



Ανάκτηση πληροφορίας

Ενότητα 1: Εισαγωγή στην Ανάκτηση Πληροφορίας

Απόστολος Παπαδόπουλος
Τμήμα Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.





ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

ΑΝΟΙΚΤΑ
ΑΚΑΔΗΜΑΪΚΑ
ΜΑΘΗΜΑΤΑ



Εισαγωγή



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Περιεχόμενα ενότητας

1. Εισαγωγή





ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

Εισαγωγή

Αντικείμενο IR

- Η Ανάκτηση Πληροφορίας (ΑΠ) (Information Retrieval - IR) μελετά προβλήματα που σχετίζονται με την:
 - Αναπαράσταση
 - Αποθήκευση
 - Οργάνωση
 - Προσπέλαση

στοιχείων πληροφορίας (έγγραφα κειμένου, εικόνες, μουσικά αρχεία, βίντεο).



ΑΠ και ΒΔ 1

- Οι περιοχή της Ανάκτησης Πληροφορίας (ΑΠ) έχει πολλά κοινά με την περιοχή των Βάσεων Δεδομένων (ΒΔ).
- Οι δύο περιοχές για πολλά χρόνια εξελίσσονταν παράλληλα και έτσι η κάθε μία έχει αναπτύξει διαφορετικές τεχνικές.
- Οι διαφορές οφείλονται στις διαφορές που έχουν τα δεδομένα και τα ερωτήματα.
- Οι ερευνητές κατευθύνονται προς τη μελέτη συστημάτων που θα έχουν δυνατότητες ΣΔΒΔ και θα μπορούν να εκτελούν αποτελεσματικά και αποδοτικά τις λειτουργίες των συστημάτων ΑΠ.



ΣΑΠ και ΣΔΒΔ

Χαρακτηριστικό	ΣΔΒΔ	ΣΑΠ
είδος δεδομένων	απολύτως δομημένα	αδόμητα, ημι-δομημένα
τύπος δεδομένων	αριθμητικά, αλφαριθμητικά	έγγραφα (κειμένου)
γλώσσα ερωτημάτων	SQL, QBE	φυσική γλώσσα
ερώτημα	σαφές	ασαφές
αποτελέσματα	χωρίς βαθμολόγηση	βαθμολογημένα



ΑΠ και ΒΔ 2

Ερώτημα SQL

```
SELECT όνομα  
FROM πλανήτες  
WHERE δορυφόροι = 0  
      OR δορυφόροι = 1  
      OR δορυφόροι = 2
```

κωδικός	όνομα	διάμετρος	δορυφόροι
1	Ερμής	4880	0
2	Αφροδίτη	12103.6	0
3	Γη	12756.3	1
4	Άρης	6794	2
5	Δίας	142984	63
6	Κρόνος	120536	34
7	Ουρανός	51118	21
8	Ποσειδώνας	49532	13
9	Πλούτωνας	2274	3

Το ερώτημα είναι σαφές, προσδιορίζει επακριβώς τη συνθήκη που πρέπει να ικανοποιεί κάθε record που εμφανίζεται στην έξοδο.



ΑΠ και ΒΔ 3

Συλλογή εγγράφων

- d1 : Ο κομήτης του Χάλλεϋ μας επισκέπτεται περίπου κάθε εβδομήντα έξι χρόνια.
- d2 : Ο κομήτης του Χάλλεϋ πήρε το όνομά του από τον αστρονόμο Έντμοντ Χάλεϋ.
- d3 : Ένας κομήτης διαγράφει ελλειπτική τροχιά.
- d4 : Ο πλανήτης Άρης έχει δύο φυσικούς δορυφόρους, το Δείμο και το Φόβο.
- d5 : Ο πλανήτης Δίας έχει 63 γνωστούς φυσικούς δορυφόρους.
- d6 : Ένας κομήτης έχει μικρότερη διάμετρο από ότι ένας πλανήτης.
- d7 : Ο Άρης είναι ένας πλανήτης του ηλιακού μας συστήματος.

Πληροφοριακή ανάγκη (ερώτημα): πληροφορίες για τον κομήτη του Χάλλεϋ
Διαισθητικά αντιλαμβανόμαστε ότι τα έγγραφα d1 και d2 σχετίζονται περισσότερο με το ερώτημα απ' ότι τα υπόλοιπα έγγραφα.



Αναζήτηση και Περιήγηση-1

- **Αναζήτηση (Search)**

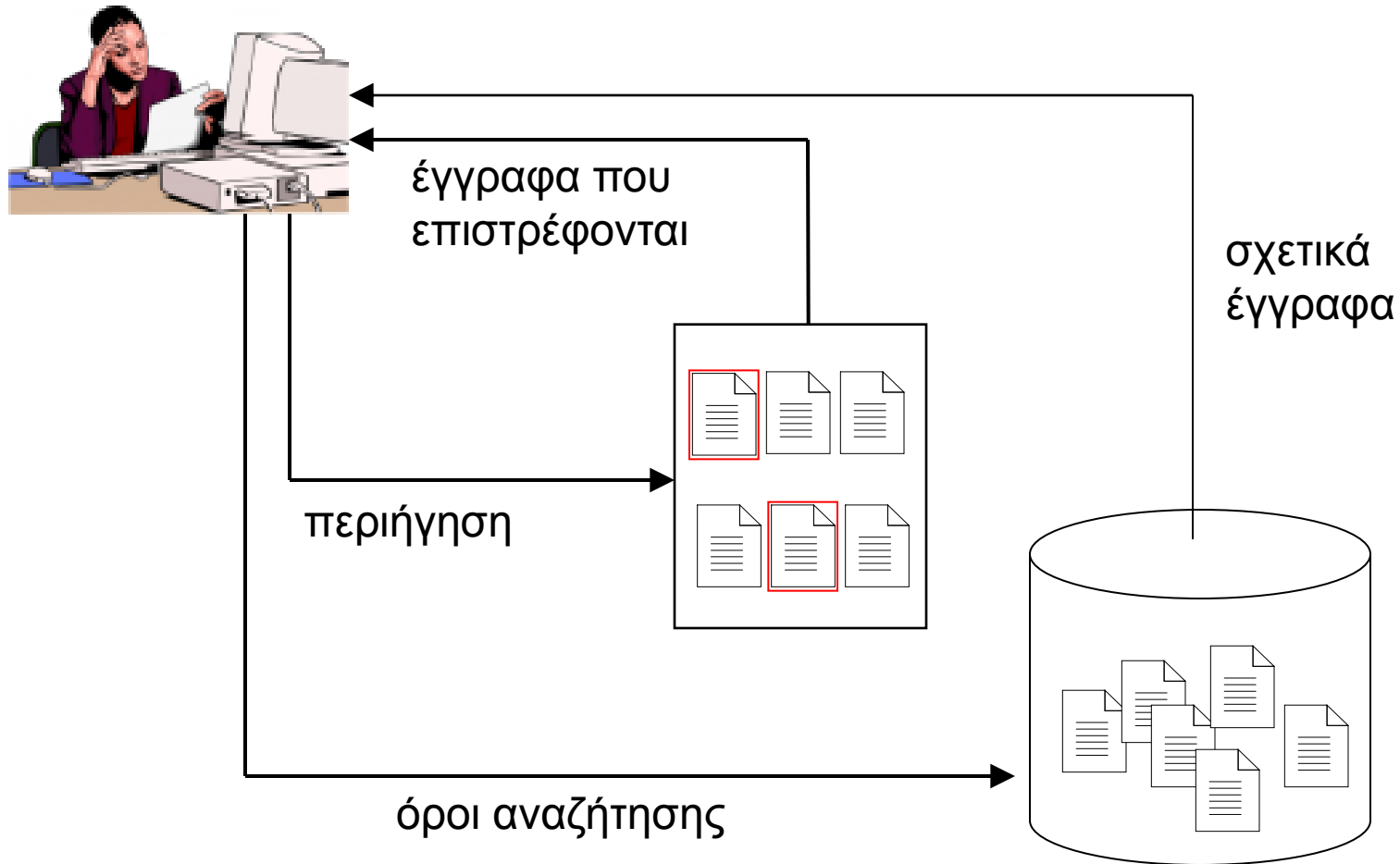
- Η δυνατότητα εύρεσης πληροφορίας σχετική με την ερώτηση του χρήστη. Ο χρήστης εκφράζει την πληροφοριακή ανάγκη (ερώτημα) και το σύστημα επιστρέφει έγγραφα με μεγάλο βαθμό ομοιότητας ως προς το ερώτημα.

