



HY463 - Συστήματα Ανάκτησης Πληροφοριών Information Retrieval (IR) Systems

Διεπαφές Χρήστη & Οπτικοποίηση (User Interfaces & Visualization)

Γιάννης Τζιτζίκας

Διάλεξη : 23

Ημερομηνία : 13-6-2007



Διάρθρωση

- Εισαγωγή
 - Χρήστης και Προσπέλαση Πληροφορίας
 - Το Γενικό Πλαίσιο
 - Πληροφοριακές Ανάγκες Χρηστών
- Διεπαφές Χρήσης για
 - (a) **Starting Points** (Σημεία Εκκίνησης)
 - Τρόποι Πλοήγησης
 - Ιεραρχική, Οπτική Scatter/Gather
 - Browsing vs Searching
 - (b) **Query Specification**
 - Formed-based Query Specification
 - Boolean Queries
 - Parametric Search
 - (c) **Viewing Retrieval Results** (in Context)
 - Exploring the Results (Cat-a-cone, Summarization)
 - (d) **Interactive Relevance Feedback**

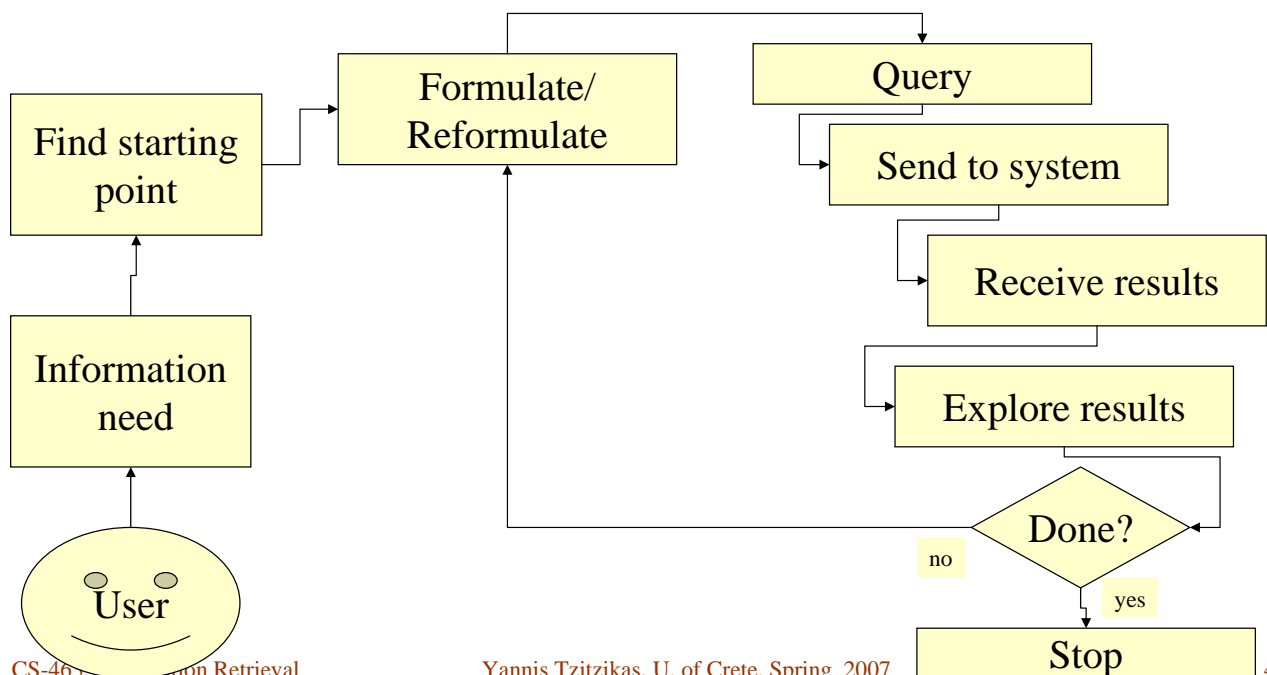


Χρήστης και Προσπέλαση Πληροφορίας

- Information seeking is an imprecise process
 - users may have only a *fuzzy understanding* of how they can achieve their goals
- UI should help users
 - understand and express their information needs
 - formulate their queries and understand the results



Χρήστης και Προσπέλαση Πληροφορίας



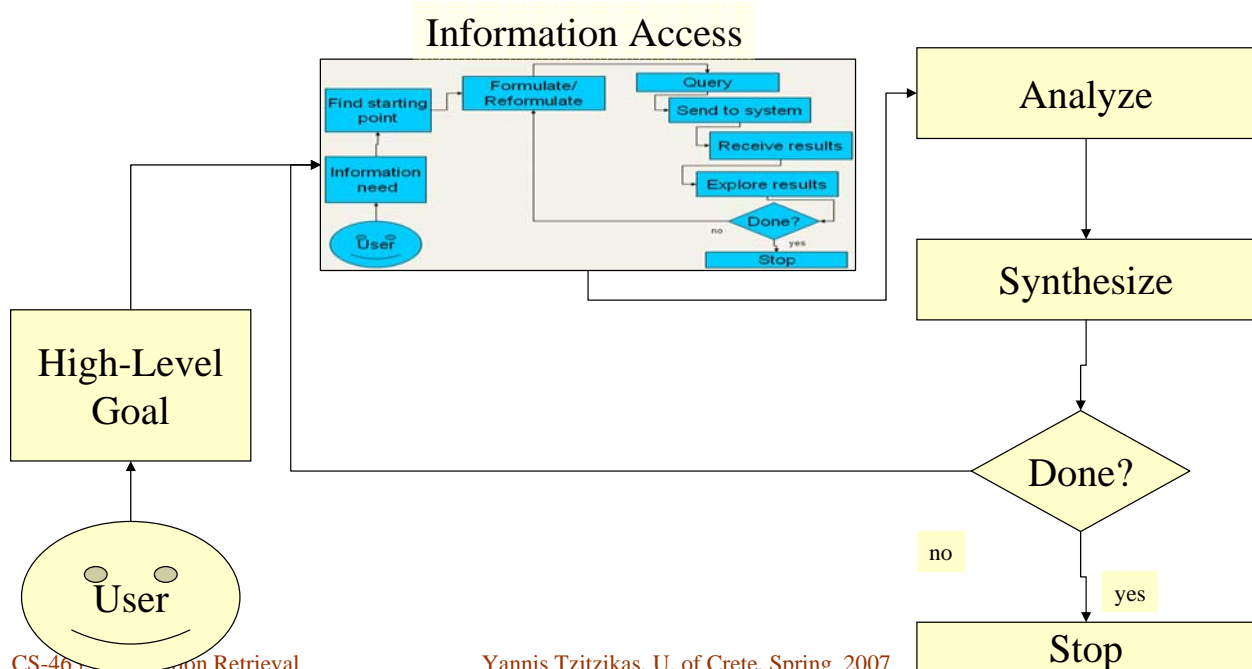


Χρήστης και Προσπέλαση Πληροφορίας Το Γενικότερο Πλαίσιο

- [O'Day and Jeffries]: Information seeking is only one part of the full work process. In between searching sessions many different work was done with the retrieved information, including reading and annotating and analysis.
- Decomposition of the **analysis step**
 - 80%
 - Finding trends
 - Making comparisons
 - Aggregating information
 - Identifying a critical subset
 - Assessing and interpreting
 - 20%
 - Cross-referencing, Summarizing
 - Finding evocative visualizations for reports, Miscellaneous activities

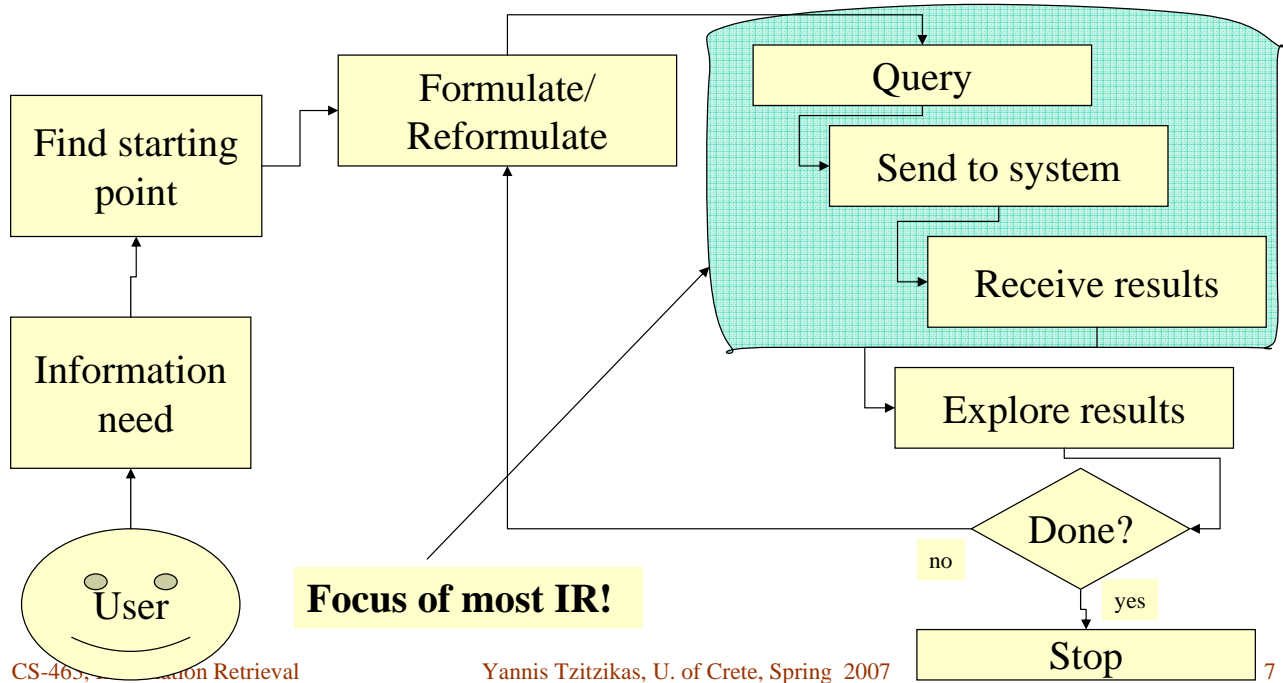


Χρήστης και Προσπέλαση Πληροφορίας Το Γενικότερο Πλαίσιο

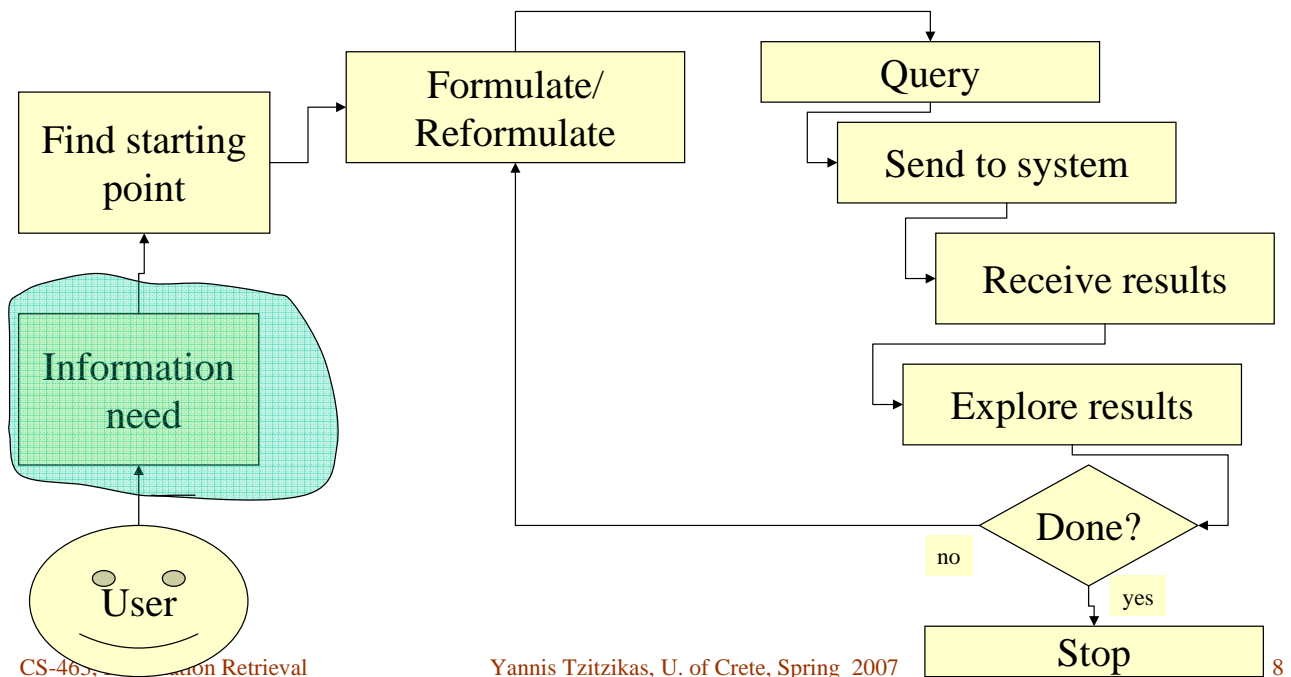




Που εστιάζει η πλειονότητα των Συστημάτων Ανάκτησης Πληροφορίας;



Πληροφοριακές Ανάγκες Χρηστών





Τύποι Πληροφοριακών Αναγκών Types of Information Needs

- Most search engines do not disclose their search logs, so information about what users are searching for on the Web is difficult to come by. Nevertheless, a study in 2001 analyzed the queries from the Excite search engine showed some interesting characteristics of web search:
 - The average length of a search query was 2.4 terms.
 - About half of the users entered a single query while a little less than a third of users entered three or more unique queries.
 - Close to half of the users examined only the first one or two pages of results (10 results per page).
 - Less than 5% of users used advanced search features (e.g Boolean Operators like AND, OR, and NOT).
- The top three most frequently used terms were and, of, and sex.



