης power law κατανομή

# Πως μπορούμε να καταλάβουμε τις προθέσεις (intent) του χρήστη;

Guess user intent *independent of context*:

- Spell correction
- Precomputed “typing” of queries

Better: Guess user intent *based on context*:

- Geographic context
- Context of user in this session (e.g., previous query)
- Context provided by personal profile

# Γεωγραφικό Περιεχόμενο

Τρεις σχετικές τοποθεσίες

1. Server (nytimes.com → New York)
2. Ιστοσελίδα (άρθρο των nytimes.com σχετικά με την Albania)
3. Χρήστης (βρίσκεται στο Palo Alto)

Εύρεση της τοποθεσίας του χρήστη

- IP address
- Πληροφορία που παρέχει ο χρήστης (πχ, στο user profile)
- Κινητό τηλέφωνο

*Geo-tagging*: Parse text and identify the coordinates of the geographic entities

Example: East Palo Alto CA → Latitude: 37.47 N, Longitude: 122.14 W

- ✓ NLP πρόβλημα

# Χρήση context

Πως μπορούμε να χρησιμοποιήσουμε τα συμφραζόμενα (context) για να τροποποιήσουμε τα αποτελέσματα ενός ερωτήματος;

- Result restriction: Αγνοούμε τα ακατάλληλα αποτελέσματα
  - Στο χρήστη της google.fr δείξε μόνο .fr αποτελέσματα
- Ranking modulation: χρησιμοποίησε μια γενική διάταξη και επανα-διέταξε τα αποτελέσματα με βάση τα συμφραζόμενα

Personalization/contextualization

# Τι άλλο θα δούμε

- Διαφημίσεις

# ΔΙΑΦΗΜΙΣΕΙΣ

# Διαφημίσεις

*Graphical graph banners* σε δημοφιλείς ιστοσελίδες (branding)

- ***cost per mil (CPM) model***: το κόστος προβολής του banner 1000 φορές (γνωστό και ως impressions)
- ***cost per click (CPC) model***: ο αριθμός των clicks στη διαφήμιση (που οδηγεί σε σελίδα από την οποία μπορεί να γίνει μια αγορά)
- *brand promotion vs transaction-oriented advertising*

# Ιστορία

- Αρχικές μηχανές βασισμένες σε λέξεις κλειδιά γύρω στο 1995-1997
  - Altavista, Excite, Infoseek, Inktomi, Lycos
- Paid search διάταξη: Goto (μετεξέλιξη σε Overture.com → Yahoo!)
  - Το αποτέλεσμα της αναζήτησης εξαρτάται από το πόσο πλήρωσες
  - Δημοπρασία (Auction) για λέξεις κλειδιά: η λέξη **casino** ακριβή!



# Διαφημίσεις στο Goto

Ως αποτέλεσμα μια ερώτησης  $q$ , η Goto

- επιστρέφει τις σελίδες των διαφημιζόμενων που πόνταραν (bid) για το  $q$ , σε διάταξη με βάση την προσφορά τους.
- Όταν ο χρήστης επέλεξε (clicked) ένα από τα αποτελέσματα, ο αντίστοιχος διαφημιζόμενος πλήρωνε τη Goto
  - Αρχικά, η πληρωμή ίση με την προσφορά bid για το  $q$

*Sponsored search* or *Search advertising*

# Διαφημίσεις στο Goto

www.goto.com/d/search;jsessionid=5A04214AAA00R50FIEF30FUQ?type=home&tm=1&Keywords=Wilmington+

Wilmington real estate.

Access 75% of all users now!  
Premium Listings reach 75% of all  
Internet users. [Sign up](#) for Premium  
Listings today!