- **Περιήγηση (Browsing)**

- Η δυνατότητα μετακίνησης μεταξύ των αποτελεσμάτων ή όλων των εγγράφων της συλλογής (π.χ. Hypertext).



Αναζήτηση και Περιήγηση-2



Πηγή φωτογραφίας: <http://www.picturesof.net/pages/100115-042047-964042.html>



Αναζήτηση Εγγράφων

- Η ανάκτηση εγγράφων τα οποία **σχετίζονται** με την ερώτηση του χρήστη, και η αποφυγή ανάκτησης εγγράφων που **δε σχετίζονται** με την ερώτηση του χρήστη.
- Τα τελικά αποτελέσματα ταξινομούνται ανάλογα με το ποσοστό συσχέτισης (relevance) δηλαδή το βαθμό ομοιότητας.



Παραδείγματα εφαρμογών-1

The screenshot shows the Amazon.com website interface. The search bar contains the text "information retrieval". The results page displays a list of books under the heading "information retrieval".

Category
Any Category
Books (125,424)
Kindle Store (343)
Magazine Subscriptions (2)
Office Products (1)
MP3 Downloads (1)

Listmania!

The Semantic Web: A list by calvinme "Texan refugee"

Building Effective Web Sites: A list by William Slawski

Create a Listmanialist

Search Listmania! GO!

ADVERTISMENT

WHY DAY

Ολοκληρώθηκε

"information retrieval"

Showing Top Results < Previous | Page: 1 2 3 | Next >

- Information Retrieval: Algorithms and Heuristics (The Information Retrieval Series)(2nd Edition)** by David A. Grossman and Ophir Frieder (**Paperback** - Dec 20, 2004)
Buy new: \$38.00 \$31.54 32 Used & new from \$28.99
Get it by **Friday, Feb 8** if you order in the next **13 hours** and choose one-day shipping.
Eligible for **FREE** Super Saver Shipping.
★★★★☆ (4)
Excerpt - page 2: "... **INFORMATION RETRIEVAL:ALGORITHMS AND HEURISTICS** Figure 1.1. Document ..."
Surprise me! See a random page in this book.
Books: See all 125,424 items
- Text Information Retrieval Systems, Third Edition (Library and Information Science) (Library and Information Science)** by Charles T. Meadow, Bert R. Boyce, Donald H. Kraft, and Carol L Barry (**Hardcover** - Mar 26, 2007)
Buy new: \$89.95 \$73.00 19 Used & new from \$68.00
Get it by **Friday, Feb 8** if you order in the next **13 hours** and choose one-day shipping.
Eligible for **FREE** Super Saver Shipping.
★★★★☆ (1)
Books: See all 125,424 items
- Modern Information Retrieval** by Ricardo Baeza-Yates and Berthier Ribeiro-Neto (**Paperback** - May 15, 1999)
15 Used & new from \$75.99
★★★★☆ (9)
Books: See all 125,424 items
- Introduction to Modern Information Retrieval** by G. G. Chowdhury (**Paperback** - Dec 1, 2003)
Buy new: \$89.95 \$79.16 6 Used & new from \$68.00
Get it by **Friday, Feb 8** if you order in the next **13 hours** and choose one-day shipping.
Eligible for **FREE** Super Saver Shipping.
Books: See all 125,424 items
- Information Retrieval: Data Structures and Algorithms** by William B. Frakes and Ricardo Baeza-Yates (**Paperback** - Jun 22, 1992) - Facsimile



Παραδείγματα εφαρμογών-2

Library of Congress Online Catalogs - Mozilla Firefox

http://catalog.loc.gov/

The Library of Congress >> Switch to Library of Congress Authorities

LIBRARY OF CONGRESS ONLINE CATALOG

About Displaying and Searching Using Non-Roman Characters
Frequently Asked Questions - Help Table of Contents - What's New

Basic Search

Using a fill-in box, search by:

- Title or Author/Creator
- Subject
- Call number
- LCCN, ISSN, or ISBN
- Keywords

Note: Search limits are available only for title and keyword searches.

Guided Search

Using a series of forms and menus:

- Construct keyword searches
- Restrict all or part of the search to a particular index
- Combine search words or phrases with Boolean operators

Note: Search limits are available for all searches.