Τύποι Πληροφοριακών Αναγκών Types of Information Needs

In the web context the "need behind the query" is often not informational in nature. [1] classified web queries according to their intent into 3 classes:

1. **Navigational.** The immediate intent is to reach a particular site.
2. **Informational.** The intent is to acquire some information assumed to be present on one or more web pages.
3. **Transactional.** The intent is to perform some web-mediated activity.

From a user survey and a log analysis

Type of Query	User Survey	Query Log Analysis
Navigational	24%	20%
Informational	39%	48%
Transactional	36%	30%

- Source:
 - [1] A. Broder, "A Taxonomy of Web Search", <http://www.acm.org/sigs/sigir/forum/F2002/broder.pdf>



Τύποι Πληροφοριακών Αναγκών Types of Information Needs

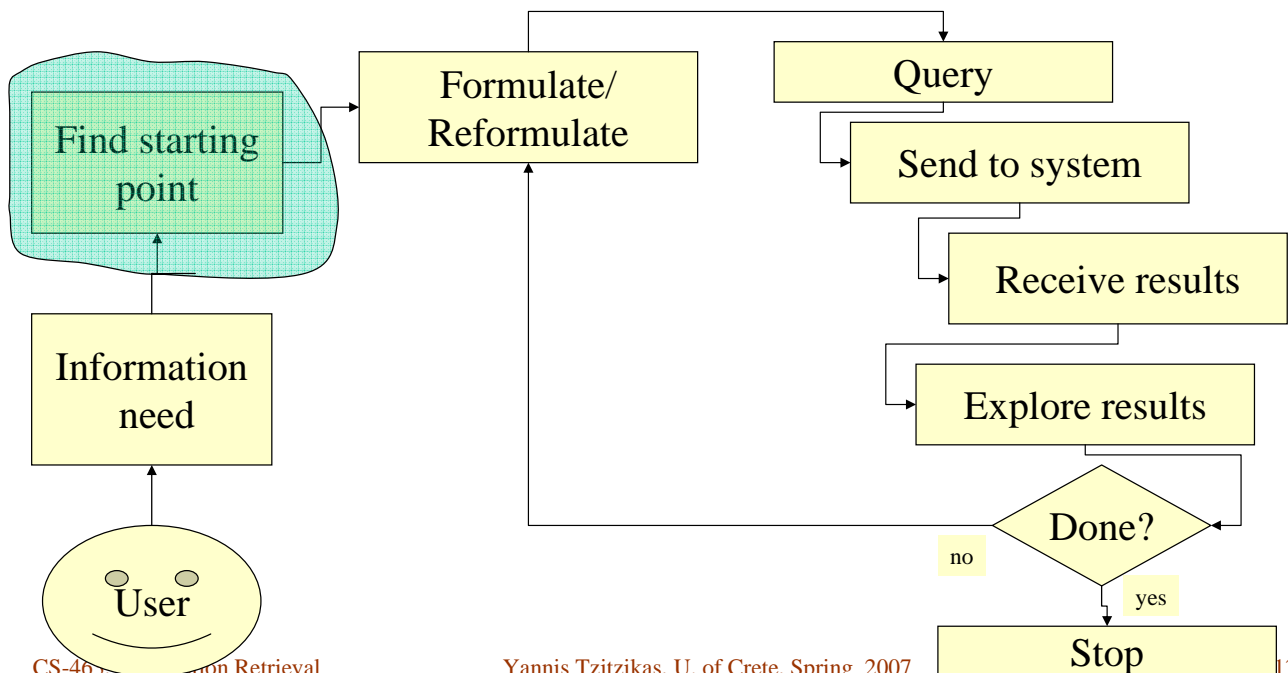
- Need **answer** to question
- **Re-find** a particular document
- **Find** a good recipe for tonight's dinner
- **Exploration** of new area (browse sites about Mexico City)
- **Authoritative summary** of information (HIV review)

In most cases, only one interface!

Current Web searching research tries to *answer the need behind the query.*



Χρήστης και Προσπέλαση Πληροφορίας





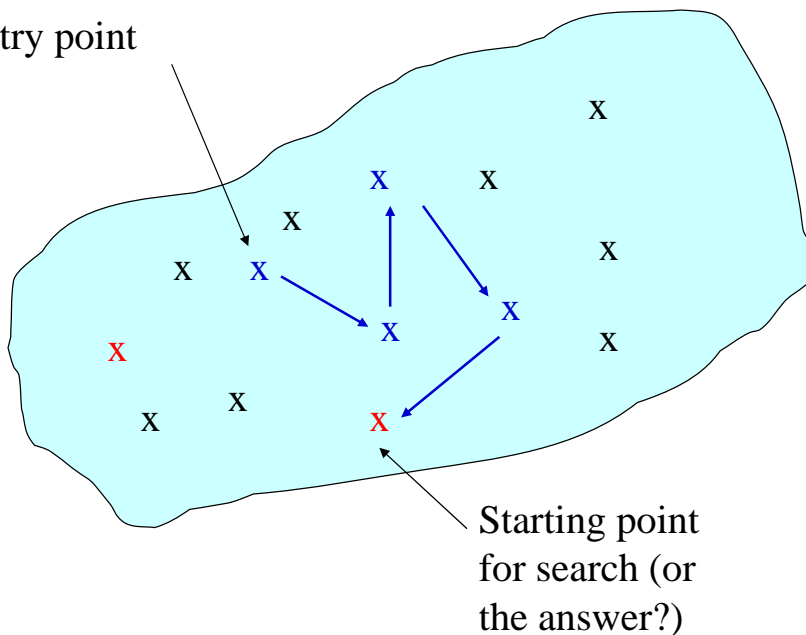
(A) Επαφές Χρήσης για Σημεία Εκκίνησης

- Methods for finding a starting point
 - Select collections from a list
 - Overviews
 - Category Hierarchies
 - Clustering Techniques
 - Co-citation Analysis



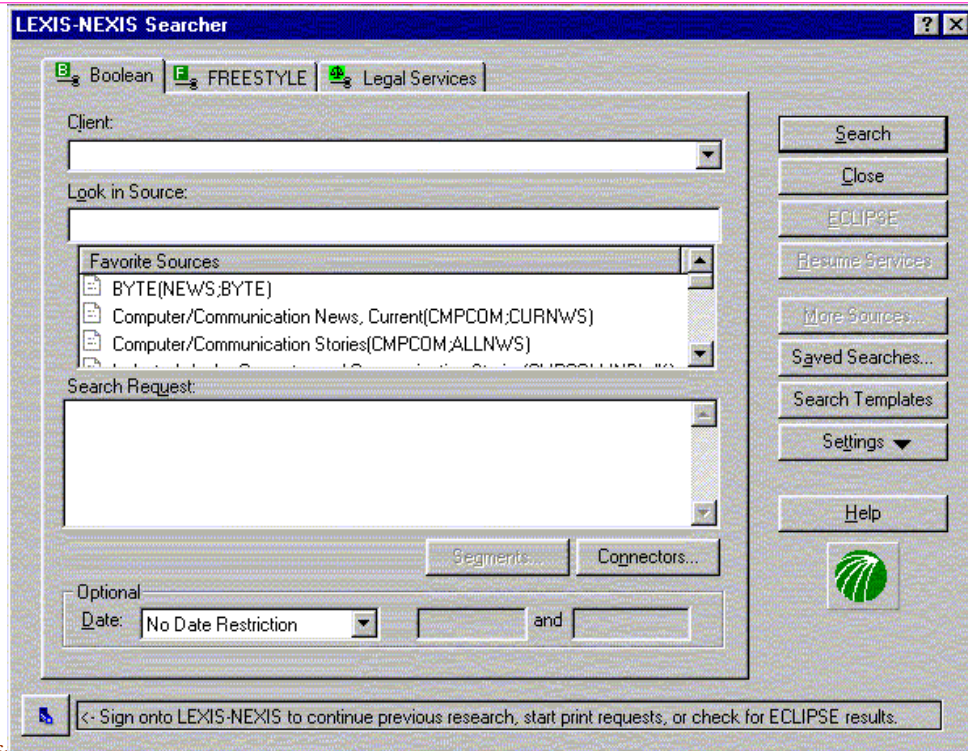
(A) Επαφές Χρήσης για Σημεία Εκκίνησης

Entry point





(A) Επαφές Χρήσης για Σημεία Εκκίνησης Select collections from a list



CS-463, Information Retrieval

Yannis Tzitzikas, U. of Crete, Spring 2007

16

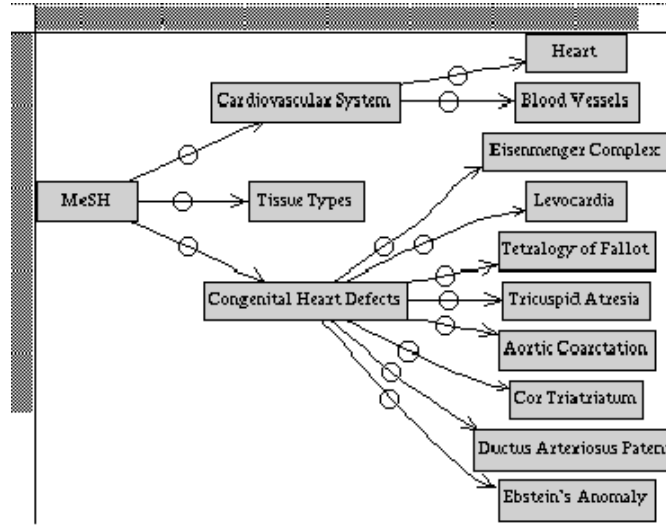


(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews

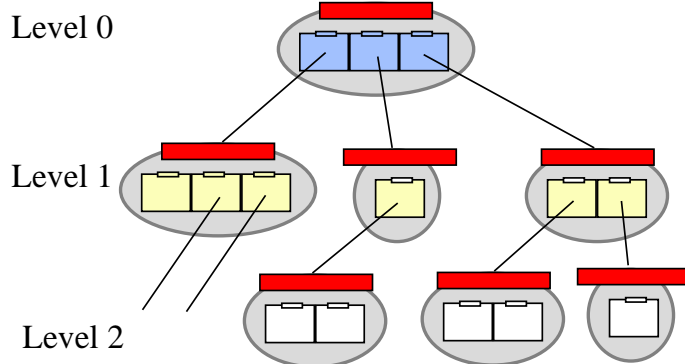
- Category Hierarchies
- Clustering Techniques
- Co-citation Analysis



(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>Category Hierarchies: MeSHBrowse

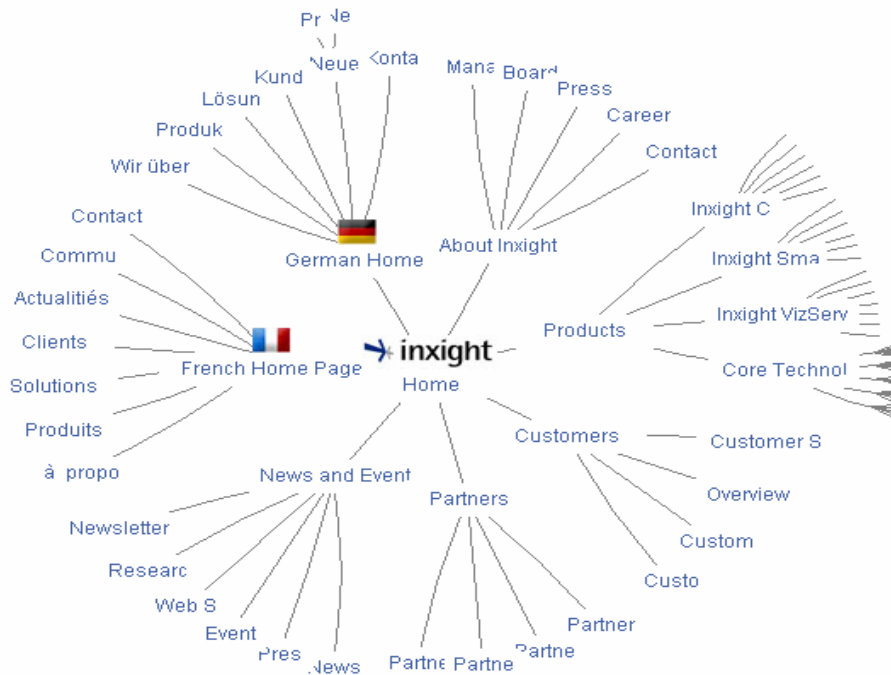


(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>Category Hierarchies: Yahoo!

