



ΤΕΙ Ιονίων Νήσων
Τμήμα Τεχνολογίας Τροφίμων
Πληροφορική
Σημειώσεις
Τεύχος 6



Μάκης Σταματελάτος
makiss@teiion.gr



WEB – ΔΙΕΥΘΥΝΣΕΙΣ...

Ιστο-*.... (Web-*....)

- Ιστοσελίδα (webpage)
 - ένα είδος εγγράφου του παγκόσμιου ιστού (WWW) που περιλαμβάνει πληροφορίες με την μορφή κειμένου, υπερκειμένου, εικόνας, βίντεο και ήχου.
 - Πολλές ιστοσελίδες μαζί συνθέτουν έναν ιστότοπο (εναλλακτικές ονομασίες: ιστοχώρος ή δικτυακός τόπος, αγγλ. web site ή Internet site). Οι σελίδες ενός ιστοτόπου εμφανίζονται κάτω από το ίδιο όνομα χώρου (domain) π.χ. microsoft.com.
 - Οι ιστοσελίδες αλληλοσυνδέονται και μπορεί ο χρήστης να μεταβεί από τη μία στην άλλη κάνοντας «κλικ», επιλέγοντας δηλαδή συνδέσμους που υπάρχουν στο κείμενο ή στις φωτογραφίες της ιστοσελίδας.
- Ιστότοπος, ιστοχώρος ή διαδικτυακός τόπος (αγγλ. web site)
 - συλλογή από ιστοσελίδες, εικόνες, βίντεο και άλλα ψηφιακά στοιχεία, τα οποία φιλοξενούνται στο ίδιο domain (περιοχή) του Παγκόσμιου Ιστού. Βασίζεται στην υπηρεσία www (world wide web - παγκόσμιος ιστός), μια από τις υπηρεσίες που παρέχονται στο Διαδίκτυο, με τη χρησιμοποίηση του πρωτοκόλλου http. Η υπηρεσία αυτή δίνει τη δυνατότητα στους χρήστες του ίντερνετ να δημιουργήσουν οποιουδήποτε είδους περιεχόμενο στις ιστοσελίδες τους. Το σύνολο των ιστοτόπων αποτελεί το world wide web (www).
 - Κάθε ιστοσελίδα είναι συνήθως της μορφής www.τίτλος.TLD που υποδηλώνει ότι βασίζεται στην υπηρεσία www. Το πρόθεμα www δεν είναι αναγκαίο, αλλά χρησιμοποιείται συνήθως σαν το όνομα του διακομιστή (web server) ο οποίος παρέχει την υπηρεσία www.

Υπερκείμενα (hypertexts)

- Τρόπος οργάνωσης πληροφοριών (που βρίσκονται σε μορφή κειμένου) ο οποίος υπερβαίνει έναν βασικό περιορισμό των υπαρχόντων μέσων οργάνωσης: της γραμμικότητάς τους. Το πρόθεμα "υπέρ" σηματοδοτεί ακριβώς την υπέρβαση αυτή. Το υπερκείμενο θεωρείται μη γραμμικό κείμενο σε αντίθεση με το βιβλίο που θεωρείται γραμμικό.
- Η γραμμικότητα του βιβλίου έγκειται στο γεγονός ότι ο αναγνώστης οφείλει να διαβάσει τη μία σελίδα (ή παράγραφο) μετά την άλλη προκειμένου να κατανοήσει το περιεχόμενό του, δίχως να του δίδεται η δυνατότητα να "πλοηγείται" ελεύθερα εντός αυτού. Ο αναγνώστης έτσι αναγκάζεται να υπακούσει στους περιορισμούς που επιβάλλει ο συγγραφέας και το βιβλίο ως μέσο.
- Το υπερκείμενο ως μέσο έρχεται να ξεπεράσει τους περιορισμούς αυτούς επιτρέποντας την ελεύθερη πλοήγηση του αναγνώστη. Η πλοήγηση επιτυγχάνεται με την χρήση **υπερσυνδέσμων**. Επειδή οι υπερσύνδεσμοι αποτελούν τον κύριο μηχανισμό απόκλισης από τη γραμμικότητα αποτελούν θεμελιώδη έννοια στο υπερκείμενο. Ο Παγκόσμιος Ιστός χτίστηκε πάνω στις ιδέες του υπερκειμένου και αποτελεί μία ενσάρκωση (υλοποίηση) τέτοιου τρόπου διασύνδεσης και οργάνωσης πληροφοριών .

Διαδικτυακές Διευθύνσεις

- Τι είναι η διεύθυνση IP (Internet Protocol)
 - Κάθε συσκευή που συνδέεται σε ένα οποιοδήποτε δίκτυο, αποκτάει μία διεύθυνση IP: μία αριθμητική διεύθυνση που είναι μοναδική για τη συγκεκριμένη συσκευή στο συγκεκριμένο δίκτυο.
 - Αυτό ισχύει για οποιαδήποτε συσκευή, από τον υπολογιστή μας, το router, το κινητό, το tablet, το PlayStation, ή τον δικτυακό εκτυπωτή, μέχρι τους μεγάλους servers που φιλοξενούν ιστοσελίδες και υποστηρίζουν το Internet.
 - Μία IP διεύθυνση είναι ένας δυαδικός αριθμός 32-bit που για να γίνει περισσότερο κατανοητός στους ανθρώπους, χωρίζεται σε 4 ομάδες των 8 bit και κατόπιν κάθε ομάδα μεταφράζεται στον αντίστοιχο δεκαδικό αριθμό.
 - IPv4