Other Online Catalogs:

- Prints and Photographs Online Catalog (PPOC)
[About - Start Searching](#)
- Sound Online Inventory & Catalog (SONIC)
[About - Start Searching](#)
- Alternative Interface to the LC Online Catalog (Z39.50)
[About - Start Searching](#)
- [Copyright Records: Registrations and Documents](#)
- [Other Libraries' Catalogs](#)

Information about the images: Two pendentive paintings by Edward J. Holslag are displayed from the Librarian's Room (Librarian's Ceremonial Office) located in the Thomas Jefferson Building of the Library of Congress. On the left, "Efficiunt clarum studio" (Study, the watchword of fame); on the right, "Dulce ante omnia musae" (The Muses, above all things, delightful).

The Library of Congress
July 25, 2007

[Contact Us](#)

Παραδείγματα εφαρμογών-3

The screenshot shows the Entrez PubMed website in a browser window. The address bar displays <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>. The page features the NCBI logo on the left, the PubMed logo in the center, and the National Library of Medicine (NLM) logo on the right. A navigation bar includes links for All Databases, PubMed, Nucleotide, Protein, Genome, Structure, OMIM, PMC, Journals, and Books. Below this is a search bar with a dropdown menu set to 'PubMed' and a search button. A sidebar on the left contains links for 'About Entrez', 'Text Version', 'Entrez PubMed' (Overview, Help | FAQ, Tutorial, New/Noteworthy, E-Utilities), and 'PubMed Services' (Journals Database, MeSH Database, Single Citation Matcher, Batch Citation Matcher, Clinical Queries, Special Queries, LinkOut, My NCBI (Cubby)). The main content area includes instructions for searching: 'To get started, enter one or more search terms. Search terms may be [topics](#), [authors](#) or [journals](#).' Below this is a highlighted box for 'My NCBI' with the text 'Set up an automated PubMed update in less than 5 minutes.' and a list of steps: '(1) Get a [My NCBI account](#). (2) Save your search. (3) Your PubMed updates can be e-mailed directly to you.' Further down, it states 'PubMed is a service of the [National Library of Medicine](#) that includes over 15 million citations from MEDLINE and other life science journals for biomedical articles back to the 1950s. PubMed includes links to full text articles and other related resources.' At the bottom, there is a banner for 'New Global NCBI Search Engine'. The browser's status bar at the bottom left shows 'Done'.

Παραδείγματα εφαρμογών-4

Yahoo! - Mozilla Firefox
http://www.yahoo.com/

Get Y! Toolbar | Make Y! your home page

Web Images Video Local Shopping more

Web Search

Yahoo! Home My Yahoo! Feb 7, 2008 Page Options

Answers Autos Finance Games Groups HotJobs Maps Mobile Web Movies Music Personals Real Estate Shopping Sports Tech TV Yellow Pages More Yahoo! Services

Small Business
Get a Web Site
Domain Names
Sell Online
Search Ads

Featured Services
Downloads
Health
Kids
GeoCities
Mail Plus
Y! International

Featured
Entertainment Sports Video

Marriage: Worse over time?
Couples find each other more irritating the longer they're together, a study shows. >> Silver lining
Find secrets to a long marriage
Signs that you've found the right mate

In the News
World Local Finance

As of 11:50 a.m. GST
Taliban militants declare cease-fire with Pakistan forces
Israeli troops launch raid in Gaza, killing five Hamas gunmen
Congress extends program to prohibit calls from telemarketers
One of the last surviving U.S. veterans of WWII dies at 108
N.C. primary could be key to Democratic nomination '08 race
Bad weather threatens to delay space shuttle launch
HBO cancels 'Inside the NFL' after more than 30 years on the air
NBA · NCAA Hoops · NHL · NASCAR · NFL · Soccer · MLB
>> More: News | Popular | Election '08

Markets: Dow: -0.5% Nasdaq: -1.3% Sponsored by: Scottrade

Marketplace
Expand your business globally
Buy and sell products on the worlds leading B2B online marketplace. Set up a free account now.
Go to Yahoo! Shopping and compare prices on thousands of phones. Find your new cells phone now.
Get a quote on a new car in your area with Yahoo! Autos.
Bring a little romance into your life. Play Dream Day Honeymoon and Ciao Bella at Yahoo! Games.

Check your mail status: Sign In Free mail: Sign Up
Mail Messenger Radio
Weather Local Horoscopes

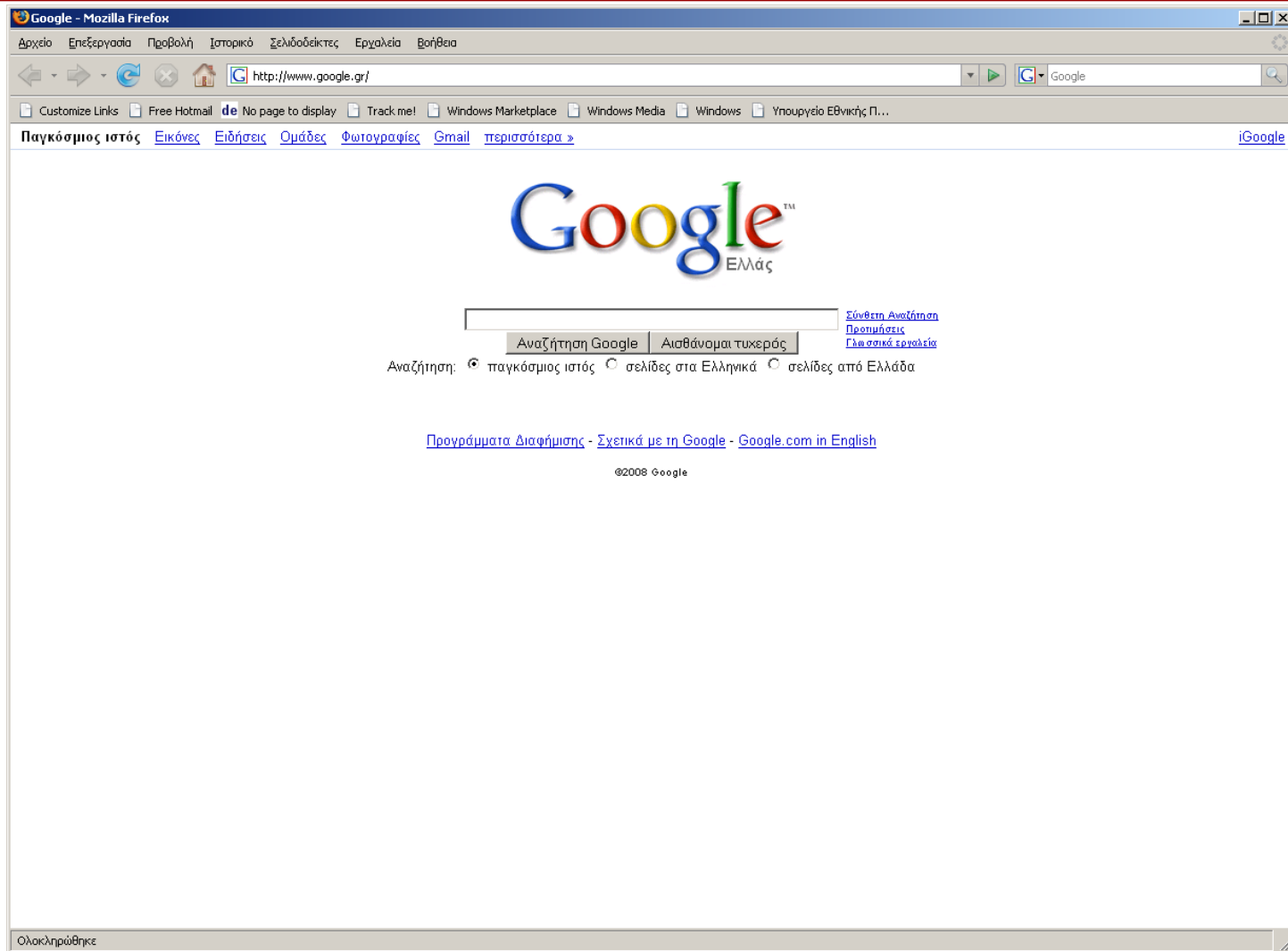
Yahoo! Shopping
Find Hot Deals, Coupons & Buying Guides > Shop now
Computers Game Consoles Bedding HDTVs
Toys Shoes Fitness GPS