- [Wilmington Real Estate - Buddy Blake](#)**  
Wilmington's information and real estate guide. This is your on  
anything to do with Wilmington.  
[www.buddyblake.com](http://www.buddyblake.com) (Cost to advertiser: **\$0.28**)
- [Coldwell Banker Sea Coast Realty](#)**  
Wilmington's number one real estate company.  
[www.cbseacoast.com](http://www.cbseacoast.com) (Cost to advertiser: **\$0.27**)
- [Wilmington, NC Real Estate Becky Bullard](#)**  
Everything you need to know about buying or selling a home c  
on my Web site!  
[www.iwwc.net](http://www.iwwc.net) (Cost to advertiser: **\$0.25**)

# Διαφημίσεις

Συνήθως παρέχονται

- ***pure search results*** (γνωστά και ως αλγοριθμικά ή ***organic search*** αποτελέσματα) ως βασική απάντηση στο ερώτημα του χρήστη,
- μαζί με ***sponsored search results*** τα οποία παρουσιάζονται ξεχωριστά και διακριτά στα δεξιά των αλγοριθμικών αποτελεσμάτων

Google  
Web Images Groups News Froogle Local more »  
nigritude ultramarine Search  
Advanced Search Preferences

Web Results 1 - 10 of about 185,000 for **nigritude ultramarine**. (0.35 seconds)

**Anil Dash: Nigritude Ultramarine**  
Do me a favor: Link to this post with the phrase **Nigritude Ultramarine**. ... Just placed a link to your **Nigritude Ultramarine** article on my weblog. Cheers! ...  
[www.dashes.com/anil/2004/06/04/nigritude\\_ultra](http://www.dashes.com/anil/2004/06/04/nigritude_ultra) - 101k - Mar 1, 2006 -  
[Cached](#) - [Similar pages](#)

**Nigritude Ultramarine FAQ**  
**Nigritude Ultramarine** FAQ - frequently asked questions about **nigritude ultramarine** and the realted SEO contest.  
[www.nigritudeultramarines.com/](http://www.nigritudeultramarines.com/) - 59k - [Cached](#) - [Similar pages](#)

**SEO contest - Wikipedia, the free encyclopedia**  
The **nigritude ultramarine** competition by SearchGuild is widely acclaimed as ...  
Comparison of search results for **nigritude ultramarine** during and after the ...  
[en.wikipedia.org/wiki/Nigritude\\_ultramarine](http://en.wikipedia.org/wiki/Nigritude_ultramarine) - 37k - [Cached](#) - [Similar pages](#)

**Slashdot | How To Get Googled, By Hook Or By Crook**  
The current 3rd result showcases the "**Nigritude Ultramarine** Fighting Force" who ... When discussing **nigritude ultramarine** [slashdot.org] it is important to ...  
[slashdot.org/article.pl?sid=04/05/09/1840217](http://slashdot.org/article.pl?sid=04/05/09/1840217) - 110k - [Cached](#) - [Similar pages](#)

**The Nigritude Ultramarine Search Engine Optimization Contest**  
It's sweeping the web -- or at least search engine optimizers -- a new contest to rank tops for the term **nigritude ultramarine** on Google.  
[searchenginewatch.com/sereport/article.php/3360231](http://searchenginewatch.com/sereport/article.php/3360231) - 57k - [Cached](#) - [Similar pages](#)

Paid Search Ads

Sponsored Links

**Business Blogging Seminar**  
...ing to L.A. March 16  
Top bloggers reveal key techniques  
[www.blogbusinesssummit.com](http://www.blogbusinesssummit.com)  
Los Angeles, CA

**Full-Time SEO & SEM Jobs**  
Find companies big & small hiring full-time SEO & SEM pros right now  
[CareerBuilder.com](http://CareerBuilder.com)

**SEO Contests**  
Information on SEO Contests like the **Nigritude Ultramarine** contest.  
[www.seo-contests.com/](http://www.seo-contests.com/)

**The SEO Book**  
**Nigritude Ultramarine** & SEO secrets  
Fun, free, raw, & different.  
[www.seobook.com](http://www.seobook.com)

Music Dance Electronic  
Overstock.com

Algorithmic results.



travel ioannina



Evaggelia

All Images Maps News Videos More Search tools



About 450,000 results (0.34 seconds)

**Aegean Airlines Flights - aegeanair.com**

www.aegeanair.com/ioannina
Book your flight to Ioannina Athens, Heraklion and Thessaloniki
Frequent Flights - 34 Greek Destinations - Direct Flights - Star Alliance
Destinations: Santorini, Athens, Thessaloniki, Heraklion, Naxos, Paros, Rhodes, Ioannina
Flight Booking & Check-in Book Flights Now
Aegean Destinations Low Fare Calendar