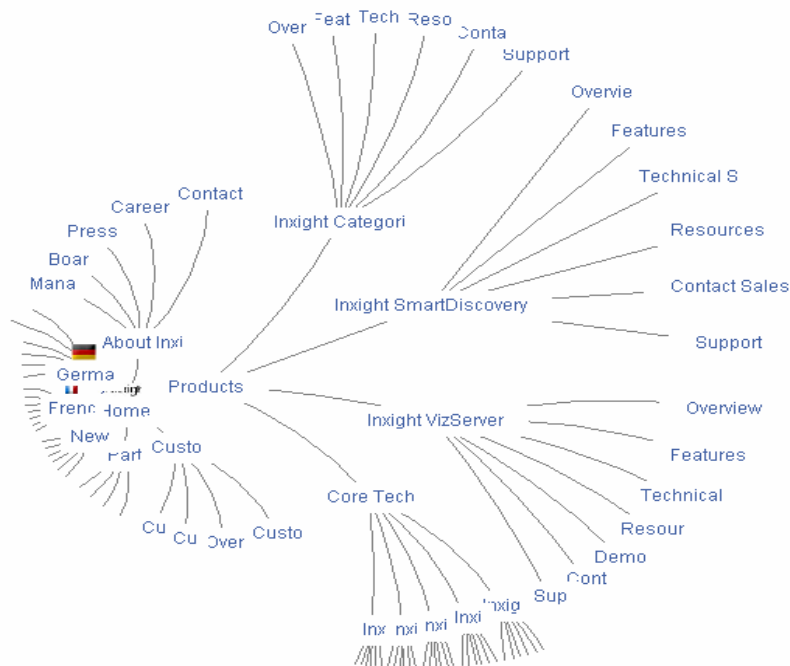




(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>Hyperbolic Trees

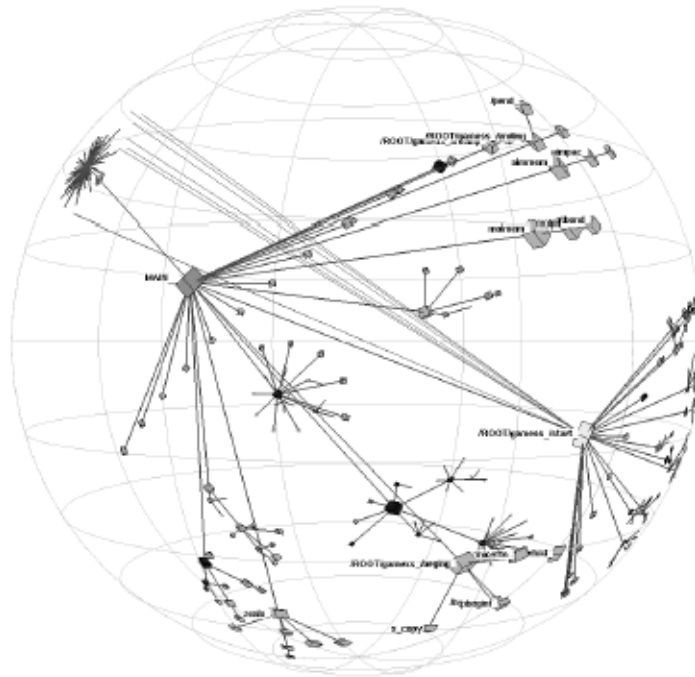


(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>Hyperbolic Trees

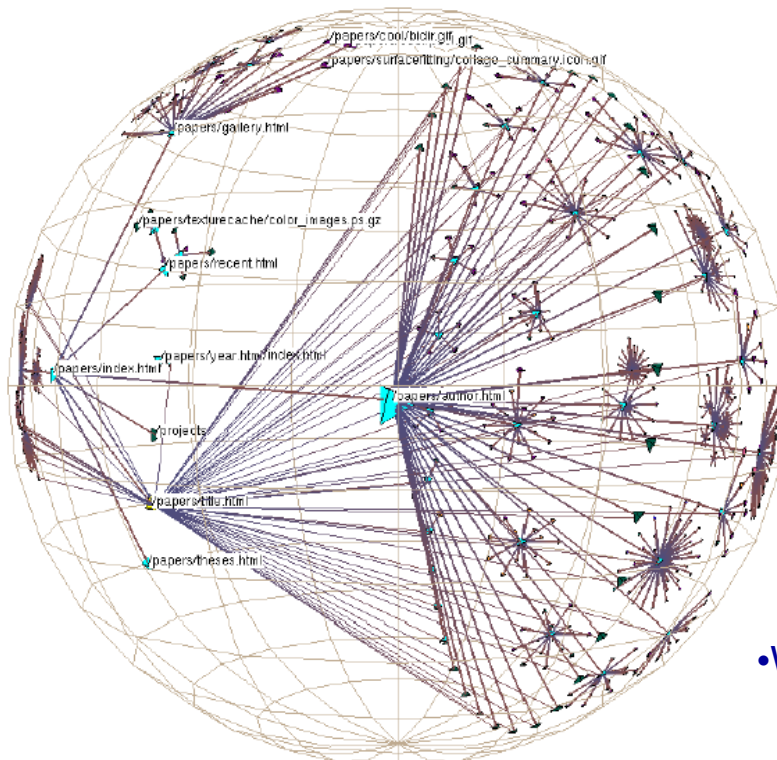




(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>Hyperbolic Trees



(A) Επαφές Χρήσης για Σημεία Εκκίνησης Ονει



•Web site



(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>**RINGS**

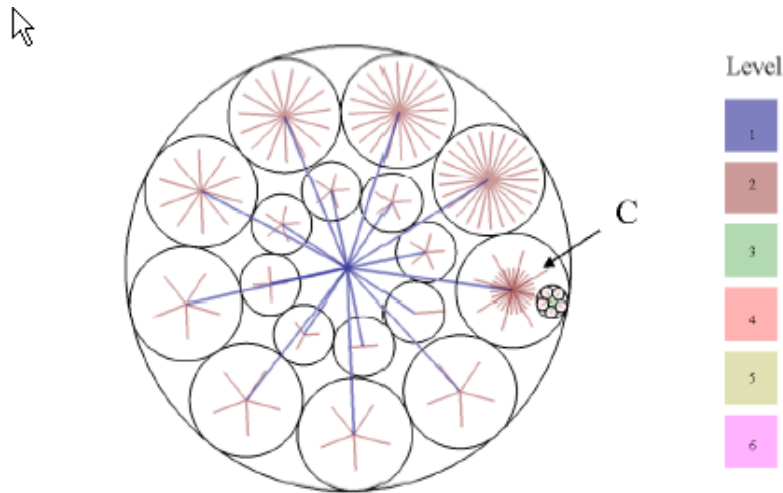
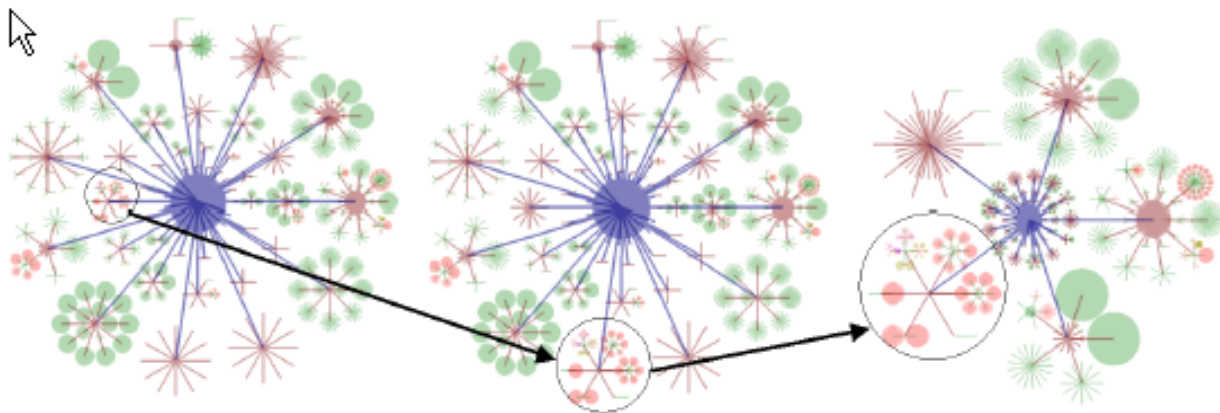


Fig. 2. Each node is placed at the center of its circle, and edges are drawn between each pair of connected nodes. Edges are colored according to their distance from the root. Level 1 edges are those directly connected to the root.