IPv4

- Μια διεύθυνση IP περιέχει δύο κομμάτια πληροφορίας.
 - Ο αριθμός δικτύου στο οποίο ανήκει ο υπολογιστής: το Internet αποτελείται από πολλά διαφορετικά δίκτυα. Κάθε δίκτυο χαρακτηρίζεται από έναν μοναδικό αριθμό που αποτελεί την “ταυτότητά” του στο Internet.
 - Ένας τοπικός αριθμός υπολογιστή που προσδιορίζει τον υπολογιστή μέσα στο συγκεκριμένο δίκτυο.
 - 18.75.0.10 το 18 είναι ο αρ. δικτύου και το 75.0.10 είναι ο αρ. υπολογιστή
 - 147.102.154.12 το 147.102 είναι ο αρ. δικτύου και το 154.12 είναι ο αρ. υπολογιστή
- Οι διαφορετικές διευθύνσεις που θα μπορούσαν να παραχθούν, από τη 0.0.0.0 μέχρι τη 255.255.255.255 θα ήταν περίπου 4.3 δισεκατομμύρια (2^{32}).
- Με τόσους υπολογιστές, κινητά, tablets, κονσόλες, routers, web servers και λοιπές συσκευές με πρόσβαση στο Internet, οι διαθέσιμες διευθύνσεις IP έχουν εξαντληθεί από το Φεβρουάριο του 2011.

ip address

124.456.789.123

IPv6

- Για να αντιμετωπιστεί η εξάντληση των διευθύνσεων IPv4, έχει αναπτυχθεί μία νέα έκδοση του πρωτοκόλλου IP, με την ονομασία IPv6.
- Μία διεύθυνση IPv6 είναι της μορφής 2001:0db8:85a3:0042:1000:8a2e:0370:7334, και μπορούν με αυτή να παραχθούν πάνω από $3.4 \cdot 10^{38}$ διαφορετικές διευθύνσεις (2^{128}).
- Η μετάβαση από τις διευθύνσεις IPv4 στις διευθύνσεις IPv6 έχει καθυστερήσει σημαντικά, καθώς απαιτούνται σημαντικές αλλαγές σε επίπεδο υποδομής και εξοπλισμού στους παρόχους Internet (Internet Service Providers, ISP).

Πώς μεταφράζεται η IP σε όνομα



- η IP της Google είναι 173.194.66.94.
- Για να μπούμε στο Google ή σε οποιαδήποτε άλλη διεύθυνση, θα έπρεπε να γράψουμε την IP της στη μπάρα των διευθύνσεων.
- Αυτό όμως θα δυσκόλευε σημαντικά τη χρήση του Internet. Άντε να θυμάσαι τόσα νούμερα!
- Για το λόγο αυτό δημιουργήθηκαν οι DNS Servers (Domain Name System Servers). Η δουλειά τους είναι να αντιστοιχούν την διεύθυνση www.google.com στην IP 173.194.66.94, και να κάνουν το ίδιο για οποιαδήποτε κατοχυρωμένο όνομα domain στο Internet.

Εσωτερική IP, Εξωτερική IP

- Εξωτερική Διεύθυνση IP
 - Ένας υπολογιστής που συνδέεται απευθείας στο Internet με ένα modem (και άρα δεν ανήκει σε τοπικό δίκτυο) έχει μια εξωτερική διεύθυνση IP στο Internet. Το ίδιο και ένα router.
 - Σε διεθνές επίπεδο, η ρυθμιστική αρχή για όλες τις εξωτερικές διευθύνσεις IP είναι ο οργανισμός Internet Assigned Numbers Authority (IANA).
 - Ο IANA μοιράζει τις διαθέσιμες εξωτερικές IP στα 5 Regional Internet Registries, το καθένα εκ των οποίων αναλαμβάνει μια μεγάλη γεωγραφική περιοχή.
 - Οι RIRs αναθέτουν εξωτερικές διευθύνσεις IP ανατίθενται στους αντίστοιχους παρόχους Internet και άλλους οργανισμούς, προκειμένου να τις μοιράσουν στους χρήστες τους.
 - Με δεδομένο πως είναι γνωστό ποια μπλοκ διευθύνσεων έχουν διατεθεί σε ποια χώρα, με αυτόν τον τρόπο μπορεί να καταλάβει μια ιστοσελίδα, όπως το YouTube, σε ποια χώρα βρισκόμαστε.



Εσωτερική IP, Εξωτερική IP

- Εσωτερική Διεύθυνση IP
 - Εσωτερική IP ονομάζεται η διεύθυνση που έχει μια συσκευή συνδεδεμένη ένα τοπικό δίκτυο (Local Area Network, LAN).
 - Τις διευθύνσεις αυτές τις διαχειρίζεται το router και έχουν τρεις πιθανές μορφές:
 - 192.168.XXX.XXX
 - 172.16-31.XXX.XXX
 - 10.XXX.XXX.XXX
 - Όλες οι διευθύνσεις που ξεκινούν με τους παραπάνω αριθμούς είναι "ιδιωτικές διευθύνσεις" και δεν μπορούν να χρησιμοποιηθούν στο Internet.



IP διευθύνσεις στην Ελλάδα

<u>From IP</u>	<u>To IP</u>	<u>Total IPs</u>	<u>Assign Date</u>	<u>Owner</u>
2.84.0.0	2.87.255.255	262144	12/05/10	Ote SA (Hellenic Telecommunications Organisation)
5.54.0.0	5.55.255.255	131072	29/05/12	Hellas OnLine Electronic Communications S.A.
5.144.192.0	5.144.255.255	16384	11/07/12	Cosmote Mobile Telecommunications S.A.
5.172.192.0	5.172.207.255	4096	31/07/12	LANCOM LTD.
5.203.0.0	5.203.255.255	65536	30/08/12	Cosmote Mobile Telecommunications S.A.
31.152.0.0	31.152.255.255	65536	21/02/11	Cosmote Mobile Telecommunications S.A.
31.217.160.0	31.217.191.255	8192	17/08/11	Cosmote Mobile Telecommunications S.A.
37.6.0.0	37.6.255.255	65536	12/12/11	TELLAS Telecommunication Services S.A.
37.32.128.0	37.32.255.255	32768	03/01/12	Greek Research and Technology Network S.A
46.12.0.0	46.12.255.255	65536	14/09/10	FORTHnet SA
46.176.0.0	46.177.255.255	131072	06/12/10	Hellas OnLine Electronic Communications S.A.
46.190.0.0	46.190.127.255	32768	07/12/10	TELLAS Telecommunication Services S.A.
46.246.128.0	46.246.255.255	32768	01/02/11	FORTHnet SA
62.1.0.0	62.1.255.255	65536	16/02/00	FORTHnet SA
62.38.0.0	62.38.255.255	65536	30/05/01	Hellas OnLine Electronic Communications S.A.
62.68.64.0	62.68.95.255	8192	15/06/00	VP Internet Hellas S.A.