**Cheap Tickets to Ioannina - Φθηνά Αεροπορικά για Ιωάννινα**

www.airtickets.gr/ioannina Cheap Tickets
Κλείστε Φθηνά Εισιτήρια, Ώρα!
Ασφαλής Επικοινωνία Όλα τα Σεισοδοχεία Τηλεφωνική Εξυπηρέτηση Οι Φθηνότερες Πτήσεις
-30% Πτήση & Ενοδοχείο - Olympic για Ιωάννινα - Aegean για Ιωάννινα - Υπηρεσία Today®

**IOANNINA | OLD.NEW.YOU**

www.travelioannina.com
Welcome to Ioannina, Greece. A centuries old town on a beautiful lake in a ring of high mountains by the sunny beaches of the mediterranean sea.

**Pamvotis Travel - Γραφείο Γενικού Τουρισμού - Όλες οι αναχωρήσεις ...**

www.pamvotistravel.gr
Pamvotis Travel - Γραφείο Γενικού Τουρισμού - Όλες οι αναχωρήσεις από Ιωάννινα. Εκδρομές ταξίδια στο εξωτερικό και στο εσωτερικό. Επιστολείματα ...

**Ioannina Travel Guide - VirtualTourist**

https://www.virtualtourist.com/travel/ioannina/ioannina
Ioannina Travel Guide: 182 real travel reviews, tips, and photos from real travelers and locals in Ioannina, Greece at VirtualTourist.

**Travel Ioannina - Facebook**

www.facebook.com
Ioannina, Greece Tourist Information
Rating: 4.8 - 8 votes
Travel Ioannina, Ioannina, Greece. 1827 likes 119 talking about this 5 were here. Επιτροπή Τουριστικής Ανάπτυξης & Προβολής Δήμου Ιωαννίνων ...

**TRAVEL GUIDE: IOANNINA, GREECE - S Marks The Spots**

www.smarksthespots.com/travel-guide-ioannina-greece/
Discover the best sights, foodie spots and hidden gems in beautiful Ioannina!

**Travel Agencies Ioannina | vrisko.gr**

www.vrisko.gr/en/dir/travel-agencies/ioannina/
Looking for Travel Agencies in Ioannina? Find Travel Agencies in Ioannina in the largest Greek business directory.

**Armonia Travel**

armoniatravel.gr
Armonia Travel - Armonia Travel Αρχική Το γραφείο μας - Εκδρομές - Εκδρομές εσωτερικού - Εκδρομές εξωτερικού - Επικοινωνία - Menu ...

**Visit Greece | Ioannina**

www.visitgreece.gr/en/main\_cities/ioannina
Ioannina - a journey in a magical land! Ioannina, the capital of Epirus, spreads out around beautiful Lake Pamvotida. The natural ... Travel Information Before you ...

**DASKALOPOULOS YIANNIS - TOUR WORLD**

www.daskalopoulos-travel.com
Yiannis Daskalopoulos - Tour World is a travel agency in Ioannina, Epirus Greece. Ο Γιάννης Δασκαλόπουλος θεωρείται το τουριστικό γραφείο TOUR WORLD.

**Ioannina Tourism: Best of Ioannina, Greece - TripAdvisor**

www.tripadvisor.com
Ioannina Tourism: TripAdvisor has 7566 reviews of Ioannina hotels, ... Reviews and advice on hotels, resorts, flights, vacation rentals, travel packages, and more ...

Searches related to travel ioannina

ioannina greece photos
lake ioannina

2016

location



Neochoropoulo - From your Internet address - Use precise location - Learn more

Help Send feedback Privacy Terms

Περίπου 1.330.000 αποτελέσματα (0,65 δευτερόλεπτα)

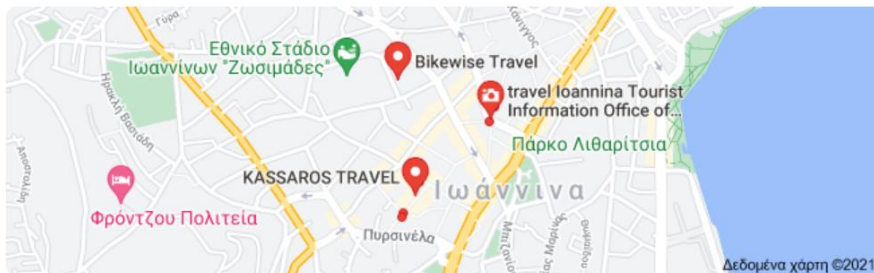
Μήπως εννοείτε: **ιωαννινα** travel

Βρείτε αποτελέσματα σε

Facebook  
[Travel Ioannina - Home](#)