Yahoo! Movies: In Theaters
Discover great new films
Rambo The Air I Breathe Over Her Dead Body Cloverfield

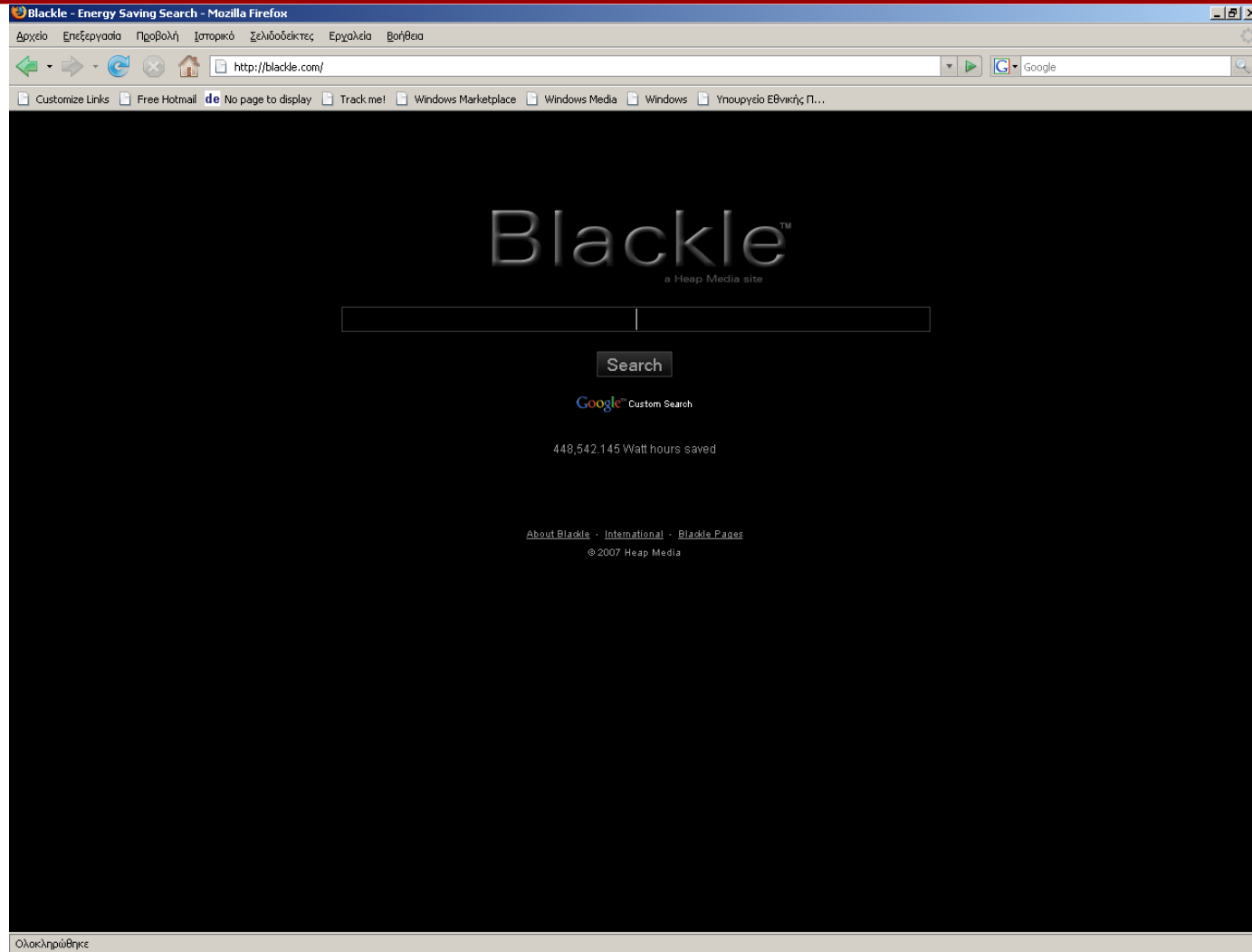
Pulse - What Yahoos Are Into
Popular Candid Celebrity Photos
All Eyes on Jessica
Eva's L.A. Premiere
Wild Wild West
'Young' Love
Britney's Breakdown
Paris Takes Flight
>> More omg!