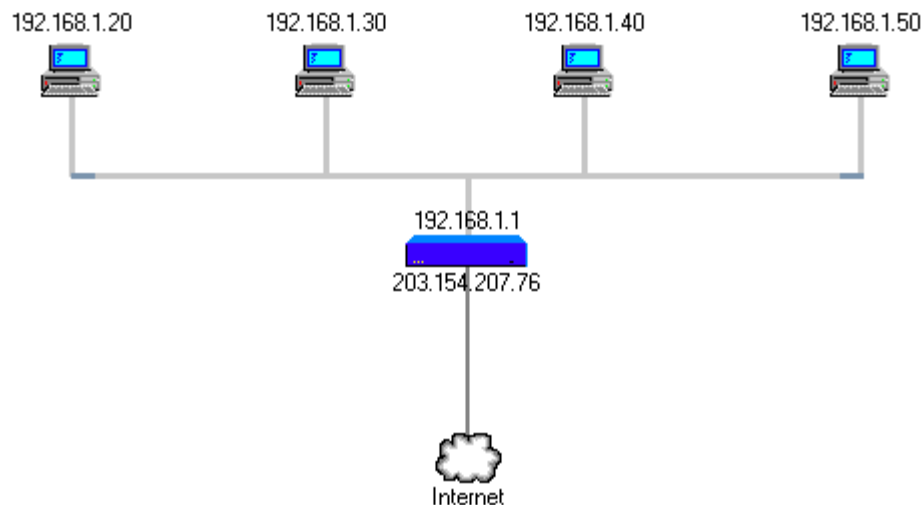
Οι διευθύνσεις IP του Router



- Το router είναι μια ιδιαίτερη συσκευή, καθώς ανήκει ταυτόχρονα σε δύο δίκτυα.
- Κάθε οικιακό router είναι ένας συνδυασμός modem/router/switch και άλλων συσκευών.
- Σε ένα μεγάλο δίκτυο (πχ ένα εταιρικό δίκτυο ή έναν πάροχο ίντερνετ) οι συσκευές αυτές θα ήταν ξεχωριστές (rack).
- Λειτουργία
 - Το Modem κομμάτι του οικιακού router συνδέεται με το Internet και παίρνει μια εξωτερική διεύθυνση IP.
 - Το Switch κομμάτι διαχειρίζεται τη μεταφορά δεδομένων μεταξύ των συσκευών στο τοπικό δίκτυο, πχ μέσω μιας οικιακής ομάδας.
 - Το router κομμάτι αναλαμβάνει τη μεταφορά των δεδομένων ανάμεσα στα δύο ξεχωριστά δίκτυα: το Internet και το τοπικό δίκτυο.
- Κάθε οικιακό router έχει δύο IP: την εσωτερική για το LAN, που συνήθως είναι 192.168.1.1 ή 192.168.0.1 και την εξωτερική στο Internet.
- Οι υπολογιστές και οι άλλες συσκευές που είναι συνδεδεμένες στο τοπικό δίκτυο (κινητό, tablet κλπ) έχουν μόνο εσωτερικές IP και χρησιμοποιούν μία εξωτερική IP, αυτή του router.
- Πώς γίνεται όμως πάνω από μία συσκευές να έχουν την ίδια εξωτερική IP;

Network Address Translator (NAT)

- Το σύστημα του Network Address Translator αποτελεί μέρος του router.
- Πρόκειται για μια μεθοδολογία που αλλάζει την πληροφορία σχετικά με τη διεύθυνση IP σε κάθε πακέτο δεδομένων, καθώς περνούν από το router.
- Το NAT "μαρκάρει" από ποιόν υπολογιστή ήρθε το αίτημα για την πληροφορία, και γνωρίζει αν η πληροφορία από το Internet πρέπει να πάει στον υπολογιστή 192.168.0.20 ή στο κινητό 192.168.0.50 και τη δρομολογεί κατάλληλα.



Στατική IP και Δυναμική IP (Εξωτερική IP)

- Η βασική διαφορά της διεύθυνσης IP με την ταχυδρομική διεύθυνση είναι πως η IP μας δεν παραμένει αναγκαστικά σταθερή.
- Δυναμική IP
 - Η συντριπτική πλειοψηφία των συνδέσεων Internet που αγοράζουμε από τους παρόχους διαθέτουν δυναμική IP.
 - Αυτό σημαίνει πως κάθε φορά που συνδεόμαστε στο Internet, αν πχ κλείσουμε το router και το ξανανοίξουμε μετά από 1-2 λεπτά, μας ανατίθεται μία νέα εξωτερική διεύθυνση IP.
 - Ακόμα όμως κι αν δεν κλείσουμε το router, ο ίδιος ο πάροχος Internet μας αλλάζει τη διεύθυνση IP κάθε κάποιο διάστημα (συνήθως μερικές ημέρες).
- Στατική IP
 - Αν έχουμε μια σύνδεση με στατική IP, τότε η εξωτερική μας IP προς το Internet δεν αλλάζει ποτέ.
 - Οι πάροχοι Internet χρεώνουν επιπλέον για το προνόμιο μιας στατικής IP, για την ίδια ταχύτητα σύνδεσης. Συνήθως αποκαλούν τις συνδέσεις αυτές business ή Office.