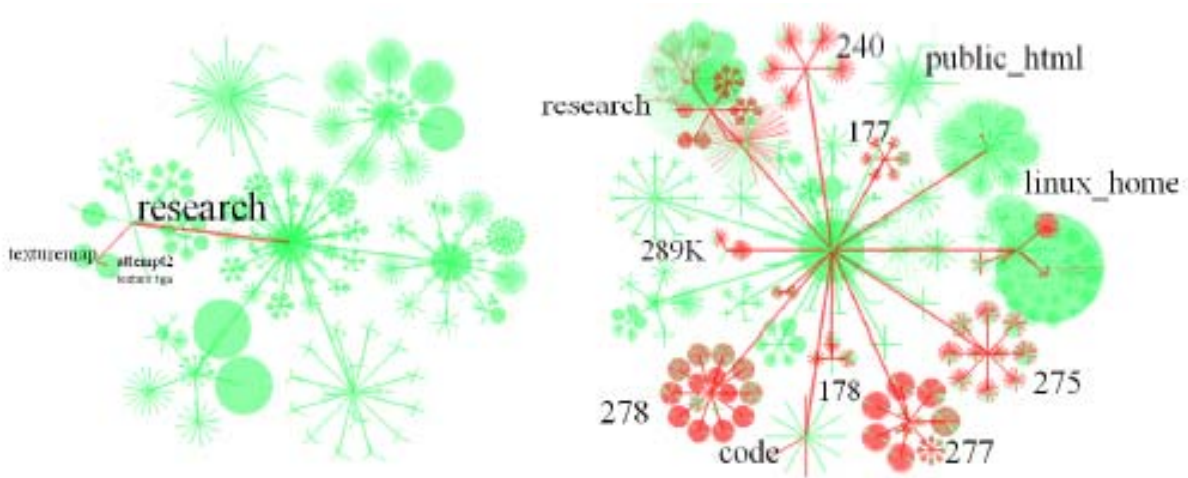


(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>**RINGS**

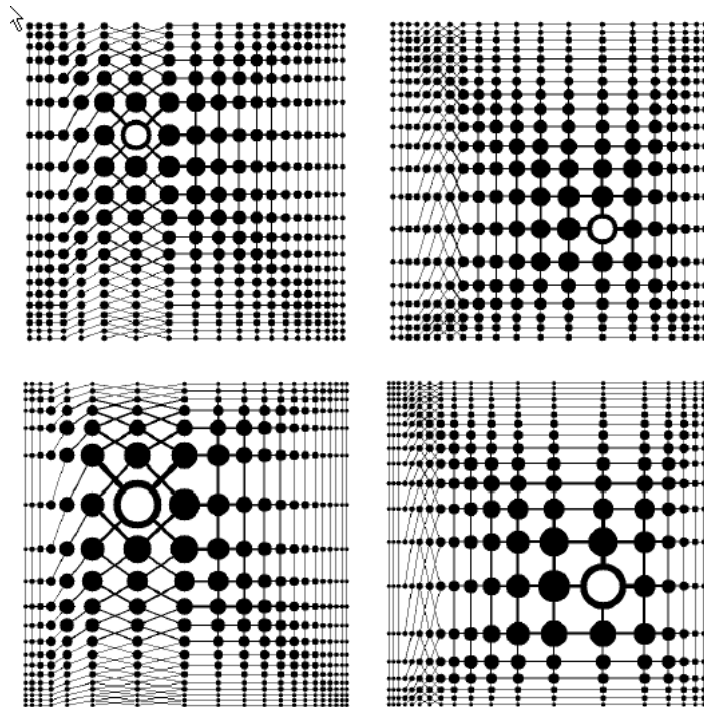




(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>RINGS

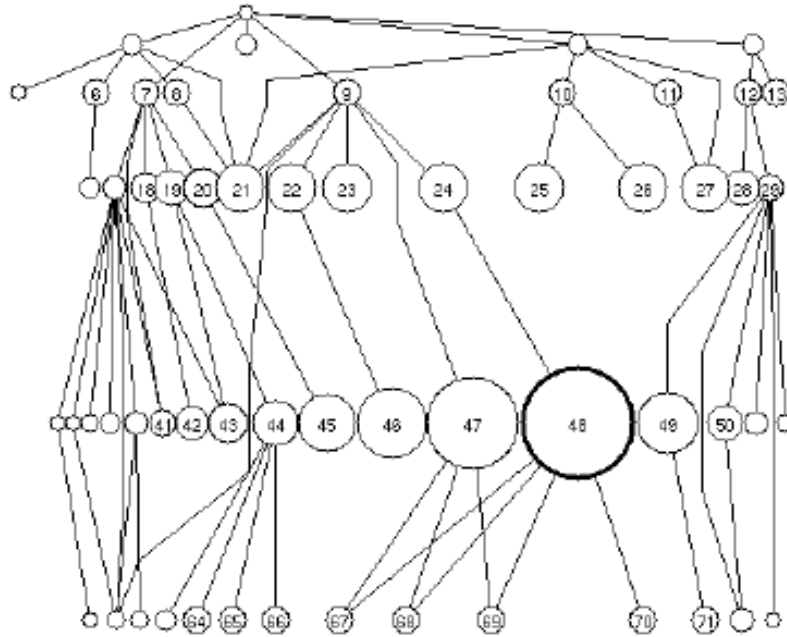


A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>Fish Eye View





A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>Fish Eye View



A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews>

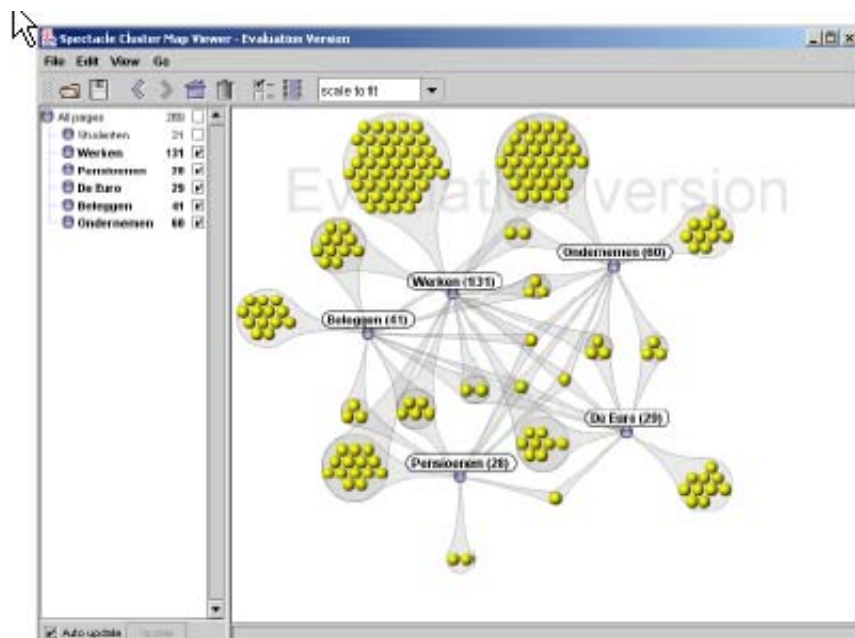


Figure 1: Spectacle



(A) Επαφές Χρήσης για Σημεία Εκκίνησης Overviews > **Category Hierarchies: HiBrowse**

The screenshot displays the HiBrowse software interface with three active windows showing hierarchical category views:

- physical disease by anatomical structure**: 12661 items. Categories include abdominal disease (313), abnormal body bu: (189), breast disease (34), cardiovascular d: (2209), connective tissu: (248), digestive system (2395), ear nose throat (774), endocrine diseas: (1195), eye disease (840), head and neck di: (764), hematologic dise: (2648), mouth disease (376), musculoskeletal (1620), neurologic disea: (3236), pelvic disease (17), respiratory tract (2527), skin disease (1509), soft tissue dise: (48), thorax disease (56), and urogenital tract (1537).
- therapy**: 12661 items. Categories include therapy (in ge: (2468), acupuncture (11), biological the: (1976), cancer therapy (724), computer assis: (2), conservative t: (481), counseling (116), detoxification (58), disease contro: (171), drug therapy (8467), drug therapy (8282), adjuvant the: (53), antibiotic pi: (88), antibiotic tl: (300), anticoagulan: (19), anticonvulsai: (60), antihyperten: (15), antimicrobia: (46), bone marrow ! (72), chelation th: (15), chemical sym: (1), chemoprophyl: (14), chemotherapy (629), and diuretic the: (5).
- child**: 12661 items. Categories include child (in genera: (8411), brain damaged ch: (1), handicapped child: (15), infant (3696), preschool child (1855), and school child (2371).

At the bottom left, it says "3 view[s]" and at the bottom right, the number "30" is visible.



(A) Επαφές Χρήσης για Σημεία Εκκίνησης **Automatic overviews: Scatter/Gather** (~ Clustering)

Scatter/Gather browsing paradigm:

- Allows the user to find a set of documents of interest through browsing.
- It iterates:
 - **Scatter**: Take the collection and scatter it into **n clusters**.
 - **Gather**: Pick the clusters of interest and **merge** them



A) Επαφές Χρήσης για Σημεία Εκκίνησης Automatic overviews: Scatter/Gather

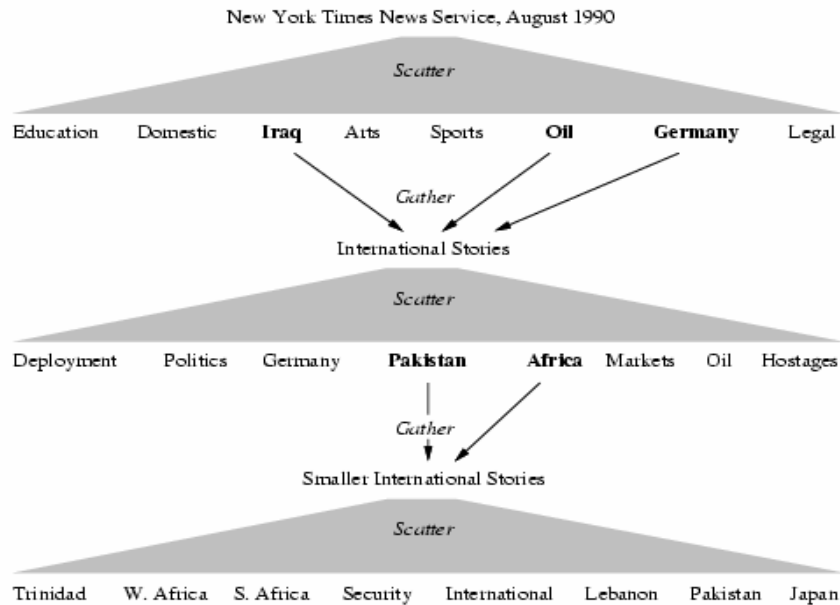


Figure 1: Illustration of Scatter/Gather

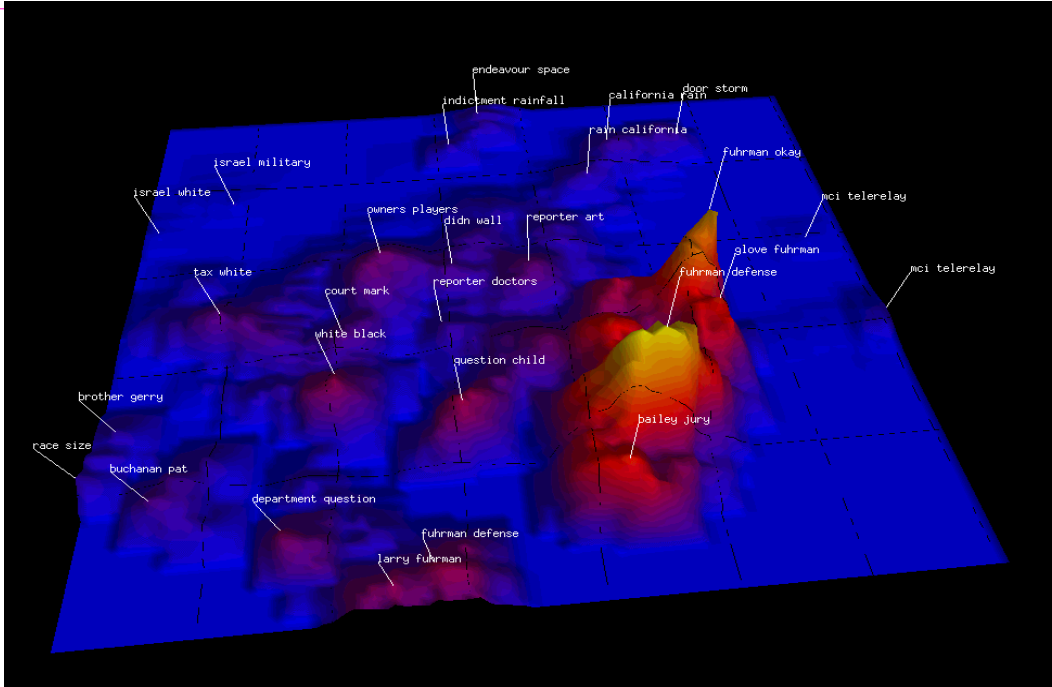


A) Επαφές Χρήσης για Σημεία Εκκίνησης Automatic overviews: Scatter/Gather

<input type="checkbox"/> Cluster 1 Size: 8	key army war francis spangle banner air song scott word poem british
<input type="radio"/>	Star-Spangled Banner, The
<input type="radio"/>	Key, Francis Scott
<input type="radio"/>	Fort McHenry
<input type="radio"/>	Arnold, Henry Harley
<input type="radio"/>	Millican, Arthur
<input type="checkbox"/> Cluster 2 Size: 68	film play career win television role record award york popular stage p
<input type="radio"/>	Burstyn, Ellen
<input type="radio"/>	Stanwyck, Barbara
<input type="radio"/>	Berle, Milton
<input type="radio"/>	Zukor, Adolph
<input type="radio"/>	Doublet, Talcott
<input type="checkbox"/> Cluster 3 Size: 97	bright magnitude cluster constellation line type contain period spectr
<input type="radio"/>	star
<input type="radio"/>	Galaxy, The
<input type="radio"/>	extragalactic systems
<input type="radio"/>	interstellar matter
<input type="radio"/>	cluster star
<input type="checkbox"/> Cluster 4 Size: 67	astronomer observatory astronomy position measure celestial telesco
<input type="radio"/>	astronomy and astrophysics
<input type="radio"/>	astrometry
<input type="radio"/>	Agema
<input type="radio"/>	astronomical catalogs and atlases
<input type="radio"/>	Hausman, C. W. House
<input type="checkbox"/> Cluster 5 Size: 10	family specie flower animal arm plant shape leaf brittle tube foot hor
<input type="radio"/>	blazing star
<input type="radio"/>	brittle star
<input type="radio"/>	bishop's-cap
<input type="radio"/>	feather star