Παραδείγματα εφαρμογών-5



Παραδείγματα εφαρμογών-6



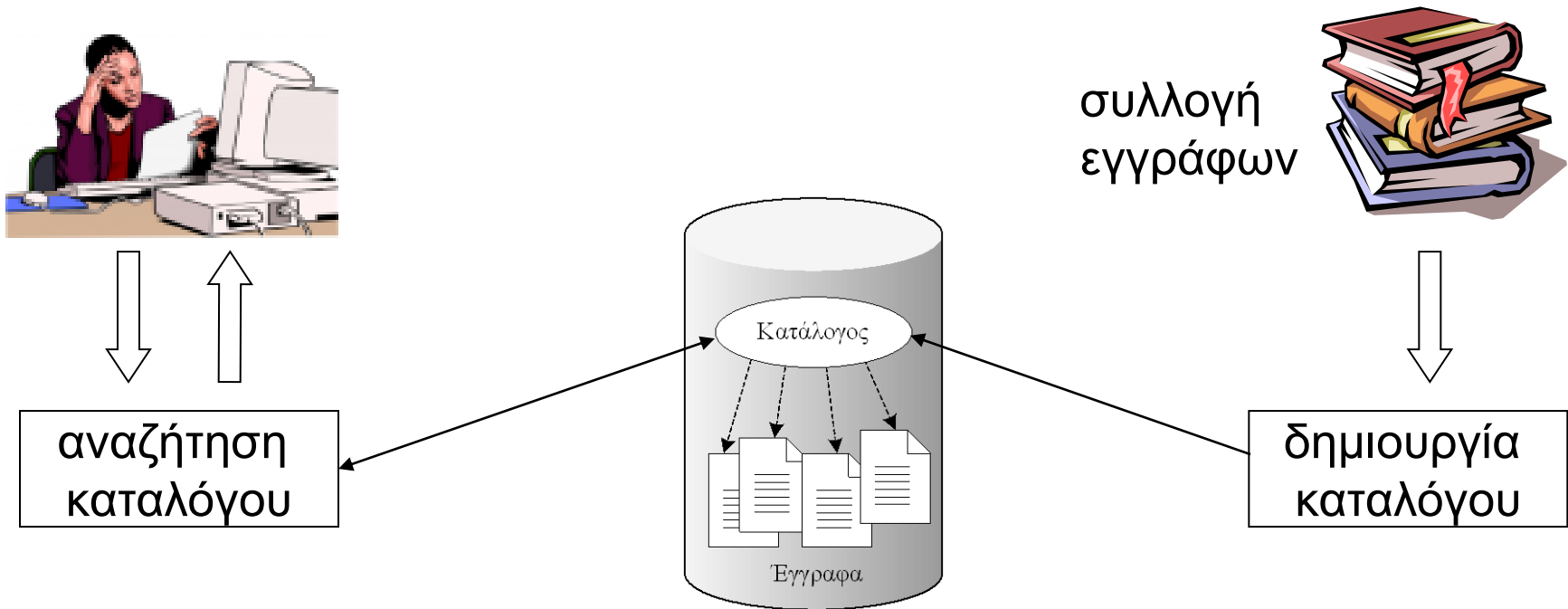
Ταξινόμηση και Βαθμολόγηση

- Όταν ένας χρήστης υποβάλλει ένα ερώτημα σε ένα σύστημα ανάκτησης, το σύστημα επιστρέφει ένα σύνολο αποτελεσμάτων. Εάν η συλλογή κειμένων είναι μεγάλη, ενδέχεται και τα αποτελέσματα να είναι πολλά (π.χ. αναζήτηση σε μία μηχανή αναζήτησης). Επομένως, έχει μεγάλη σημασία η **σειρά εμφάνισης** των αποτελεσμάτων. Υπάρχουν τρεις μέθοδοι:
 - **Ταξινόμηση** των αποτελεσμάτων π.χ. ως προς την ημερομηνία.
 - **Βαθμολόγηση** των αποτελεσμάτων ως προς την ομοιότητα με το ερώτημα.
 - **Βαθμολόγηση** των αποτελεσμάτων ως προς τη σημαντικότητά τους.



Χρήση Καταλόγων

- Τα συστήματα ανάκτησης σπάνια αναζητούν την πληροφορία απευθείας στη συλλογή εγγράφων. Συνήθως, χρησιμοποιούνται **κατάλογοι** οι οποίοι **επιταχύνουν** τη διαδικασία αναζήτησης.