Ποιος ο λόγος που χρεώνεται ακριβότερα η στατική IP;

- Αν έχουμε στατική IP στη σύνδεσή μας, μπορούμε από οπουδήποτε στον κόσμο να έχουμε πρόσβαση στο router μας, και κατ' επέκταση στο προσωπικό μας τοπικό δίκτυο.
- Μπορούμε πχ να έχουμε IP κάμερες για να παρακολουθούμε το χώρο μας από όπου κι αν είμαστε, να μπούμε εξ' αποστάσεως στον υπολογιστή μας και να έχουμε πρόσβαση στα αρχεία μας, ή ακόμα και να ανάψουμε τον σβηστό υπολογιστή μέσω Internet.

2play 4Mbps (upload 512Kbps)	2play 8Mbps (upload 1Mbps)	2play 24Mbps (upload 1Mbps)	2play 24Mbps (upload 1Mbps)	VDSL 2play 35Mbps (upload 3Mbps)	VDSL 2play 50Mbps (upload 10Mbps)
€16	€19	€23	€33	€39	€45
+ Δωρεάν εξοπλισμός Wi-Fi Modem / Router	+ Δωρεάν εξοπλισμός Wi-Fi Modem / Router	+ Δωρεάν εξοπλισμός Wi-Fi Modem / Router	+ 1 κανάλι φωνής & Απεριόριστα* Σταθερά Ελλάδα & 1 στατική IP	+ 1 κανάλι φωνής & Απεριόριστα* Σταθερά Ελλάδα & 1 στατική IP	+ 1 κανάλι φωνής & Απεριόριστα* Σταθερά Ελλάδα & 1 στατική IP



ΑΝΑΖΗΤΗΣΗ ΣΤΟ ΔΙΑΔΙΚΤΥΟ



ΤΕΙ ΙΟΝΙΩΝ
ΝΗΣΩΝ

Αναζήτηση Πληροφοριών στο Διαδίκτυο

- Το διαδίκτυο αποτελείται από εκατοντάδες εκατομμύρια ιστοσελίδες.
- Δύσκολη για ένα χρήστη η ανάκτηση της πληροφορίας που τον ενδιαφέρει.
- Ο όγκος της πληροφορίας είναι τεράστιος.
- Κίνδυνος να αναζητούμε μια πληροφορία και να περιπλανιόμαστε επί ώρες άσκοπα στον Παγκόσμιο Ιστό!



Ποιοτική Πληροφορία

- Αξιολογούμε τους δικτυακούς τόπους
- Αναζητούμε την επιστημονική ιδιότητα του συγγραφέα
- Εξετάζουμε την ποιότητα του άρθρου (μεθοδολογία, πηγές)
- Ελέγξτε το πόσο επίκαιρη είναι η πληροφορία (ημερομηνία εμφάνισης στο διαδίκτυο)





Υπηρεσίες Αναζήτησης

- Δύο (2) είδη υπηρεσιών Αναζήτησης:
 - Μηχανές Αναζήτησης (Search Engines)
 - Θεματικοί Κατάλογοι



Those were the days....

[Lycos Sports Fans Only \\$1,000,000 in Cash and Prizes!](#)
 For serious sports fans only! Play Fantasy Football!

LYCOS It's amazing where
 Go Get It will get you.

Find:

[Enhance your search.](#)





[New Search](#) - [Top News](#) - [Sites by Subject](#) - [Top 50 Sites](#) - [City Guide](#) - [Flowers & Sounds](#)
[FossilFind](#) - [Point Expense](#) - [Road Maps](#) - [Software](#) - [About Lycos](#) - [Child Lycos](#) - [Help](#)

[Add Your Site to Lycos](#)

Copyright © 1998 Lycos®, Inc. All Rights Reserved.
 Lycos is a trademark of Carnegie Mellon University.
[Questions & Comments](#)



YAHOO!

NEW! COOL! FUNDOO!
 NEWS YAHOO! ADD URL

Now Open: Yahoo! Start Page!  [Web Links](#)

- AIM**
[Messenger](#), [File Transfer](#), [Archives](#), ...
- Business and Economy (Beta)**
[Company](#), [Securities](#), [Classification](#), [Tools](#), ...
- Computers and Internet (Beta)**
[Internet](#), [WWW](#), [Software](#), [Hardware](#), ...
- Education**
[Thesaurus](#), [E-Sci](#), [Course](#), ...
- Entertainment (Beta)**
[TV](#), [Movies](#), [Music](#), [Magazines](#), ...
- Government**
[Politics](#) (Beta), [Awards](#), [Law](#), [Money](#), ...
- Health**
[PillSite](#), [Drugs](#), [Dieting](#), [Fitness](#), ...
- News (Beta)**
[World \(Beta\)](#), [Daily](#), [Current Events](#), ...
- Recreation**
[Sports \(Beta\)](#), [Games](#), [Trivia](#), [Arts](#), ...
- Reference**
[Library](#), [Dictionaries](#), [Phrase Books](#), ...
- Regional**
[Countries](#), [Europe](#), [U.S. News](#), ...
- Science**
[Ox](#), [Biology](#), [Astronomy](#), [Engineering](#), ...
- Social Science**
[Anthropology](#), [Psychology](#), [Economics](#), ...
- Society and Culture**
[Fairytales](#), [Folklore](#), [Idioms](#), ...

[Try Only Yahoo! - Countdown](#)



AltaVista HOME - BetaPage

File Edit View Go Communication Help

Back Forward Reload Home Search Mailings Print Security

Bookmarks Location [http://www.altavista.com/](#)

AltaVista The most powerful and useful guide to the Web October 23, 1999 (EST)

Connections [My AltaVista](#) [Connections](#) [Helpdesk](#)

Ask AltaVista® a question. Or enter a few words in (any language)

Search For:

Search For:

Example: When precisely will the new millennium begin?