A) Επαφές Χρήσης για Σημεία Εκκίνησης Automatic overviews: Themescape

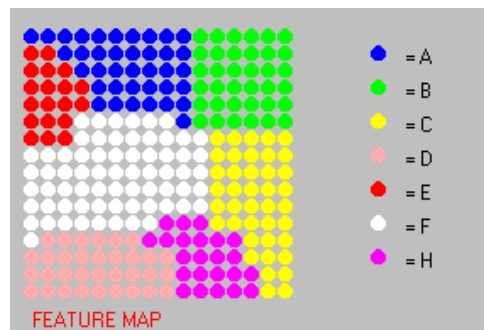


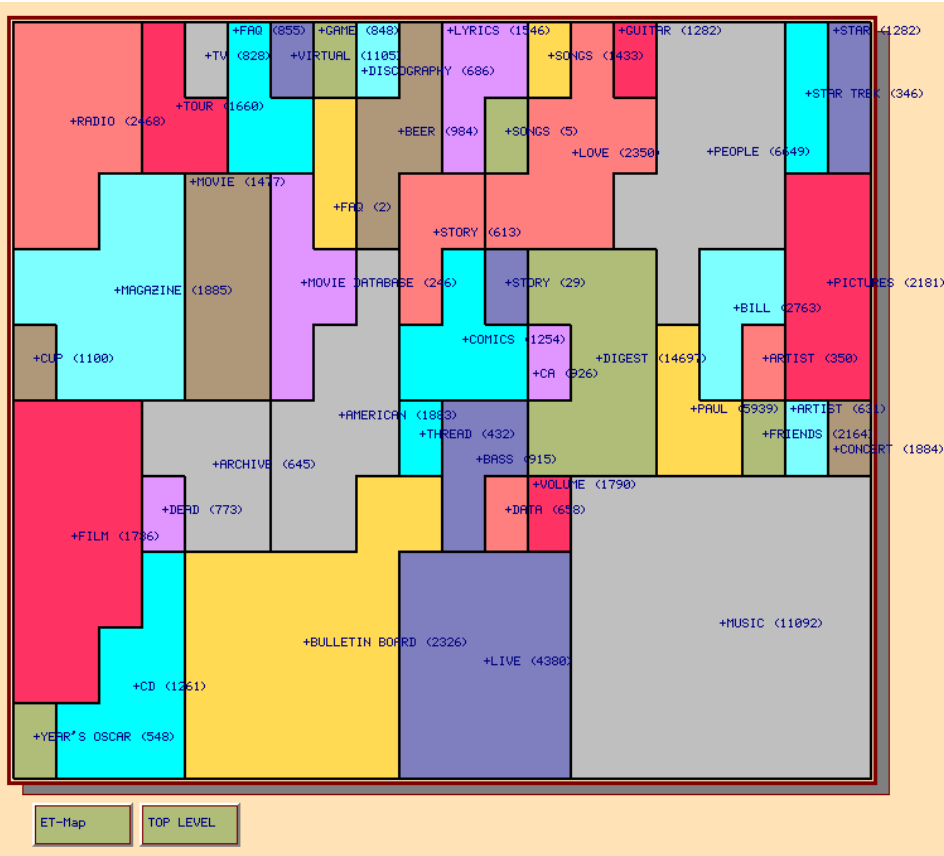
Three-dimensional overview based on clustering



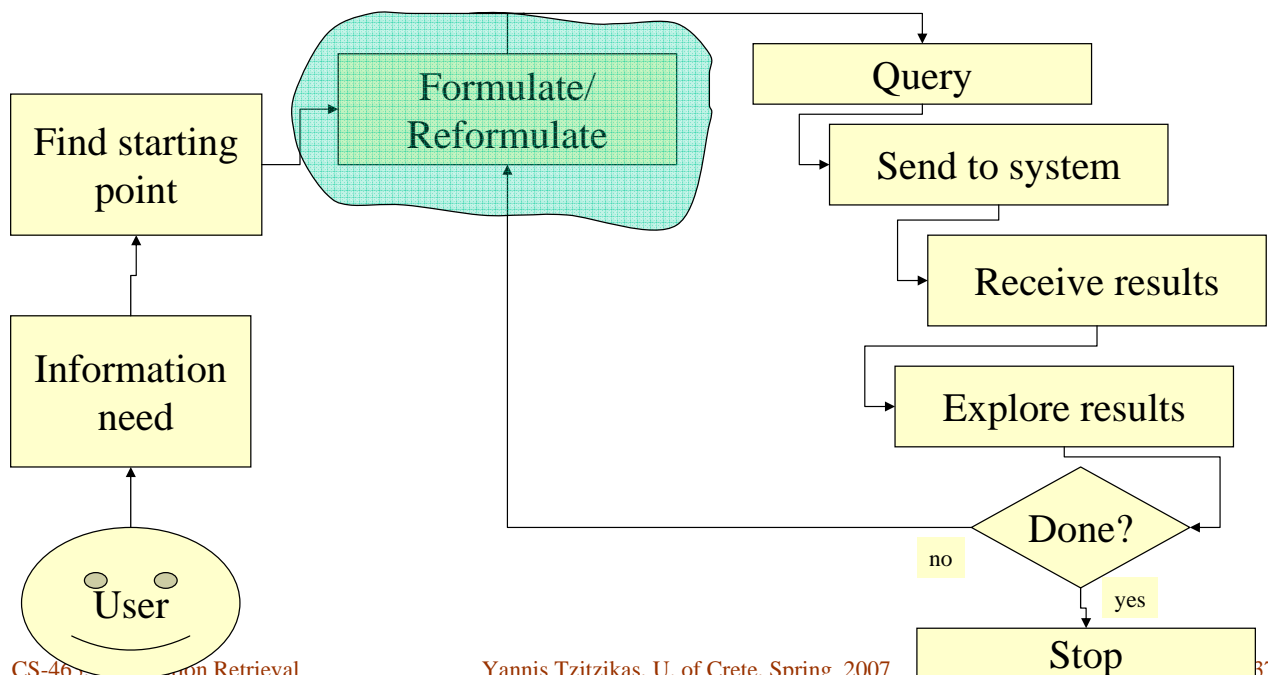
A) Επαφές Χρήσης για Σημεία Εκκίνησης Automatic overviews: Kohonen's feature map

- Algorithm for creating maps that graphically characterize the overall content of a document collection
 - the regions of the 2D map vary in size and shape corresponding to how frequently documents assigned to the corresponding themes occur within the collection
 - Regions are characterized by single words or phrases
 - adjacency of regions reflects semantic relatedness of the themes





(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων





Browsing vs. Searching

- Browsing and searching are often interleaved.
- Information need dependent
 - Open-ended (find information about Mexico City) -> browsing
 - Specific (who won the Super Bowl) -> searching
- Compare to physical world
 - Browsing vs. searching in a grocery store
- User dependent
 - Some users prefer searching, others browsing (confirmed in many studies: some hate to type)
 - Advantage of browsing: You don't need to know the vocabulary of the collection
- Browsers vs. Searchers
 - 1/3 of users do not search at all
 - 1/3 rarely search (or URLs only)
 - Only 1/3 understand the concept of search



Boolean Queries

- Boolean logic is difficult for the average user.
- Some interfaces for average users support formulation of boolean queries
- Current view is that non-expert users are best served with non-boolean or simple +/- boolean (pioneered by Altavista).
- But boolean queries are the standard for certain groups of expert users (eg, lawyers).



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων

Types of Interface:

- **Command line**
 - Single text box (google)
- **Form fillin**
- **Parametric Search**
- **Menu selection**
- **Direct manipulation**
 - continuous representation of the object of interest
 - physical actions or buttons presses instead of complex syntax
 - rapid incremental reversible operations whose impact on the object of interest is immediately visible
- **Natural Language**



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Με Συμπλήρωση Φορμών

The screenshot shows a web browser window with the URL <http://192.35.215.185/mw/mwcgi.home>. The page displays the MELVYL system search interface. On the left is a navigation menu with options: New Search, Search History, Saved Lists, Profile, Resources, Restart, Quit, and Help. The main content area shows a search form for the 'Current Contents' database. The 'Author' field contains 'swanson, d.' with a mouse cursor. Below it are 'Options and Limits' with dropdown menus for 'Another Author' (set to 'and') and 'Journal Title' (set to 'and'). There are radio buttons for search options: 'Any words' (selected), 'Exact beginning', and 'Complete title'. The 'Location' dropdown is set to 'UC Berkeley (all libraries)'. At the bottom of the form are 'Submit Search' and 'Back' buttons. A footer note says: 'Send questions, comments, or suggestions to melvyl@www.melvyl.ucop.edu. Melvyl® is a registered trademark of The Regents of the University of California'.



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Με Συμπλήρωση Φορμών

Infoseek Advanced Search: World Wide Web - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Location: http://www.infoseek.com/find?pg=advanced_www.html&ud9=advanced_www

Infoseek Advanced Search

WEB USENET PEOPLE FINDER COMPANIES STOCKS NEWS SHAREWARE REFERENCE

· Click each box to see options. · Enter query terms. · Leave unused boxes blank.

Search the Web for pages in which the:

Document must contain the phrase

Document should contain the phrase

Document should not contain the phrase

Show only results from .com

Show my results: Number of results:

Show Summaries 10 per page

Search by collection: Search by location: search clear form Tips

Entire Web Earth

This form requires JavaScript.

Document: Done

Credit: Marti Hearst



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Parametric search

- Each document has, in addition to text, some “meta-data” e.g., Make, Model, City, Color
- A parametric search interface allows the user to combine a full-text query with selections on these parameters



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Parametric search

CarFinder.com 

Over one million fictional vehicles to choose from!

Choose your search criteria from the drop down menus:

Number of results to display:

Make Model Category Year
 City Color Price



Notice that the output is a (large) table. Various parameters in the table (column headings) may be clicked on to effect a sort.

Make	Model	Year	City	Mileage	Price	Category	Description	Color
BMW	5-Series	1995	San Francisco	16100	11100	Luxury	Never driven in winter conditions. Body work makes it look like new. Keyless entry and security features. This is a bargain.	Silver
BMW	5-Series	1995	San Francisco	16600	11100	Luxury	Great first car for your teen-aged kid. Solid, dependable, affordable with 0% down and owner financing.	Blue
BMW	5-Series	1995	San Francisco	16800	11200	Luxury	Upgraded sound system really rocks. Customized interior features wood grain dash and beige leather seats. Power locks, windows, steering. Price firm.	White
BMW	5-Series	1995	San Francisco	16100	11300	Luxury	Safe choice for a young family: ABS, driver and passenger air bags. Roomy interior with power everything. Low mileage driving kids back and forth to soccer.	Maroon
BMW	5-Series	1995	San Francisco	16300	11400	Luxury	This baby's got it all: power steering, cruise, power locks, power windows, remote entry, leather interior, security alarm, AM/FM/CD/Cassette. Priced to sell!	Brown



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Parametric search

CarFinder.com 

Over one million fictional vehicles to choose from!

Choose your search criteria from the drop down menus:

We can add text search.

Number of results to display:

Make Model Category Year
 City Color Price Description

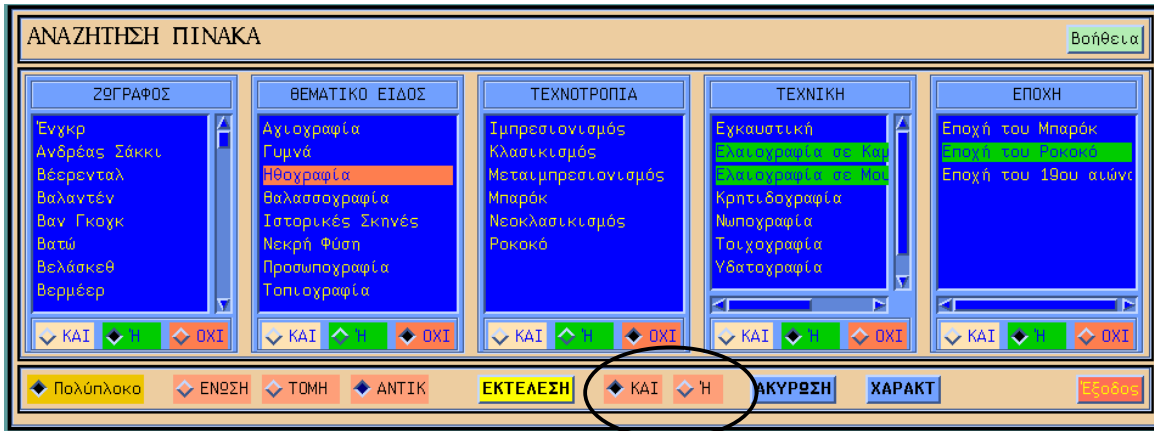


Make	Model	Year	City	Mileage	Price	Category	Description	Color
BMW	5-Series	1997	San Francisco	14300	13100	Luxury	5-speed, heavy-duty suspension, extra wide tires. Well-maintained by mechanic-owner. Cloth seats and upgraded stereo system.	White
BMW	5-Series	1997	San Francisco	14600	13100	Luxury	Is that price for real? You bet it is. Fully loaded with all factory options. Former floor model.	Beige
BMW	5-Series	1997	San Francisco	14900	13100	Luxury	Fun to drive. Manual 5-speed transmission, turbo charger. Garaged all winter and pampered the rest of the year. This is a steal!	Orange
BMW	5-Series	1997	San Francisco	14800	13200	Luxury	Fully loaded, automatic transmission. Power everything. Anti-lock brakes and full safety features. Must test drive. Price firm.	Green
BMW	5-Series	1997	San Francisco	14300	13200	Luxury	Formerly an executive's vehicle. Interior has been professionally maintained, engine factory serviced every 3000 miles. Great gas mileage. Price negotiable.	Maroon
BMW	5-Series	1997	San Francisco	15000	13200	Luxury	Sun roof, air, CD player, driver side air bag. 10% deposit required. Owner financing available. Best offer by end of weekend buy it!	Red



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Menu-based Interface for Faceted Boolean Queries