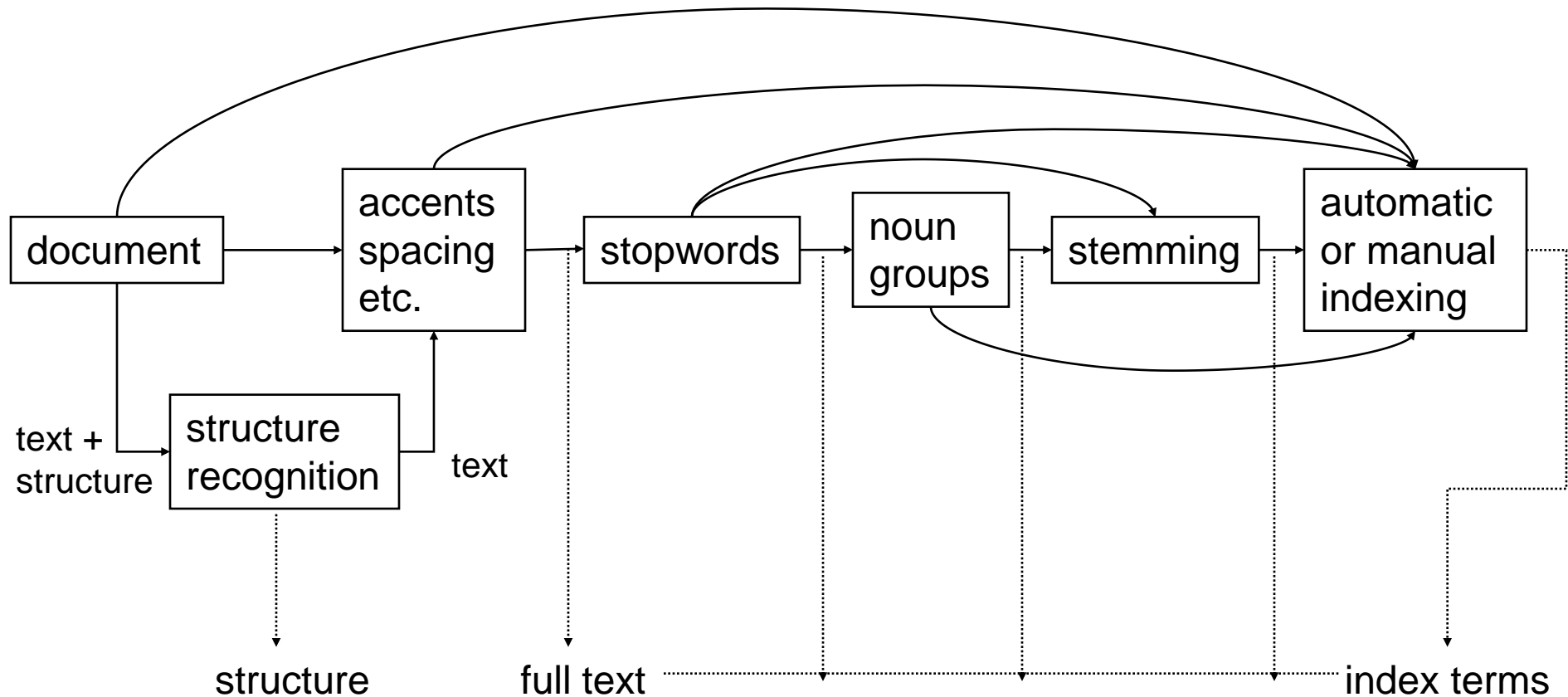


Αναπαράσταση Εγγράφων-1

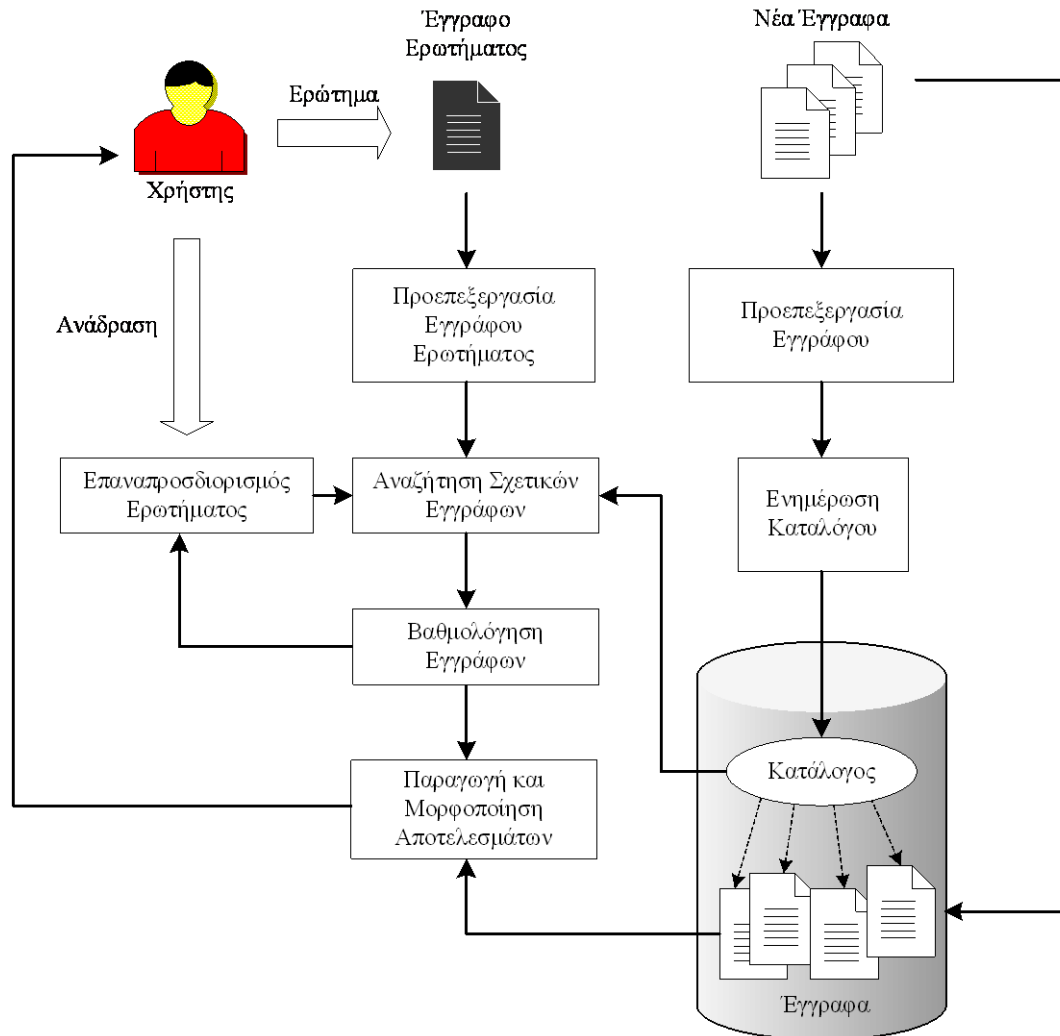
1. Πλήρες κείμενο.
 2. Απαλοιφή άρθρων και συνδέσμων.
 3. Απαλοιφή επιθέτων, ρημάτων, κλπ.
 4. Άλλοι μετασχηματισμοί.
 5. Χαρακτηριστικές Λέξεις (Keywords).
- Πολλές φορές μας ενδιαφέρει και η δομή (structure) του κειμένου (π.χ. αναπαράσταση με XML).