ALTAVISTA CHANNELS - [My AltaVista](#) - [Entire](#) - [Travel](#) - [Shopping](#) - [Careers](#) - [Health](#) - [News](#) - [Entertainment](#)

FREE INTERNET ACCESS - [Download Now](#) ^{Free} - [Support](#) **USEFUL TOOLS** - [Family Filter](#) - [Translation](#)
[Yellow Pages](#) - [People Finder](#) - [Maps](#) - [Voicemail](#) - [Check Email](#)

DIRECTORY [Automotive](#) [Business & Finance](#) [Manufacturers & Internet](#)

ALTAVISTA HIGHLIGHTS
POWER SEARCH
 100% of searches coming to AltaVista 100% job

TRY THESE SEARCHES
[Search for Halloween in photos](#)

©1999 Altavista Inc. Document Done

Θεματικοί Κατάλογοι

- Οι Θεματικοί Κατάλογοι (Directories) είναι δικτυακοί τόποι (sites) που είναι σαν ευρετήρια με κατηγορίες θεμάτων που ενδιαφέρουν το ευρύ κοινό. Κάθε κατηγορία θέματος θα μας οδηγήσει, με συνδέσμους, σε άλλους δικτυακούς τόπους, το περιεχόμενο των οποίων έχει άμεση σχέση με το συγκεκριμένο θέμα. Οι θεματικοί κατάλογοι είναι μια πολύ εύχρηστη και αποτελεσματική υπηρεσία αναζήτησης όταν ψάχνουμε για γενικές πληροφορίες για κάποιο θέμα.
- Ενδεικτικές κατηγορίες: «Ενημέρωση», «Επιστήμη», «Αθλητισμός», «Υγεία», «Εκπαίδευση».
- Η κατηγορία «Υγεία» μπορεί να έχει υποκατηγορίες: «Ιατρικές Συμβουλές», «Νοσοκομεία», «Πρώτες Βοήθειες»

Θεματικοί Κατάλογοι

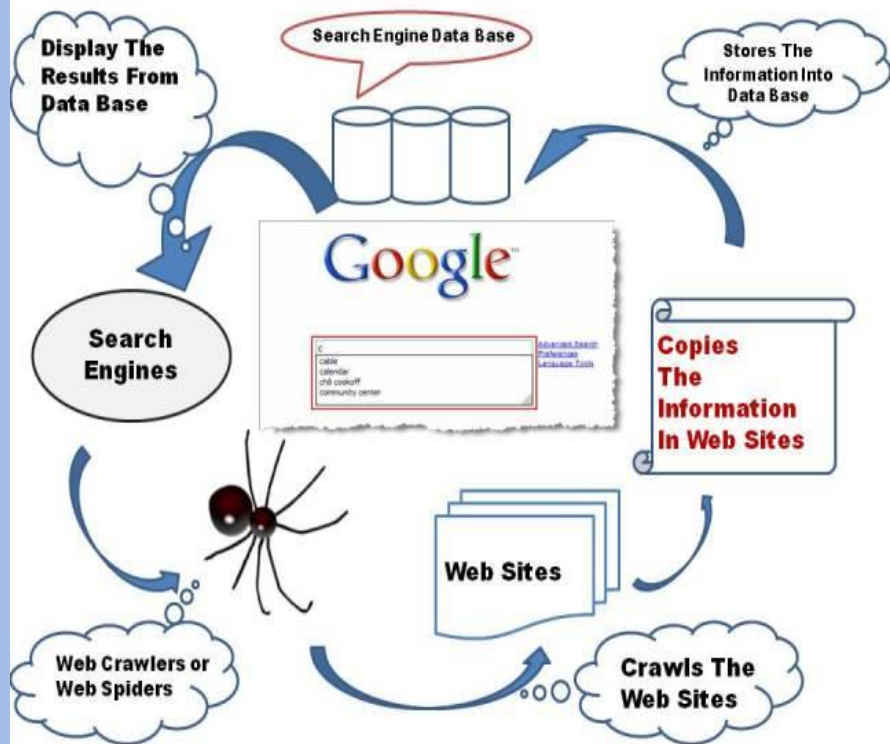
- Γενικοί
 - Best of the Web Directory – ιστοσελίδες κατηγοριοποιημένες ανα θέμα ή περιοχή (με πληρωμή)
 - [DMOZ](#) (a.k.a. [Open Directory Project](#)) – ο μεγαλύτερος θεματικός κατάλογος. Έχει ανοιχτό περιεχόμενο που το αναπαρήγαγε και η κατάλογοι της Google (μέχρι 2011).
 - [JoeAnt](#) .
 - [Sources](#) – για δημοσιογράφους, freelance συγγραφείς, εκδότες, ερευνητές.
 - [World Wide Web Virtual Library](#) (VLIB) – ο παλαιότερος θεματικός κατάλογος.
 - [S Directory](#)- startup θεματικός κατάλογος
 - [Hotfrog](#) – μεγάλη συλλογή επιχειρηματικών καταλόγων.
- Ακαδημαϊκοί
 - <https://www.academia.edu/>
 - <https://scholar.google.ca/>
 - <http://www.jurn.org/#gsc.tab=0>
 - <http://www.refseek.com/>
 - <https://www.highbeam.com/>
 - <http://link.springer.com/>
 - <https://www.base-search.net/>
- List of academic databases and search engines:
https://en.wikipedia.org/wiki/List_of_academic_databases_and_search_engines

Μηχανές Αναζήτησης

- Οι μηχανές αναζήτησης(search engines) λειτουργούν ως τεράστιες βάσεις δεδομένων οι οποίες αποθηκεύουν πληροφορίες για εκατομμύρια δικτυακούς τόπους.
- Η Αναζήτηση γίνεται με λέξεις κλειδιά.
- Η μηχανή αναζήτησης τότε θα βρει και θα ταξινομήσει τα αποτελέσματα ανάλογα με τη συνάφεια που έχουν με τις λέξεις



Λειτουργία Μηχανής Αναζήτησης



- Οι μηχανές αναζήτησης δεν ερευνούν σε πραγματικό χρόνο τον παγκόσμιο ιστό αλλά μία βάση δεδομένων που περιέχει αντίγραφα ιστοσελίδων. Οι ιστοσελίδες αυτές επιλέγονται ανάμεσα σε δισεκατομμύρια σελίδες στο ίντερνετ.