Visitgreece  
[city of the Ioannina](#)

Lonely Planet  
[Ioannina travel](#)



Αξιολόγηση ▾
Ώρες ▾

travel Ioannina Tourist Information Office of Ioannina Municipality

4,3 ★★★★★ (19) · Γραφείο τουριστικών πληροφοριών

Μουσαιμής 10 · 2651 037017

Ανοιχτά τώρα

Ιστότοπος
 Οδηγίες

KASSAROS TRAVEL

4,2 ★★★★★ (24) · Ταξιδιωτικό πρακτορείο

3+ έτη επιχειρηματικής δραστηριότητας · Σακκά 6 · 2651 039274

Ανοιχτά · Κλείνει στις 9:00 μ.μ.

Ιστότοπος
 Οδηγίες

2021

## Kassaros Travel - Ταξιδιωτικό γραφείο στα Ιωάννινα.

Ταξιδιωτικό γραφείο Kassaros Travel Agency. Εκδρομές σε Ελλάδα και εξωτερικό όλη την διάρκεια του έτους.

<https://www.smarksthespots.com> > ... > [Μετάφραση αυτής της σελίδας](#)

## TRAVEL GUIDE: IOANNINA, GREECE - S Marks The Spots

So let me give you a brief tour of **Ioannina** (Ιωάννινα or Γιάννενα in Greek), the capital of the most mountainous province in Greece, namely Epirus.

2021

### Συναφείς αναζητήσεις

- ioannina city
- ioannina synagogue
- Koni travel ιωαννινα
- ioannina refugee camp
- ioannina population
- ioannina, greece
- ioannina hotels
- ioannina to athens



Ελλάδα ● 441 00 - Από τη διεύθυνσή σας στο διαδίκτυο - Χρήση ακριβούς τοποθεσίας - Μάθετε περισσότερα

Βοήθεια Αποστολή σχολίων Απώρρητο Όροι

## Πληροφορίες για διαχείριση θέσης

[https://support.google.com/websearch/answer/179386?p=ws\\_settings\\_location&hl=el&visit\\_id=637575231184863460-3976430011&rd=1](https://support.google.com/websearch/answer/179386?p=ws_settings_location&hl=el&visit_id=637575231184863460-3976430011&rd=1)







white shoes



🔍 Δια 📄 Σελίδες 🛒 Αγορές 📁 Σελίδες 📁 Βιβλία 📄 Παραδόσεις 📄 Ρυθμίσεις 📄 Εγχειρίδιο

Πήδησε 2.140.000.000 αποτελέσματα (0,59 δευτερόλεπτα)

Διαφήμιση - <https://www.epapoutsia.gr> - 231 110 0017

### epapoutsia.gr - Shoes - Mid Season Sale έως -40%

Mid Season Sale έως -40%. Συλλογή επώνυμων παπουτσιών. δωρεάν παράδοση με courier. Άνετος και ασφαλής online αγοράς. Ευρήματα μεγάλων και γραμμάτων. Spring/Summer 2021. Παραγωγή Αποστολή εντός 24 ωρών. -10€ με το newsletter.

#### Γυναικία παπούτσια

Το μεγαλύτερο ηλεκτρονικό κατάστημα  
Παρέχουμε για τις υπηρεσίες μας

#### Ανδρικά παπούτσια

Δωρεάν παράδοση με courier  
Αγορές με έκπτωση στο epapoutsia.gr

#### Γυναικία σανδάλια

30 ημέρες για επιστροφή  
Αποστολή εντός 24 ωρών





#### Mother's Day έως -50%

Οι πιο μοντέρνες συλλογές  
Αποστολή εντός 24 ωρών

Οι γρήγοροι αναζητούν επίσης

- Best white sneakers
- White shoes
- Nike white shoes
- Kiddee white shoes
- Best white sneakers 2020
- White sneakers Women's
- Reebok white sneakers
- White sneakers slouutz