Αναπαράσταση Εγγράφων-2



Δομή ενός ΣΑΠ

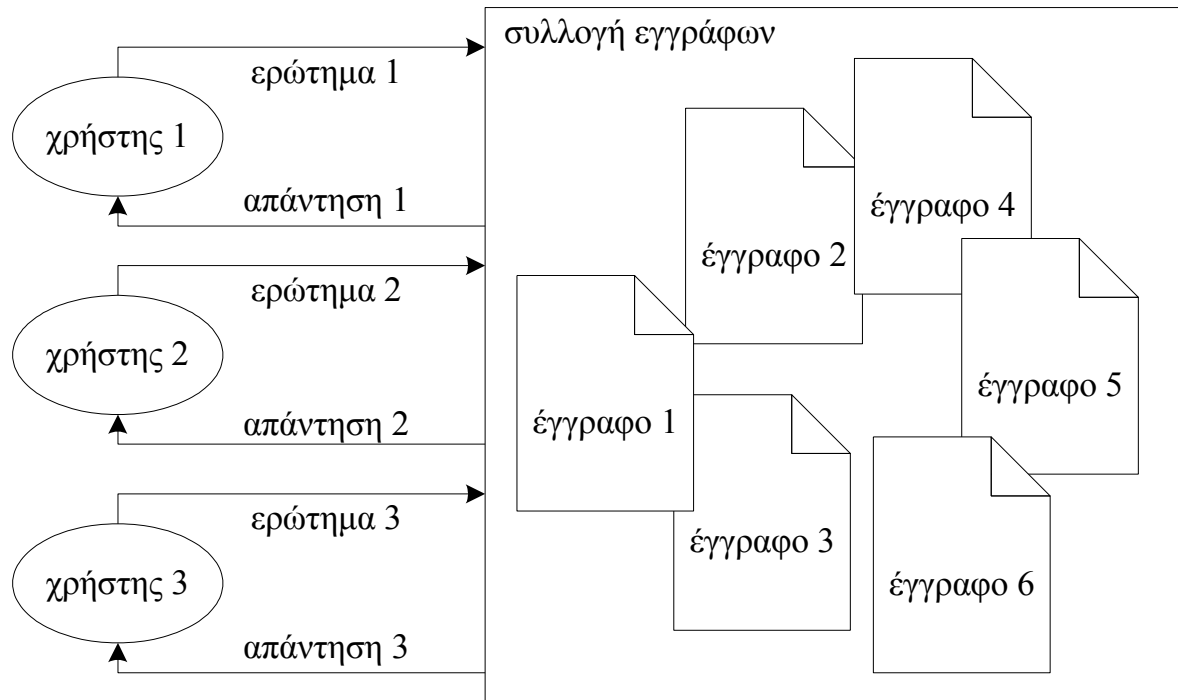


Περιστασιακή Αναζήτηση και Φιλτράρισμα

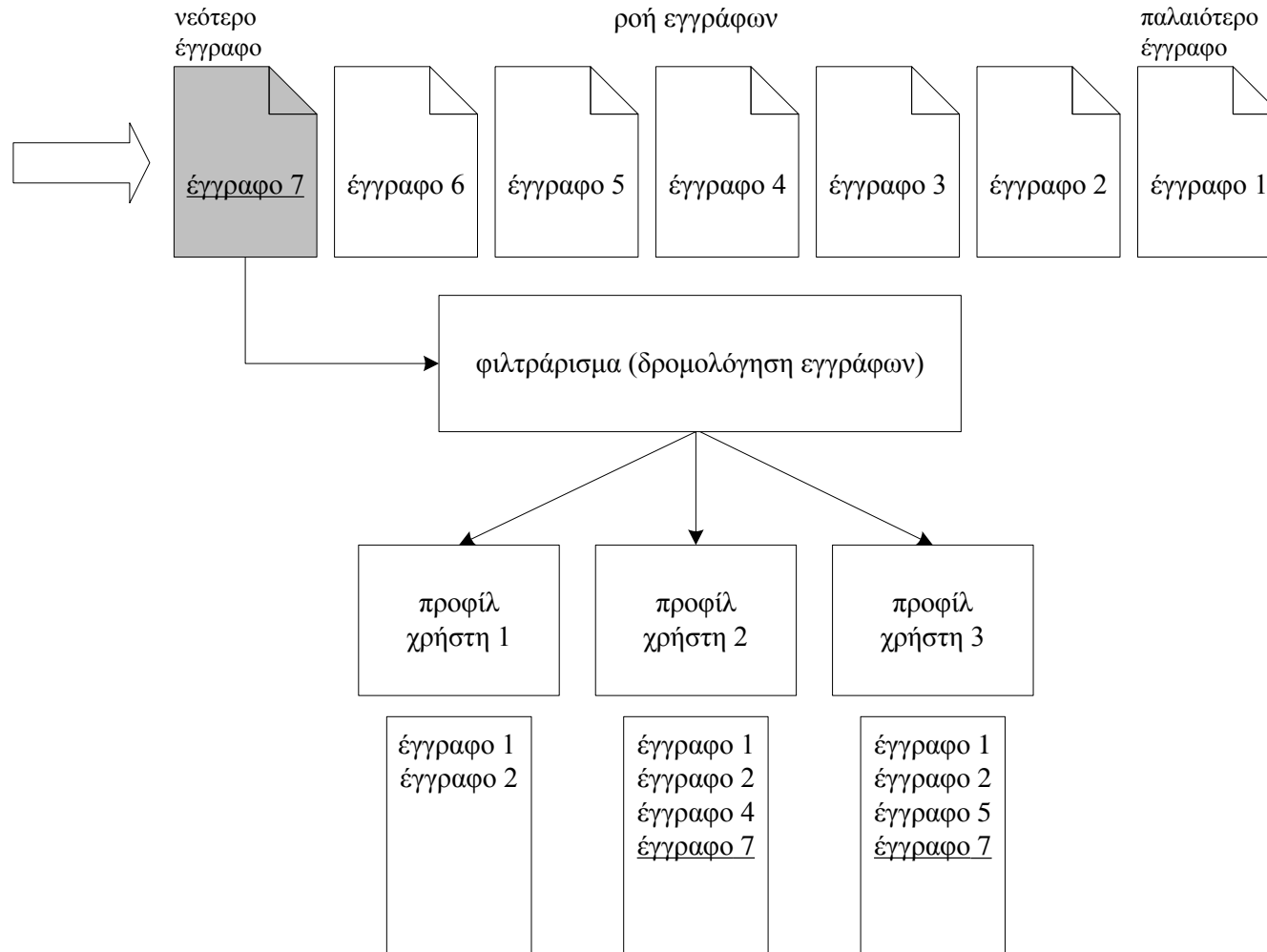
- Ένα ερώτημα μπορεί να ανήκει σε μία από δύο κατηγορίες. Η πρώτη κατηγορία αφορά σε ερωτήματα που εκτελούνται μία φορά, ενώ η δεύτερη κατηγορία αφορά σε ερωτήματα που εκτελούνται συνεχώς για κάποιο χρονικό διάστημα που συνήθως προσδιορίζεται από το χρήστη.
- Η πρώτη κατηγορία αναζήτησης καλείται **περιστασιακή** ή **εξειδικευμένη** (ad-hoc), ενώ η συνεχής επεξεργασία ενός ερωτήματος καλείται και **φιλτράρισμα** (filtering) ή **δρομολόγηση** (routing) των εγγράφων.



Περιστασιακή Αναζήτηση



Φιλτράρισμα ή Δρομολόγηση



Χθες και Σήμερα

- Τα πρώτα διαθέσιμα συστήματα IR χρησιμοποιήθηκαν από βιβλιοθήκες για αναζήτηση βιβλίων με βάση τον τίτλο, το συγγραφέα, τον εκδοτικό οίκο, κλπ.
- Αργότερα, προστέθηκε η δυνατότητα να πραγματοποιείται ανάκτηση με βάση τίτλους κεφαλαίων, λέξεις κλειδιά και δυνατότητα επεξεργασίας πολύπλοκων ερωτήσεων (π.χ. AND, OR, NOT).
- Σήμερα η έρευνα και ανάπτυξη στο χώρο προσανατολίζεται σε πιο εξελιγμένα user interfaces, πιο γρήγορες μηχανές αναζήτησης, ανοιχτή αρχιτεκτονική.