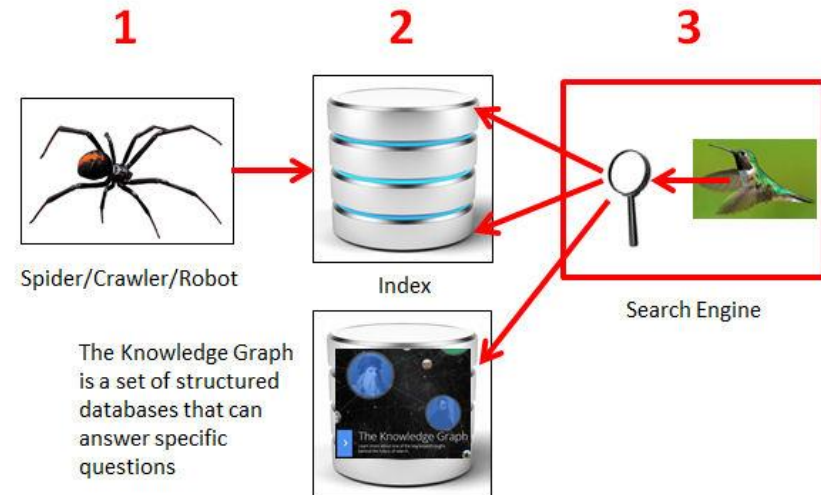
- Για την έρευνα χρησιμοποιούνται προγράμματα που ονομάζονται «ρομπότ» ή «αράχνες» (spider, crawler ή webcrawler). Τα προγράμματα «σερφάρουν» στο Διαδίκτυο σε διάφορες ιστοσελίδες και συλλέγουν πληροφορίες με βάση κριτήρια. Η λειτουργία αυτή γίνεται συνεχώς διότι οι ιστοσελίδες αλλάζουν και η βάση της μηχανής θα πρέπει να ανανεώνεται αντίστοιχα. Μια μηχανή αναζήτησης μπορεί να διαθέτει περισσότερα από ένα «ρομπότ».
- Αφού τα «ρομπότ» βρουν τις ιστοσελίδες τις περνούν σε ένα άλλο πρόγραμμα για τοποθέτηση δεικτών. Εκεί αναγνωρίζεται το κείμενο, οι σύνδεσμοι, και το υπόλοιπο περιεχόμενο της ιστοσελίδας και αποθηκεύεται στη βάση δεδομένων. Τότε μπορεί να πραγματοποιηθεί αναζήτηση πάνω στη βάση της μηχανής αναζήτησης.
- Οι μηχανές επιτρέπουν στους χρήστες να «ψάχνουν» στη βάση δεδομένων τους μέσα από περιβάλλον που παρέχει πολλαπλές δυνατότητες αναζήτησης. Η λειτουργία αυτή έχει σχέση με αυτό που αντιμετωπίζουν οι χρήστες μέσα από το περιβάλλον της μηχανής αναζήτησης.

Μηχανή Αναζήτησης – Τρία Μέρη

Οι μηχανές αναζήτησης αποτελούνται από τρία μέρη:

- Τον **spider**: Ένα πρόγραμμα το οποίο διατρέχει όλο το web και βρίσκει και διαβάζει όλες τις ιστοσελίδες.
- Το **ευρετήριο (Index)**: Είναι μία βάση δεδομένων η οποία περιλαμβάνει αντίγραφα των σελίδων που επισκέφτηκε ο spider.
- Τον **μηχανισμό αναζήτησης (search engine)**: Ένα λογισμικό που επιτρέπει στους χρήστες να κάνουν έρευνα στο ευρετήριο.

Four Parts – With The Knowledge Graph



Βασικές Συμβουλές Αναζήτησης

- Περιγράφουμε με ακρίβεια αυτό που θέλουμε να βρούμε, με βάση «Λέξεις – Κλειδιά»
- Όσες περισσότερες και εύστοχες λέξεις-κλειδιά χρησιμοποιούμε τόσο πιο σχετικά θα είναι τα αποτελέσματα που θα πάρουμε.
- Όταν χρησιμοποιούμε λιγότερες «Λέξεις – Κλειδιά» αναμένουμε μεγαλύτερο όγκο πληροφοριών
- Επιλογή όρων Αναζήτησης-Παράδειγμα
- Η επιλογή των όρων αναζήτησης είναι το κλειδί ...
 - Έστω ότι ψάχνουμε για την Πύλη Βηθλεέμ του Ηρακλείου...
 - Ψάχνουμε με τον όρο Ηράκλειο: Πολλά – γενικά αποτελέσματα.
 - Ψάχνουμε με τους όρους Ηράκλειο-Ιστορία: Καλύτερα αποτελέσματα
 - Κάνουμε την αναζήτηση πιο συγκεκριμένη: Ηράκλειο-Ιστορία-Πύλη Βηθλεέμ
- Κεφαλαία – Πεζά - Τονισμός
 - Οι αναζητήσεις στο Google ΔΕΝ επηρεάζονται από το αν θα γράψετε μια λέξη:
 - με κεφαλαία ή πεζά γράμματα
 - με τόνο ή χωρίς
 - Η σειρά με την οποία πληκτρολογούνται οι όροι επηρεάζει τα αποτελέσματα της αναζήτησης

Σύνθετη Αναζήτηση...