Διαφήμιση - Κάντε αγορές τώρα

 <p>Ανδρικά Παπούτσια... 66,00 € 99-€ epapoutsia.gr</p> <p>Από Klarna</p>	 <p>Yana OLD SKOOL 3cm... 77,00 € Sparto.gr</p> <p>Από Google</p>
 <p>SNEAKERS στίβου...</p>	 <p>Ανδρικά Αθλητικά...</p>

- Navigational
- Διαφήμιση
- Ελληνικό κατάστημα
- Οι χρήστες αναζητούν ακόμα



Αποκρύπτει ▾ Είδη ▾

**DION shoes**  
4.6 ★★★★★ (40) Κατάστημα υποδημάτων  
Λεωφόρος 2625 037581  
Ανοίγει - Κλείνει στις 2:30 μ.μ.  
Αγοράζει εντός καταστήματος - Παράδοση στο κατάστημα - Δωρεάν

**MINAT LINES shoes**  
4.8 ★★★★★ (54) Κατάστημα υποδημάτων  
Λεωφόρος - 2625 076177  
Ανοίγει ανοίγει - 10:00 π.μ.  
Αγοράζει εντός καταστήματος - Παράδοση στο κατάστημα - Δωρεάν

Famous Shoes Ιαθρονα

<https://www.gq.com/story/Μεταρρύθμιση-αυτής-της-σελίδας>  
**24 Best White Sneakers & Shoes For Men in 2021: Leather ...**  
27 Απρ 2021 — If you're looking for the best **white sneakers** for men—to wear with jeans, suits, chinos, shorts, and anything else—we've got you covered.

<https://www.finishline.com/> — Μεταρρύθμιση-αυτής-της-σελίδας  
**All White Shoes | White on White Sneakers | Finish Line**  
Shop online at Finish Line for **All White Shoes** to upgrade your look. Free Shipping on White-on-White buys like the Air Force 1, Air Maxes, and NMD.

#### Συναφείς αναζητήσεις

- Q **Best white sneakers**
- Q **Best white sneakers 2020**
- Q **White shoes**
- Q **White sneakers Women's**
- Q **New white shoes**
- Q **Heebok white sneakers**
- Q **Adidas white shoes**
- Q **White sneakers skirts**

### Άλλες ερωτήσεις χρηστών

What are the best white shoes?



Are white shoes in Style 2021?



Should I get white sneakers?



How do you choose white sneakers?



Σχόλια

Διαφορετικό layout αν η ερώτηση στα ελληνικά:  
Άσπρα παπούτσια

# Διαφημίσεις

- **Search Engine Marketing (SEM)**

Κλάδος της διαφήμισης: κατανόηση του πως γίνεται η διάταξη των αποτελεσμάτων και ποιο ποσό της διαφημιστικής καμπάνιας να δοθεί σε ποιες λέξεις και σε ποιες μηχανές

**Click spam:** clicks σε sponsored search results από μη πραγματικούς χρήστες

πχ, από έναν αντίπαλο διαφημιστή

# Διαφημίσεις

**Paid inclusion:** πληρωμή για να συμπεριληφθεί μια σελίδα στο ευρετήριο της μηχανής αναζήτησης

Διαφορετικές μηχανές ακολουθούν διαφορετικές πολιτικές σχετικά με το αν επιτρέπουν *paid inclusion* και αν αυτό επηρεάζει τη διάταξη των αποτελεσμάτων της αναζήτησης.

*Παρόμοια προβλήματα με τηλεόραση/εφημερίδες*

# Διάταξη διαφημίσεων

- Οι διαφημιστές κάνουν *προσφορές για λέξεις κλειδιά* σε δημοπρασίες
- *Ανοικτό σύστημα*: Οποιοσδήποτε μπορεί να συμμετάσχει και να κάνει προσφορές
- Οι διαφημιστές χρεώνονται μόνο όταν κάποιος επιλέγει (*clicks*) στη διαφήμισή τους
- Σημαντική περιοχή για τις μηχανές αναζήτησης – *computational advertising*.
  - μια μικρή αύξηση της χρέωσης σε κάθε διαφήμιση μπορεί να οδηγήσει σε δις επιπρόσθετο κέρδος