Ιστορική Αναδρομή-1

- 1960-70's:
 - Αναζήτηση σε μικρές συλλογές κειμένων σχετικές με επιστημονικά άρθρα και νομικά κείμενα.
 - Ανάπτυξη των βασικών μοντέλων boolean και vector.
 - Από τους σημαντικότερους ερευνητές στο χώρο είναι ο καθ. Salton και οι συνεργάτες του στο Πανεπιστήμιο Cornell.



Ιστορική Αναδρομή-2

- 1980's:
 - Ανάπτυξη μεγάλων συστημάτων ανάκτησης, πολλά από τα οποία διαχειρίζονται από εταιρίες:
 - Lexis-Nexis
 - Dialog
 - MEDLINE



Ιστορική Αναδρομή-3

- 1990's:
 - Αναζήτηση κειμένων στο Internet μέσω FTP
 - Archie
 - WAIS
 - Αναζήτηση στο World Wide Web
 - Lycos
 - Yahoo
 - Altavista



Ιστορική Αναδρομή-4

- 1990's συνέχεια:
 - Διαγωνισμοί
 - NIST TREC
 - Recommender Systems
 - Ringo
 - Amazon
 - NetPerceptions
 - Αυτόματη κατάταξη και ομαδοποίηση κειμένων.



Ιστορική Αναδρομή-5

- 2000's
 - Ανάλυση συνδέσμων (link analysis)
 - Google
 - Αυτόματη εξαγωγή πληροφορίας
 - Whizbang
 - Fetch
 - Burning Glass
 - Απάντηση ερωτήσεων (question answering)
 - TREC Q/A track



Ιστορική Αναδρομή-6

- 2000's συνέχεια:
 - Ανάκτηση πολυμεσικών δεδομένων (Multimedia IR)
 - Image
 - Video
 - Audio and music
 - Cross-Language IR
 - DARPA Tides
 - Document Summarization



Σύγχρονες Τάσεις

- Παγκόσμιος Ιστός.
- Γεωγραφική Ανάκτηση Πληροφορίας.
- Πολυμεσικά Δεδομένα.
- Ολοκλήρωση Τεχνικών Ανάκτησης και Βάσεων Δεδομένων.
- Ομότιμα Συστήματα (P2P systems).



Περιεχόμενα Διαλέξεων

- Μοντέλα ανάκτησης πληροφορίας.
- Προεπεξεργασία εγγράφων.
- Αποτίμηση αποτελεσματικότητας συστημάτων.
- Κατάλογοι για αποδοτική ανάκτηση.
- Συμπύεση εγγράφων.
- Επεξεργασία ερωτημάτων.
- Ανάδραση σχετικότητας.
- Παράλληλη και κατανεμημένη ανάκτηση.
- Ανάκτηση πολυδιάστατων δεδομένων.
- Ανάκτηση πληροφορίας στον παγκόσμιο ιστό.
- Ειδικά θέματα.



Συνέδρια

- Text Retrieval Conference (TREC).
- ACM Special Interest Group on Information Retrieval (SIGIR) Conference.
- European Conference on Information Retrieval (ECIR).
- European Conference on Research and Advanced Technology for Digital Libraries (ECDL).
- Joint Conference on Digital Libraries (JCDL).
- ACM International Workshop on Multimedia Information Retrieval (MIR).
- International Symposium on Music Information Retrieval (ISMIR), και άλλα.

Ακόμη, πολλά από τα μεγάλα συνέδρια που αναφέρονται στη διαχείριση δεδομένων έχουν ειδικές συνεδρίες για την Ανάκτηση Πληροφορίας.

Χαρακτηριστικά αναφέρουμε:

- ACM Conference on Information and Knowledge Management (CIKM).
- ACM Special Interest Group on Management of Data (SIGMOD) Conference.
- International Conference on Very Large Databases (VLDB).



Περιοδικά

- Τέλος, αναφέρουμε και επιστημονικά περιοδικά που εστιάζουν στην περιοχή:
 - ACM Transactions on Information Systems.
 - Information Retrieval.
 - Information Processing and Management.
 - Information Systems.
 - International Journal on Digital Libraries, και άλλα.



Βιβλιογραφία - βιβλία

- Manning, C.D. and Raghavan, P. and Schütze, H., “An Introduction to Information Retrieval”, Cambridge University Press”, 2007.
- Baeza-Yates, R. and Ribeiro-Neto, B., “*Modern Information Retrieval*”, Addison Wesley, 1999.
- Witten, I.H and Moffat, A. and Bell, T.C., “*Managing Gigabytes: Compressing and Indexing Documents and Images*”, Morgan Kaufmann, 1999.
- Korfhage, R., “*Information Storage and Retrieval*”, John Wiley & Sons, 1997.
- Frakes, W.B. and Baeza-Yates (eds), “*Information Retrieval: Data Structures and Algorithms*”, Prentice Hall, Englewood Cliffs, NJ, 1992.
- van Rijsbergen, C.J., “*Information Retrieval*”, Butterworths, 1979.



Χρήσιμοι Ιστότοποι

- <http://www-csli.stanford.edu/~hinrich/information-retrieval.html>
- <http://documents.cfar.umd.edu/resources/ir/>
- <http://www.glue.umd.edu/~dlrg/clir/>
- <http://www.daviddlewis.com/resources/>
- <http://www.cs.umbc.edu/~crowder/pubs/IR.html>
- http://www.google.com/Top/Computers/Software/Information_Retrieval/
- <http://ir.dcs.gla.ac.uk/resources.html>
- <http://www.searchtools.com/info/info-retrieval.html>



Σημείωμα Αναφοράς

Copyright Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Απόστολος Παπαδόπουλος. «Ανάκτηση πληροφορίας. Εισαγωγή στην Ανάκτηση Πληροφορίας». Έκδοση: 1.0. Θεσσαλονίκη 2014. Διαθέσιμο από τη δικτυακή διεύθυνση: <http://eclass.auth.gr/courses/OCRS388/>



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά - Μη Εμπορική Χρήση - Όχι Παράγωγα Έργα 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

[1] <http://creativecommons.org/licenses/by-nc-nd/4.0/>





Τέλος ενότητας

Επεξεργασία: <Μαυρίδης Απόστολος>
Θεσσαλονίκη, <Εαρινό εξάμηνο 2013-2014>



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ



ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

Σημειώματα

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