- AND
 - Το Google συνδέει αυτόματα όλους τους όρους αναζήτησης με τον λογικό τελεστή AND.
 - Στα αποτελέσματα εμφανίζονται ιστοσελίδες που περιέχουν όλους τους όρους...
- OR
 - π.χ Ηράκλειο OR Ιστορία
 - Τα αποτελέσματα περιέχουν σελίδες που περιέχουν τον ένα από τους δύο όρους...
- NOT
 - Προσθέτουμε ένα μείον (-) πριν την λέξη που θέλουμε να μην αναζητηθεί...
 - π.χ: Ελληνικά Μάρμαρα -Θάσος

Σύνθετη Αναζήτηση...

Google Σύνθετη Αναζήτηση [Συμβουλές αναζήτησης](#) | [Σ](#)

Εύρεση αποτελεσμάτων	με όλες τις λέξεις με την ακριβή φράση με τουλάχιστον μία λέξη χωρίς τις λέξεις	<input type="button" value="AND"/> <input type="text"/> <input type="button" value="OR"/> <input type="button" value="NOT"/>	10 αποτελέσματα ▾ <input type="button" value="Αναζήτηση Google"/>
Γλώσσα	Εύρεση σελίδων σε	<input type="text" value="οποιαδήποτε γλώσσα"/> ▾	
Περιοχή	Αναζήτηση σελίδων που βρίσκονται σε:	<input type="text" value="οποιαδήποτε περιοχή"/> ▾	
Μορφή αρχείου	<input type="button" value="Μόνο"/> ▾ εύρεση αποτελεσμάτων σε αρχείο	<input type="text" value="οποιαδήποτε μορφή"/> ▾	
Ημερομηνία	Επιστροφή σε ιστοσελίδες που εμφανίστηκαν για πρώτη φορά	<input type="text" value="οποτεδήποτε"/> ▾	
Εμφανίσεις	Εύρεση αποτελεσμάτων στα οποία εμφανίζονται οι όροι μου	<input type="text" value="οπουδήποτε στη σελίδα"/> ▾	
Τομέας δικτύου	<input type="button" value="Μόνο"/> ▾ επιστροφή αποτελεσμάτων από τον ιστότοπο ή το δικτυακό τομέα	<input type="text"/>	
Δικαιώματα χρήσης	Εύρεση αποτελεσμάτων που είναι	<input type="text"/>	

π.χ. google.com, στο [Περισσότερες πληροφορίες](#)

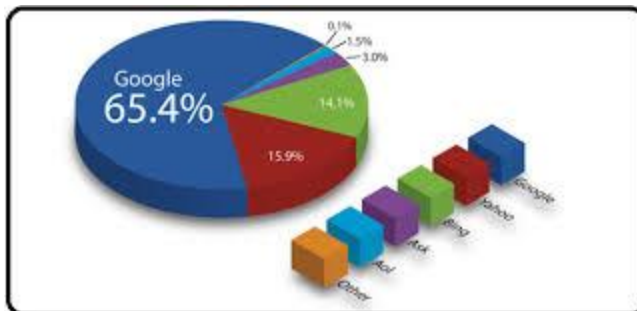
Αναζήτηση με βάση τη σελίδα

Παρόμοιος	Εύρεση σελίδων παρόμοιων με τη σελίδα	<input type="text"/> <input type="button" value="Αναζήτηση"/>
Σύνδεσμοι	Εύρεση σελίδων με συνδέσμους στη σελίδα	<input type="text"/> <input type="button" value="Αναζήτηση"/>

π.χ. www.google.com/intl/el/help.html

To Google or not to Google?

- WolframAlpha. Η μηχανή αναζήτησης των σπουδαστών (<https://www.wolframalpha.com/>)
- DogPile. Η επιστροφή μιας πρωτόπορου μηχανής αναζήτησης (<http://www.dogpile.com/>)
- Yippy. Για να μπείτε στα άδυτα του διαδικτύου (<http://yippy.com/>)
- Bing. Για τους φανατικούς της Microsoft (<http://www.bing.com/>)
- Ask.com. Ρώτησε για να μάθεις... (<http://uk.ask.com>)
- Yandex. Η ρώσικη μηχανή αναζήτησης (<https://www.yandex.com/>)
- More
 - <https://www.tsouk.gr/15-enallaktikes-michanes-anazitisis-ektos-apo-tin-google/>
 - <http://www.kathimerini.gr/796784/article/tecnologia/diadiaktyo/polemos-mhxanwn-anazhthshs>



Πως λειτουργεί το Google?

Crawlers – Spiders (ξανά)....

- Τα προγράμματα αυτά επισκέπτονται όλες τις σελίδες στο Internet, εκτός από αυτές που ο ιδιοκτήτης ενός site έχει επιλέξει να μην συμπεριλαμβάνονται στην αναζήτηση.
- Πηγαίνοντας από link σε link, οι crawlers καταγράφουν τις λέξεις-κλειδιά και τοποθετούν τις σελίδες σε κατηγορίες, με βάση το περιεχόμενό τους και άλλους παράγοντες.
- Όλες αυτές οι πληροφορίες που συγκεντρώνουν, αποθηκεύονται στο "Index", ένα "κατάλογο περιεχομένων" για το Internet, που έχει μέγεθος πάνω από 100 petabytes (**100 εκατομμύρια gigabytes**).
- Αυτή η μέθοδος είναι που κάνει την αναζήτηση στο Internet τόσο γρήγορη. Όταν ψάχνουμε κάτι στο Google, δεν ψάχνει εκείνη την ώρα σε ολόκληρο το Internet, αλλά κάνει μια αναζήτηση στο πολύ μικρότερο Index.

Πως λειτουργεί το Google?

Που αποθηκεύεται αυτό το Index;

- Η πολιτική του Google είναι να χρησιμοποιεί εκατοντάδες χιλιάδες μικρούς και σχετικά οικονομικούς **servers** σε ολόκληρο τον κόσμο για όλες τις λειτουργίες.
- Οι server αυτοί τρέχουν ένα ειδικά τροποποιημένο λειτουργικό σύστημα βασισμένο στο Linux. Ένα πρόγραμμα που ονομάζεται Google File System αναλαμβάνει τη διαχείριση των δεδομένων σε όλους τους διαφορετικούς servers.