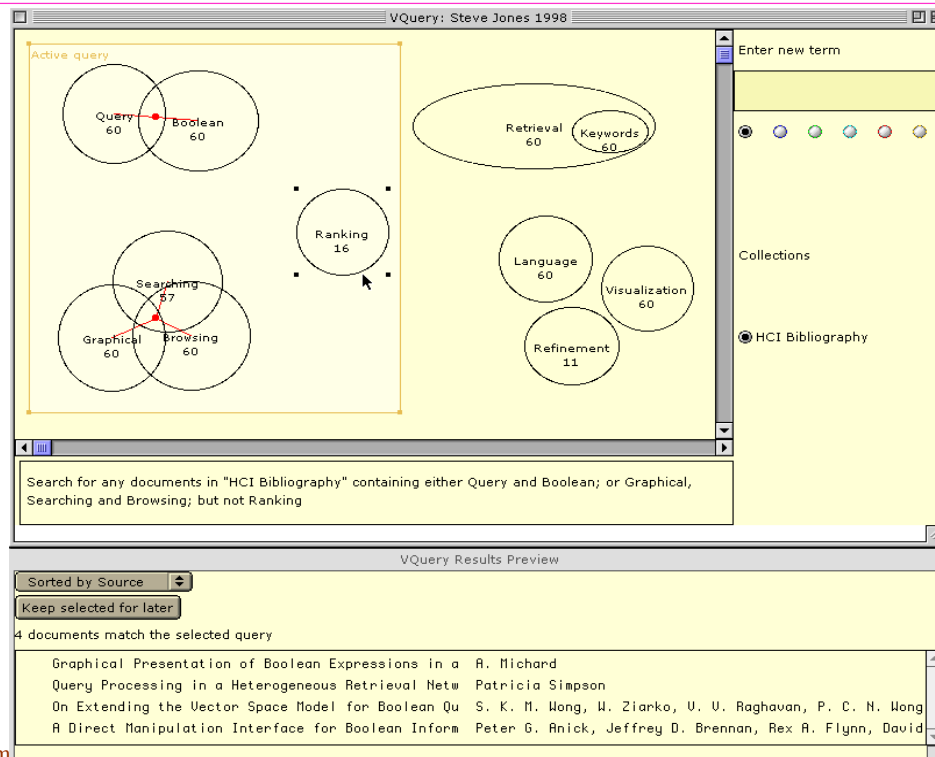
Domenicus [Tzitzikas & Theodorakis, Hypertext'96]:



Τρόπος συνδυασμού των
συνθηκών των λιστών

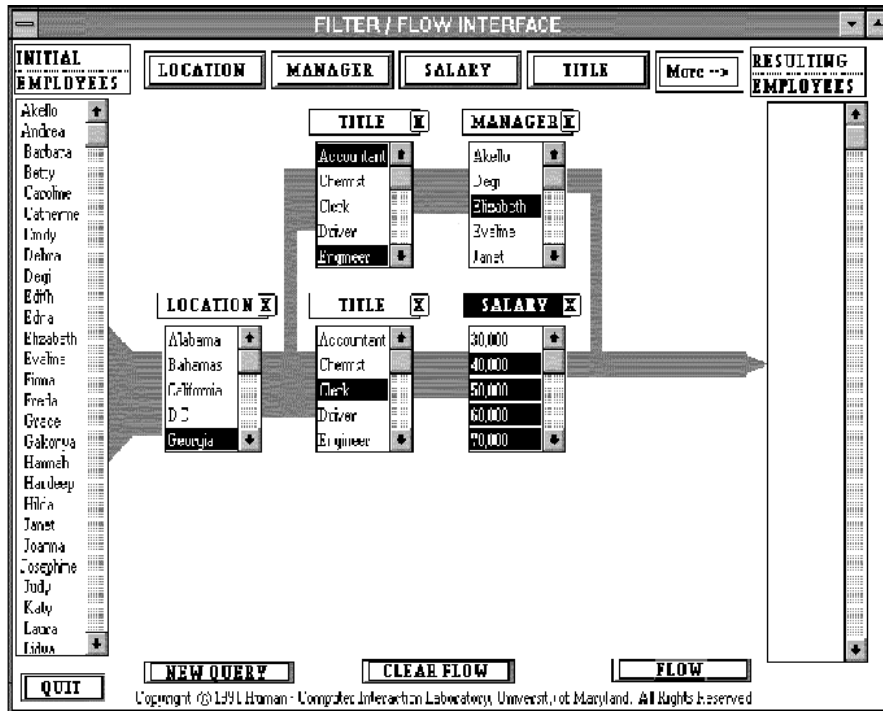


(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Direct Manipulation (VQuery Venn diagram)





(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Direct Manipulation (Filter flow visualization)



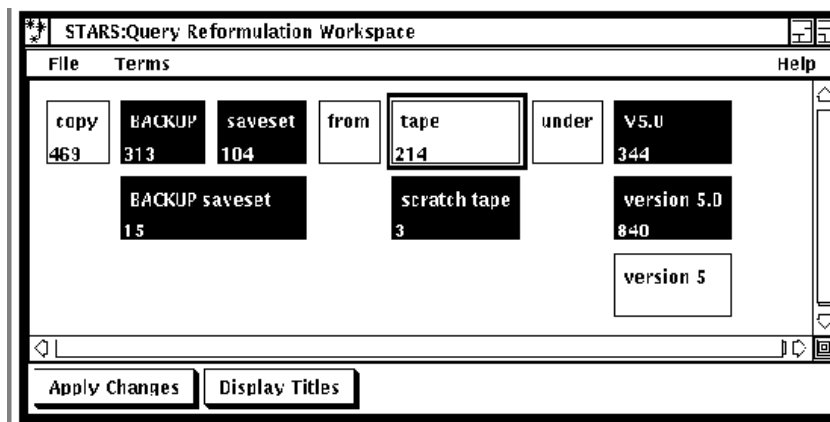
CS-463, Information Retrieval

Yannis Tzitzikas, U. of Crete, Spring 2007

48



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Direct Manipulation (Block oriented diagram)



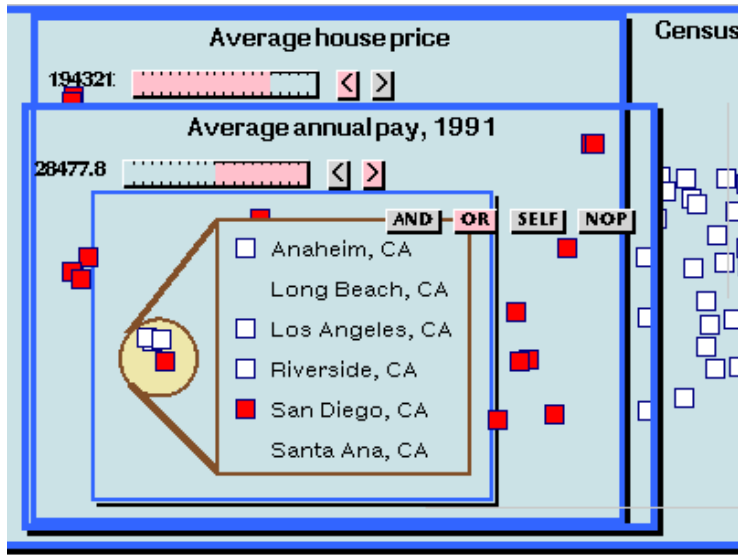
CS-463, Information Retrieval

Yannis Tzitzikas, U. of Crete, Spring 2007

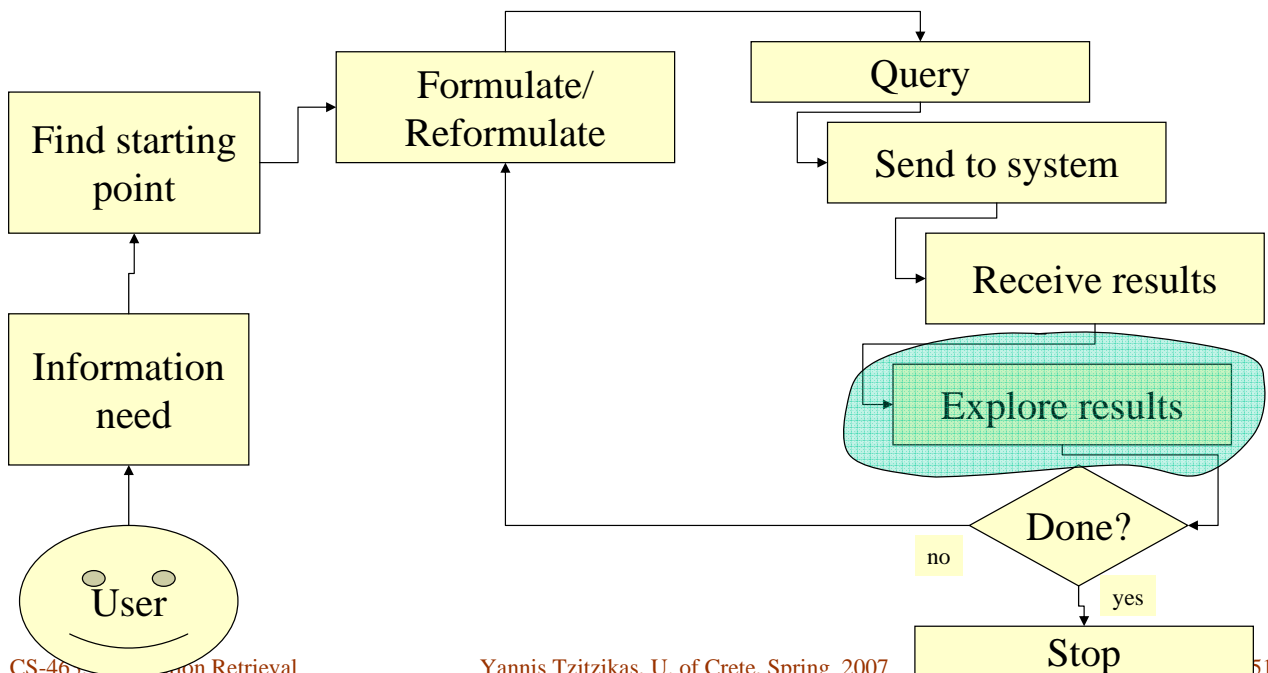
49



(B) Επαφές Χρήσης για Διατύπωση Επερωτήσεων: Direct Manipulation (Magic Lens)



(Γ) Επαφές Χρήσης για: Viewing Results in Context





(Γ) Επαφές Χρήσης για: Viewing Results in Context

User Goal: Do these results answer my question?

Methods:

- **Document Surrogates**
 - Metadata: URL, date, file size, author, Category labels
 - Summarization (?)
- **Query Term Hits Within Document Content**
 - KWIC (Keyword in Context), TileBars, SeeSoft
- **Query Term Hits Between Documents**
 - InfoCrystal, VIBE and Lyberworld, Lattices
- **SuperBook: Context via Table of Contents**
- **Categories for Results Set Context**
- **Use of Hyperlinks**
 - Cha-Cha
 - Mapuccino: Graphical Depiction of Link Structure
 - Tables



Γ) Viewing Results in Context Document Surrogates and KWIC

Surrogate

Google

Alan Turing

Αναζήτηση

Αναζήτηση: στον ιστό σελίδες γραμμένες στα Ελληνικά σελίδες από Ελλάδα

Ιστός

Αποτελέσματα 1 - 10 από περίπου 458.000 για Alan Turing. (0,29 δευτερόλεπτα)

[Alan Turing - Home Page](#)

Alan Turing Home Page. Guide to a large website maintained by Andrew Hodges, biographer of Alan Turing (1912-1954)

www.turing.org.uk/turing/ - 18k - Αποθηκευμένη Σελίδα - Παρόμοιες σελίδες

[Alan Turing](#)

Welcome to the domain of www.turing.org.uk maintained by Andrew Hodges.

www.turing.org.uk/ - 11k - Αποθηκευμένη Σελίδα - Παρόμοιες σελίδες

[Περισσότερα αποτελέσματα από το www.turing.org.uk]

[Alan Turing](#)

Biography from the MacTutor History of Mathematics archive.

www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Turing.html - 28k - Αποθηκευμένη Σελίδα - Παρόμοιες σελίδες

KWIC



(Γ) Viewing Results in Context Document Surrogates and Category Labels

Terry Winograd

Terry Winograd. Professor ...
Informatica . Publications. Books: **Terry Winograd**, Understanding Natural Language, Academic Press, 1972. ...
Description: **Terry Winograd** is one of the foreground figures in research into human/computer interaction and natural...
Category: Computers > History > Pioneers
nci.stanford.edu/~winograd/ - 5k - Nov 1, 2003 - Cached - Similar pages

- Example:
 - ODP categories on google
- Advantages:
 - Interpretable
 - Capture summary information
 - Describe multiple facets of content
 - Domain dependent, and so descriptive
- Disadvantages
 - Domain dependent, so costly to acquire
 - May mis-match users' interests