# Διάταξη διαφημίσεων

Πως αποφασίζεται μέσω της δημοπρασίας η θέση της διαφήμισης και η τιμή

- Βασίζεται στο **second price auction**

# Google's second price auction

advertiser	bid	CTR	ad rank	rank	paid
A	\$4.00	0.01	0.04	4	(minimum)
B	\$3.00	0.03	0.09	2	\$2.68
C	\$2.00	0.06	0.12	1	\$1.51
D	\$1.00	0.08	0.08	3	\$0.51

- **bid**: maximum bid του διαφημιστή για ένα click
- **CTR**: click-through rate: το ποσοστό των χρηστών που επιλέγουν (click on) μια διαφήμιση όταν η διαφήμιση προβάλλεται
  - το **CTR είναι μια μέτρηση της συνάφειας**
- **ad rank**:  $\text{bid} \times \text{CTR}$ : εξισορρόπηση ανάμεσα (i) σε πόσα χρήματα είναι διατεθειμένος κάποιος διαφημιστής να πληρώσει, και (ii) στο πόσο σχετική είναι η διαφήμιση
- **rank**: θέση στη διάταξη της δημοπρασίας
- **paid**: second price auction = price που πληρώνει ο διαφημιστής
  - *πως υπολογίζεται;*

# Google's second price auction

advertiser	bid	CTR	ad rank	rank	paid
A	\$4.00	0.01	0.04	4	(minimum)
B	\$3.00	0.03	0.09	2	\$2.68
C	\$2.00	0.06	0.12	1	\$1.51
D	\$1.00	0.08	0.08	3	\$0.51

Second price auction: Ο διαφημιστής πληρώνει το μικρότερο ποσό ώστε να διατηρήσει τη θέση του στη δημοπρασία (συν 1 cent).

$\text{price}_1 \times \text{CTR}_1 = \text{bid}_2 \times \text{CTR}_2$  (αυτό έχει ως αποτέλεσμα  $\text{rank}_1 = \text{rank}_2$ )

$$\text{price}_1 = \text{bid}_2 \times \text{CTR}_2 / \text{CTR}_1$$

$$p_1 = \text{bid}_2 \times \text{CTR}_2 / \text{CTR}_1 = 3.00 \times 0.03 / 0.06 = 1.50$$

$$p_2 = \text{bid}_3 \times \text{CTR}_3 / \text{CTR}_2 = 1.00 \times 0.08 / 0.03 = 2.67$$

$$p_3 = \text{bid}_4 \times \text{CTR}_4 / \text{CTR}_3 = 4.00 \times 0.01 / 0.08 = 0.50$$



# Λέξεις κλειδιά με τα μεγαλύτερα bids

από το <http://www.cwire.org/highest-paying-search-terms/> (2018)

\$69.1 mesothelioma treatment options

\$65.9 personal injury lawyer michigan

\$62.6 student loans consolidation

\$61.4 car accident attorney los angeles

\$59.4 online car insurance quotes

\$59.4 arizona dui lawyer

CPC: cost per

\$46.4 asbestos cancer

click

\$40.1 home equity line of credit

\$39.8 life insurance quotes

\$39.2 refinancing

\$38.7 equity line of credit

\$38.0 lasik eye surgery new york city

\$37.0 2nd mortgage

\$35.9 free car insurance quote

For 2020 check

<https://www.wordstream.com/articles/most-expensive-keywords>

Keyword	Cost per Click (CPC)
Insurance	\$54.91
Loans	\$44.28
Mortgage	\$47.12
Attorney	\$47.07
Credit	\$36.06
Lawyer	\$42.51
Donate	\$42.02
Degree	\$40.61
Hosting	\$31.91
Claim	\$45.51
Conference Call	\$42.05

# Search ads: A win-win-win?

- Η **μηχανή αναζήτησης** κερδίζει χρήματα κάθε φορά που κάποιος επιλέγει (clicks) μια διαφήμιση
- Ο **χρήστης** επιλέγει μια διαφήμιση μόνο αν τον ενδιαφέρει
  - Οι μηχανές τιμωρούν μη συναφές ή παραπλανητικό περιεχόμενο
  - Ως αποτέλεσμα οι χρήστες συνήθως είναι ικανοποιημένοι από το τι βρίσκουν όταν επιλέγουν μια διαφήμιση
- Ο **διαφημιστής** βρίσκει νέους πελάτες με ένα αποδοτικό από άποψη κόστους τρόπο.