ΤΕΙ ΙΟΝΙΩΝ
ΝΗΣΩΝ

Πως λειτουργεί το Google?

Hardware – Servers

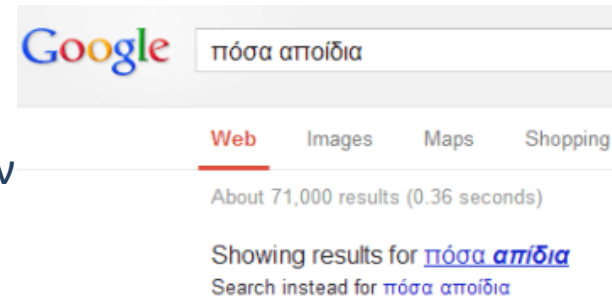
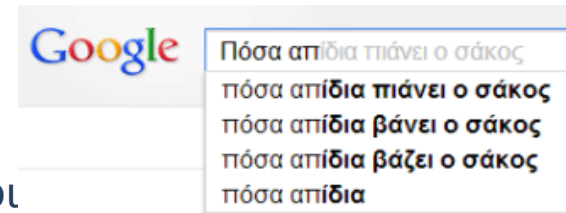
- Οι **Web Server** δέχονται και επεξεργάζονται τις ερωτήσεις των χρηστών, στέλνοντας την ερώτηση στον κατάλληλο server
- Οι **Index servers** αποθηκεύουν συνολικά το Index και τα αποτελέσματα των αναζητήσεων.
- Οι **Document servers** αποθηκεύουν τις πληροφορίες των χρηστών, τα email του Gmail, τα αρχεία του Google Drive κλπ.
- Οι **Ad servers** αποθηκεύουν τις διαφημίσεις του Google που προβάλλονται οπουδήποτε στον κόσμο.



Πως λειτουργεί το Google?

Software – Αλγόριθμοι

- Όλα τα εκατομμύρια Gigabytes του Index θα ήταν εντελώς άχρηστα χωρίς τους κατάλληλους αλγόριθμους – διαφορετικά προγράμματα που βοηθούν την αναζήτηση.
- Όταν κάνουμε μία αναζήτηση, μία σειρά από αλγόριθμοι μπαίνουν σε λειτουργία.
- **Autocomplete**: αλγόριθμος που προσπαθεί να μαντέψει τι θέλουμε να ψάξουμε, για να μας γλιτώσει χρόνο.
- Αλγόριθμος ελέγχει την ορθογραφία και μας κάνει τις σωστές προτάσεις
- Αλγόριθμοι εντοπίζουν συνώνυμα των λέξεων που ψάχνουμε, καθώς επίσης προσπαθούν να αντιληφθούν τι ακριβώς θέλουμε να βρούμε, με τις συγκεκριμένες λέξεις.
- Στη συνέχεια, έχοντας μια γενικότερη κατανόηση του τι ψάχνουμε, γίνεται μια αναζήτηση στο Index (από τον αντίστοιχο αλγόριθμο αναζήτησης).
- Στην πράξη κάθε αναζήτηση και επιστροφή αποτελεσμάτων γίνεται σε κλάσματα του δευτερολέπτου.



Πως λειτουργεί το Google?

PageRank

- Αυτό που κάνει το Google να ξεχωρίζει από άλλες μηχανές αναζήτησης είναι ο τρόπος με τον οποίο επιστρέφει τα σχετικότερα αποτελέσματα. Γι' αυτό σε κάθε αναζήτηση, συνήθως βρίσκουμε αυτό που ψάχναμε μέσα στην πρώτη ή το πολύ τη δεύτερη σελίδα των αποτελεσμάτων του Google (πολλοί από εμάς δεν ψάχνουμε καν πέρα από την πρώτη).
- Ο PageRank να δίνει μια βαθμολογία σε κάθε σελίδα για κάθε λέξη-κλειδί στην αναζήτηση ελέγχοντας μεταξύ άλλων, τις παρακάτω παραμέτρους:
 - Την συχνότητα και την τοποθεσία των λέξεων κλειδιών στη σελίδα: Αν μια λέξη-κλειδί εμφανίζεται μόνο μία φορά στο κείμενο της σελίδας, η σελίδα θα έχει χαμηλή βαθμολογία όσον αφορά τη συγκεκριμένη λέξη.
 - Πόσο καιρό υπάρχει η σελίδα και πότε έχει ανανεωθεί τελευταία φορά: Με χιλιάδες σελίδες να δημιουργούνται κάθε μέρα, και πολλές από αυτές να εξαφανίζονται μετά από κάποιο διάστημα, το Google θεωρεί πιο αξιόπιστες αυτές που έχουν περισσότερα χρόνια στην πλάτη τους.
 - Το πόσες σελίδες έχουν link προς μία συγκεκριμένη σελίδα: Το Google ελέγχει πόσες διαφορετικές ιστοσελίδες έχουν link προς μία συγκεκριμένη σελίδα, σε συνδυασμό με το Rank εκείνων των σελίδων, προκειμένου να υπολογίσουν το κατά πόσον είναι σχετική με την αναζήτησή μας.
- Από τους τρεις παραπάνω παράγοντες, ο τρίτος είναι ο πιο σημαντικός όλων: όσο πιο πολλά διαφορετικά sites έχουν σύνδεσμο προς μία συγκεκριμένη σελίδα του site μας, τόσο μεγαλύτερο θα είναι το δικό μας PageRank για τη συγκεκριμένη σελίδα, και τόσο πιο ψηλά θα εμφανίζεται στα αποτελέσματα του Google σε μια σχετική αναζήτηση.



$$PR(A) = (1-d) + \frac{d}{C(T_1)} (PR(T_1) + \dots + PR(T_n))$$