# Not a win-win-win: Keyword arbitrage

- Αγόρασε μια λέξη κλειδί στο Google
- Μετά επανα-προώθησε την κυκλοφορία σε κάποιον τρίτο που πληρώνει περισσότερα (ένα φτηνό keyword σε ένα ακριβό)
  - E.g., μια σελίδα γεμάτη διαφημίσεις
- Ad spammers keep inventing new tricks.
  - The search engines need time to catch up with them.

# Not a win-win-win: Violation of trademarks

- Example: geico
  - During part of 2005: The search term “geico” on Google was bought by competitors.
  - Geico lost this case in the United States.
  - Louis Vuitton lost similar case in Europe (2010).
- 
- It’s potentially misleading to users to trigger an ad off of a trademark if the user can’t buy the product on the site.

# Google adds options

Ρυθμίσεις διαφημίσεων

<https://adssettings.google.com/authenticated>

<http://www.youronlinechoices.com/gr/>

# Google search options

Settings: Advanced search

[https://www.google.com/advanced\\_search](https://www.google.com/advanced_search)

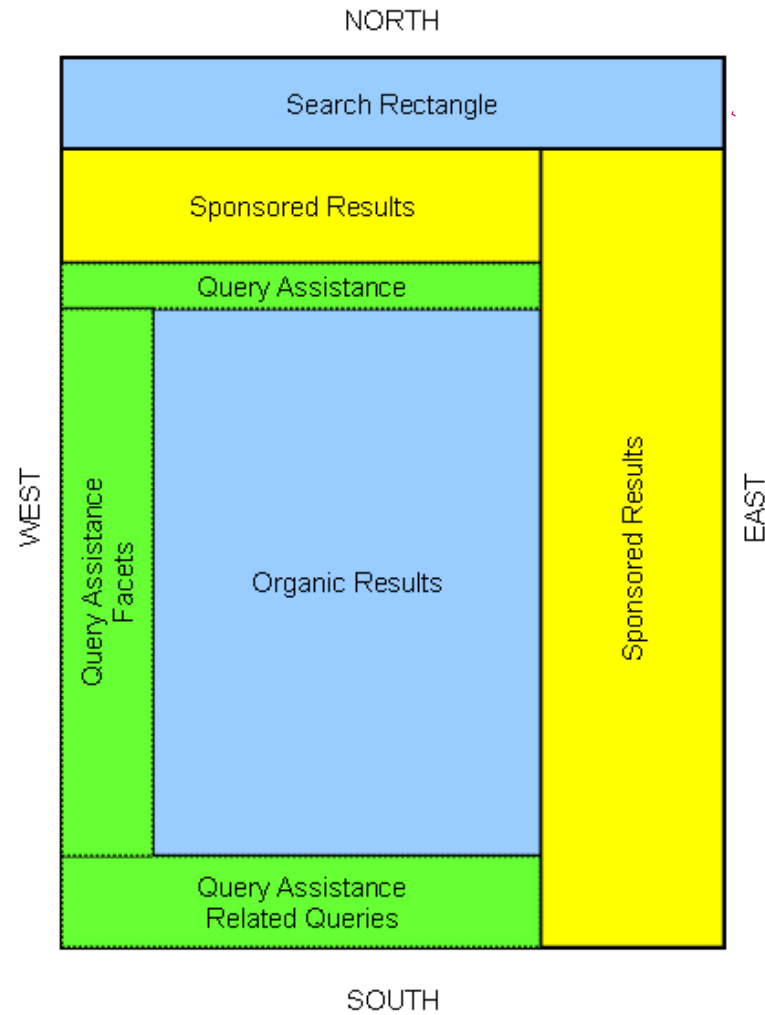
[https://www.google.com/advanced\\_image\\_search](https://www.google.com/advanced_image_search)

Turn on safe search

Hide private results

History

# SERP Layout





# ΤΕΛΟΣ (μέρους) 19<sup>ου</sup>, 20<sup>ου</sup> Κεφαλαίου

## Ερωτήσεις?

*Χρησιμοποιήθηκε κάποιο υλικό από:*

✓ *Pandu Nayak and Prabhakar Raghavan, CS276:Information Retrieval and Web Search (Stanford)*

✓ *Hinrich Schütze and Christina Lioma, Stuttgart IIR class*