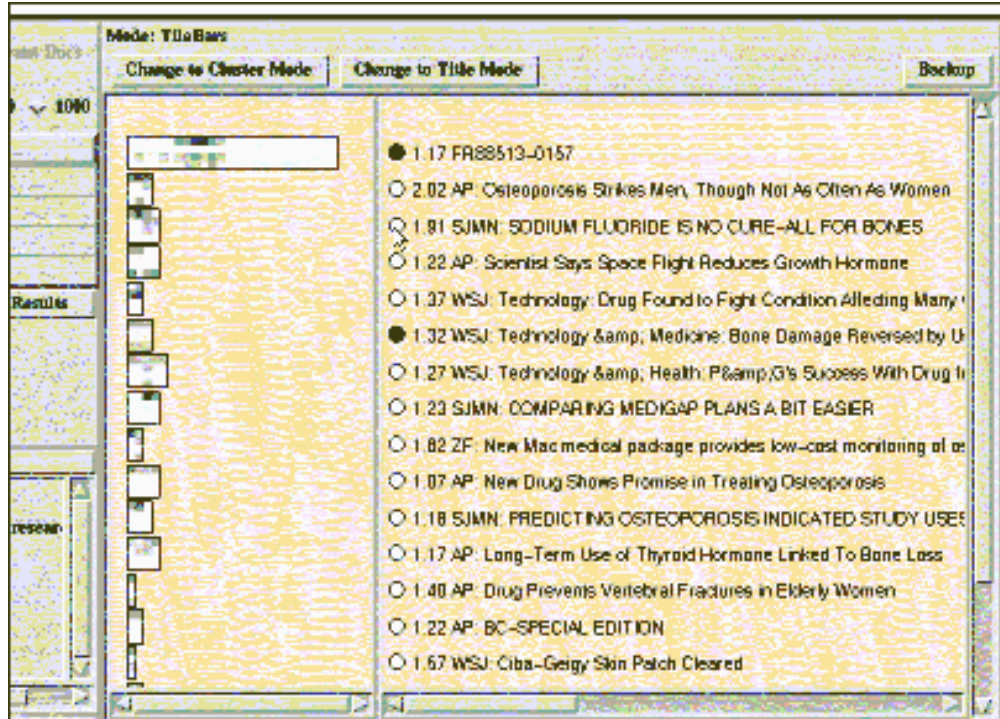


(Γ) Viewing Results in Context Document Surrogates and TileBars

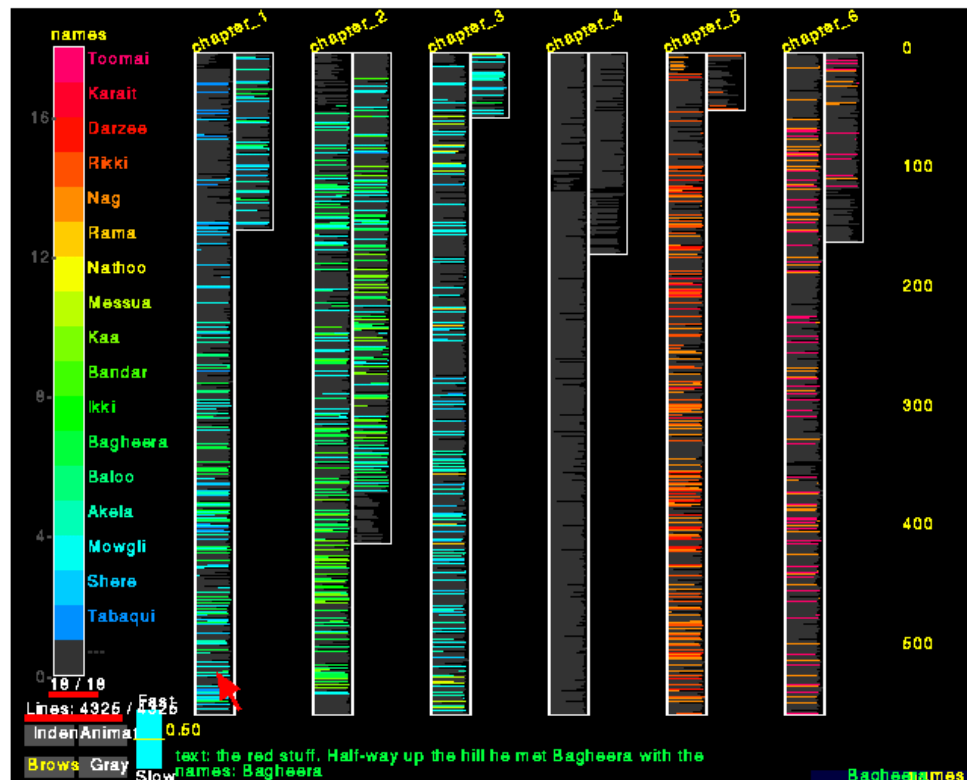
- The user enters a query in a faceted format, with one topic per line.
- A graphical bar is displayed next to the title of each document showing the degree of match for each facet.
 - Bar size: relevance
 - Bar **rows**: one row **per term** (query facet)
 - Bar **columns**: one **per text passage**
 - Bar **color**: number of **hits** (the darker the square the greater the number of hits)
- TileBars thus illustrate at a glance which passages in each article contain which topics - and moreover, how frequently each topic is mentioned (darker squares represent more frequent matches).



(Γ) Viewing Results in Context Document Surrogates and TileBars



(Γ) Viewing Results in Context Document Surrogates and TileBars: SeeSoft



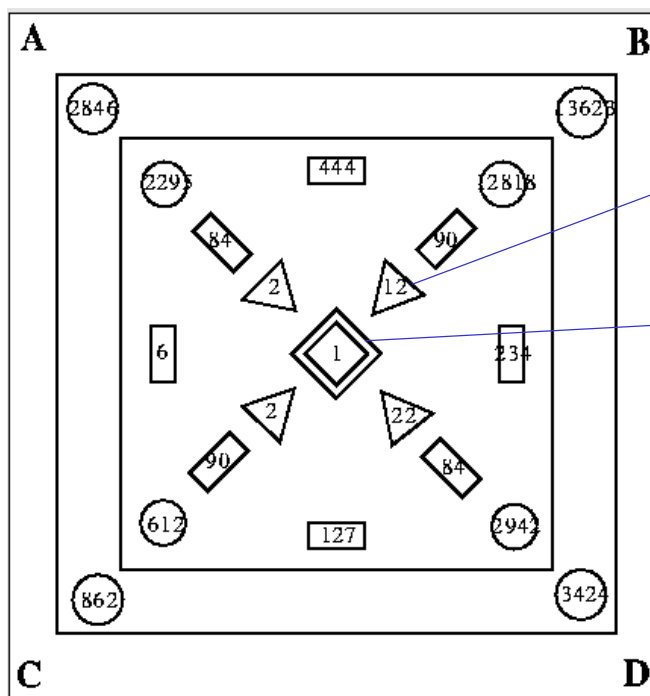


(Γ) Viewing Results in Context Query Term Hits Between Documents

- **Idea:**
 - Rather than showing how query terms appear within individual documents, as is done in KWIC interfaces and TileBars, these systems display an overview or summary of the retrieved documents according to which subset of query terms they contain.
- **Examples:**
 - InfoCrystal, VIBE and Lyberworld



(Γ) Viewing Results in Context Query Term Hits Between Documents: InfoCrystal



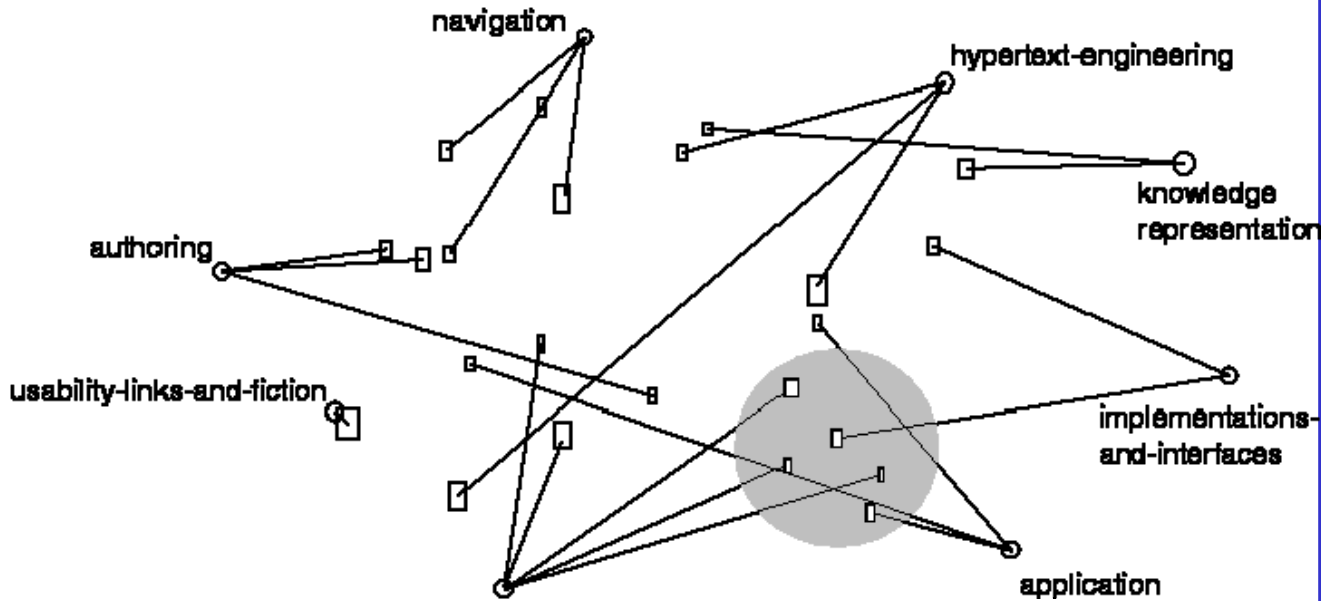
Query terms: A, B, C, D

12 documents contain
all A,B,D

1 document contains
all A,B,C,D



(Γ) Viewing Results in Context Query Term Hits Between Documents: VIBE



(Γ) Viewing Results in Context Clustering of Results (Vivisimo)

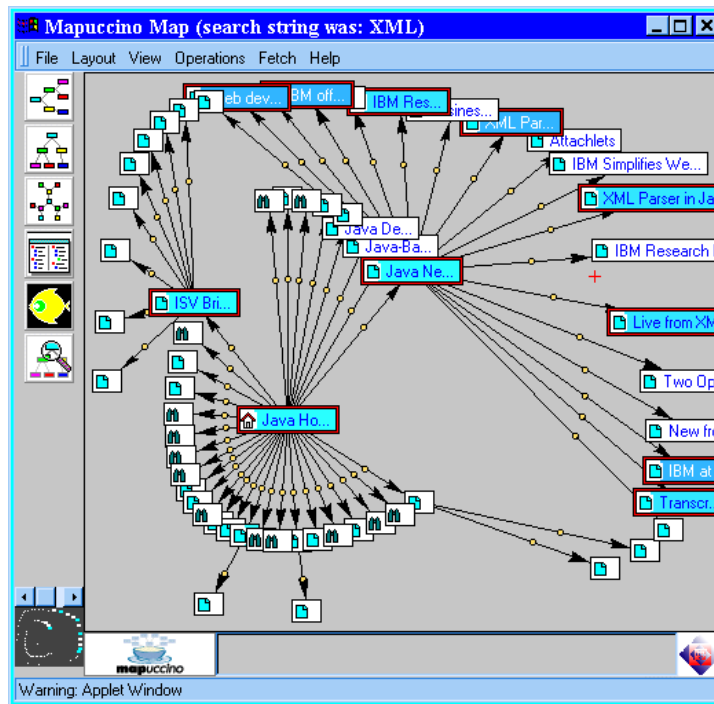
The screenshot shows a web browser displaying search results on the Vivisimo platform. The search term is 'Information Retrieval'. The results are clustered into several groups:

- Information Retrieval (250)**
 - Software (30)
 - Information Retrieval System (26)
 - Processing, Natural Language (15)
 - Research Group (16)
 - Book (15)
 - Baeza-Yates (3)
 - Online Book (2)
 - Management, Indexing (3)
 - Springer (2)
 - Storage and Retrieval (2)
 - Publicly available rate for the same hotel (3)
 - SIGIR (11)
 - Program, Databases (12)
 - Computing (13)
 - Management, Information Retrieval (8)
- Cluster Information Retrieval Group contains 7 documents.**
 - Glasgow Information Retrieval Group** [new window] [frame] [preview]
The **Information Retrieval Group** Congratulations to Prof. Keith van Rijsbergen, who has recently organising the **Information Retrieval** in Context Workshop at SIGIR 2004 ...
URL: ir.dcs.gla.ac.uk - [show in cluster](#)
Sources: [WiseNut 1](#)
 - (UK) University of Sheffield Information Retrieval Group** [new window] [frame] [preview]
The primary research areas of the group include statistical **information retrieval** techniques, multi **retrieval**, and personal **information** management and **retrieval**.
URL: ir.shef.ac.uk - [show in cluster](#)
Sources: [Open Directory 8](#)
 - The Glasgow Information Retrieval Group** [new window] [frame] [preview]
Has a research program aimed at giving better access to multi-media **information**.
URL: ir.dcs.gla.ac.uk - [show in cluster](#)
Sources: [Open Directory 14](#)
 - Retrieval Group Homepage** [new window] [frame] [preview]
... The **Retrieval Group of the Information Access Division** works with industry ... support specific sub-tasks such as cross-language **retrieval** and multimedia **retrieval** ...
URL: www.nlpir.nist.gov - [show in cluster](#)
Sources: [MSN 32](#)
 - Library and Information Science > Information Retrieval in the Yahoo! Directory** [new window]
Yahoo! reviewed these sites and found them related to Library and **Information Science > Informa** of Glasgow - **Information Retrieval Group** - **information** on the resources and people in the Glas

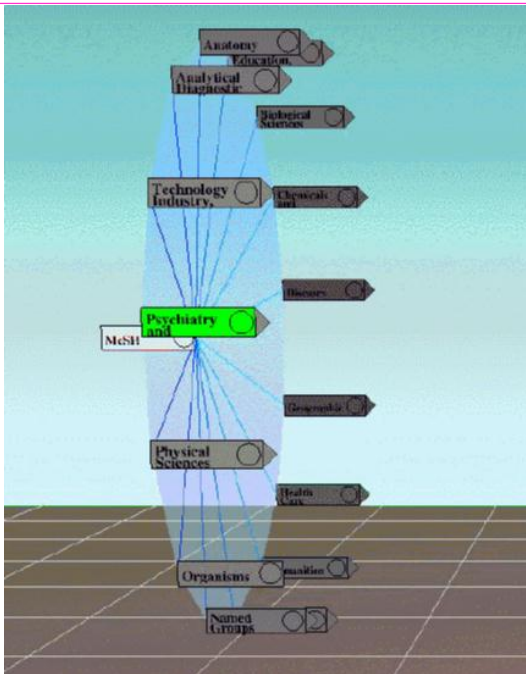
The browser's address bar shows the URL: <http://www.nlpir.nist.gov/>. The taskbar at the bottom shows several open applications, including '463_07b_Clustering.ppt' and 'Lecture6_Clustering.ppt'.



(Γ) Viewing Results in Context Mapuccino: Graphical Depiction of Link Structure



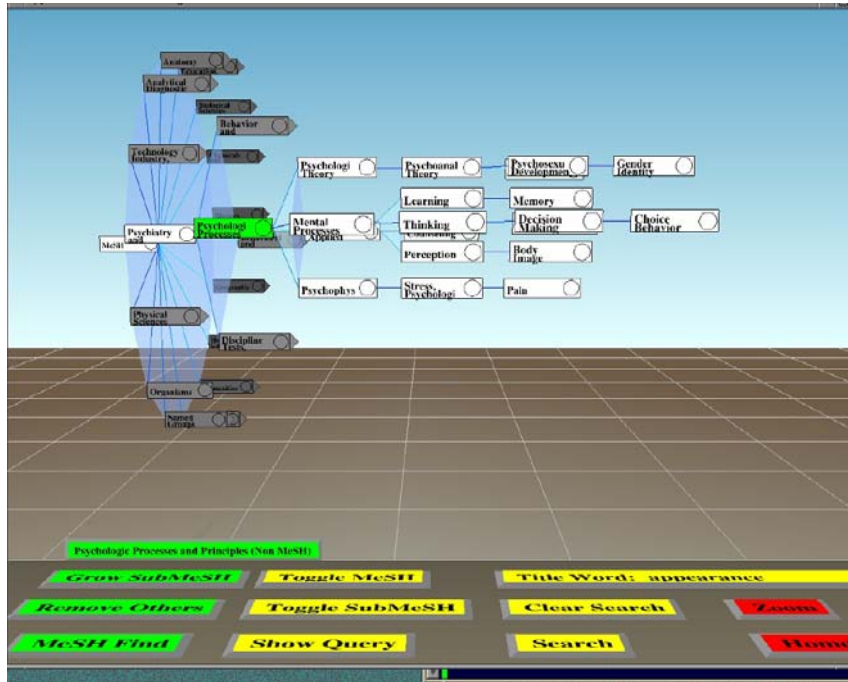
(Γ) Viewing Results in Context Integrating Browsing and Searching: Cat-a-Cone



- The top-level representation of the MeSH hierarchy as a search starting point.



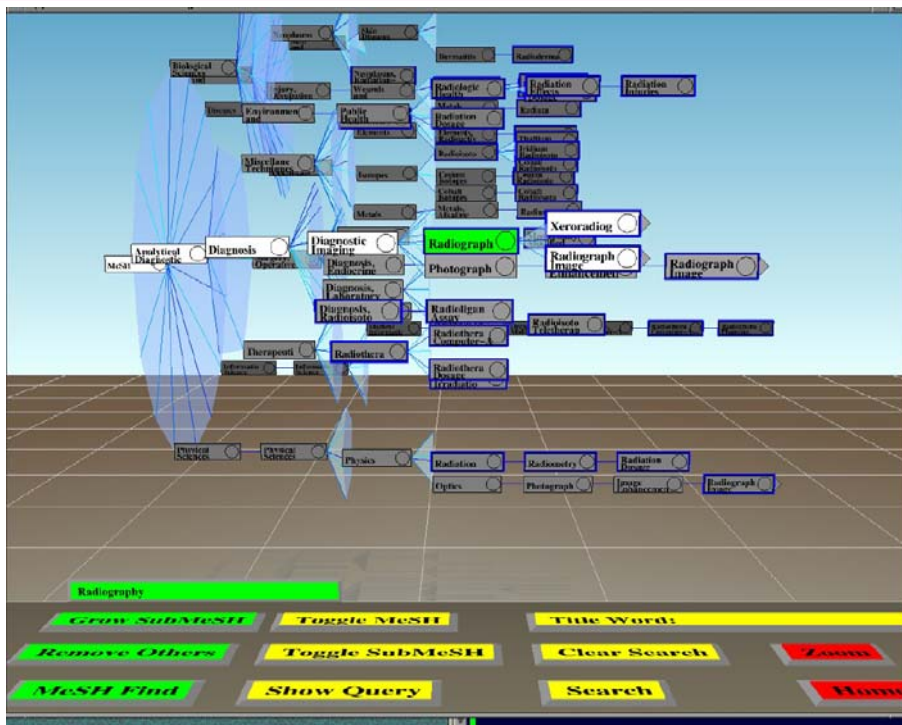
(Γ) Viewing Results in Context Integrating Browsing and Searching: Cat-a-Cone



The expansion of the *Psychiatry* subtree.



(Γ) Viewing Results in Context Integrating Browsing and Searching: Cat-a-Cone



The results of a search over MeSH labels on "radi*". In this subcollection the term occurs under four main headings. Partial matches are shown.



(Γ) Viewing Results in Context Integrating Browsing and Searching: Cat-a-Cone

The screenshot displays a 3D visualization of a concept network. Nodes are represented as rectangular boxes, and they are interconnected by a dense web of blue lines. The nodes are organized into hierarchical levels, with some nodes highlighted in yellow or green. In the foreground, a search result window is open, showing the following information:

TITLE
FEAR OF RECURRENCE,
BREAST-CONSERVING SURGERY,
AND THE TRADE-OFF
HYPOTHESIS

MeSH TERMS
Breast Neoplasms
Fear
Mastectomy, Segmental
Mastectomy, Simple
Neoplasm Recurrence, Local
Prospective Studies
Survival Analysis

ABSTRACT
Fear of recurrence has been at the heart of the controversy between surgeons favoring mastectomy versus those advocating a less radical operation. Breast-conserving surgery is thought to result in a better body image, but patients are expected to worry more about a cancer recurrence because only a small part of the breast is excised. To assess survival rates after breast-conserving intervention, patients were randomized into the National Surgical Adjuvant Breast/Bowel (NSABP) prospective clinical trial (Protocol B06) with three treatment groups: total mastectomy, mastectomy, and breast-conserving surgery.



(Γ) Viewing Results in Context Integrating Browsing and Searching: Cat-a-Cone

The screenshot displays a 3D visualization of a concept network, similar to the one in the previous slide. The nodes are arranged in a hierarchical structure, with some nodes highlighted in green. In the foreground, a search result window is open, showing the following information:

TITLE
FEAR OF RECURRENCE,
BREAST-CONSERVING SURGERY,
AND THE TRADE-OFF
HYPOTHESIS

MeSH TERMS
Breast Neoplasms
Fear
Mastectomy, Segmental
Mastectomy, Simple
Neoplasm Recurrence, Local
Prospective Studies
Survival Analysis

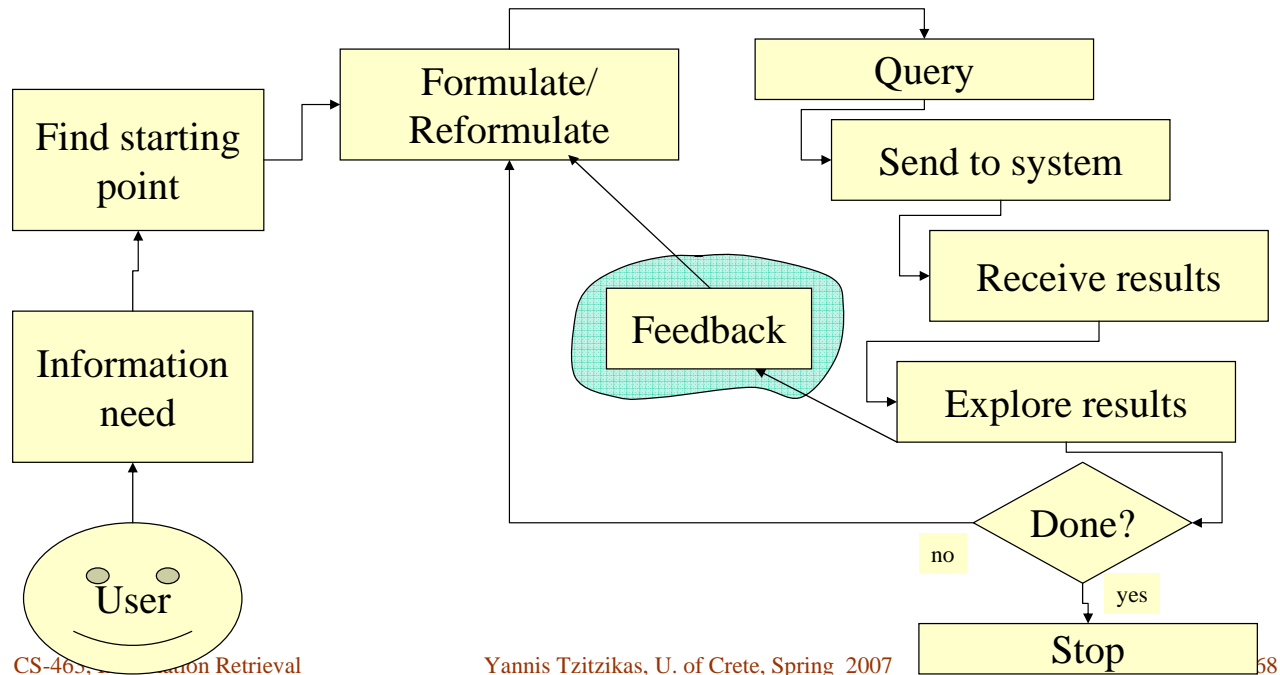
ABSTRACT
Fear of recurrence has been at the heart of the controversy between surgeons favoring mastectomy versus those advocating a less radical operation. Breast-conserving surgery is thought to result in a better body image, but patients are expected to worry more about a cancer recurrence because only a small part of the breast is excised. To assess survival rates after breast-conserving intervention, patients were randomized into the National Surgical Adjuvant Breast/Bowel (NSABP) prospective clinical trial (Protocol B06) with three treatment groups: total mastectomy, mastectomy, and breast-conserving surgery.

At the bottom of the interface, there is a control panel with the following buttons:

- Remove SubMeSH
- Toggle MeSH
- Title Words
- Remove Others
- Toggle SubMeSH
- Clear Search
- Zoom
- MeSH Full
- Show Query
- Search
- Home



(Δ) Επαφές Χρήσης για Ανατροφοδότηση / Βελτίωση της Αναζήτησης

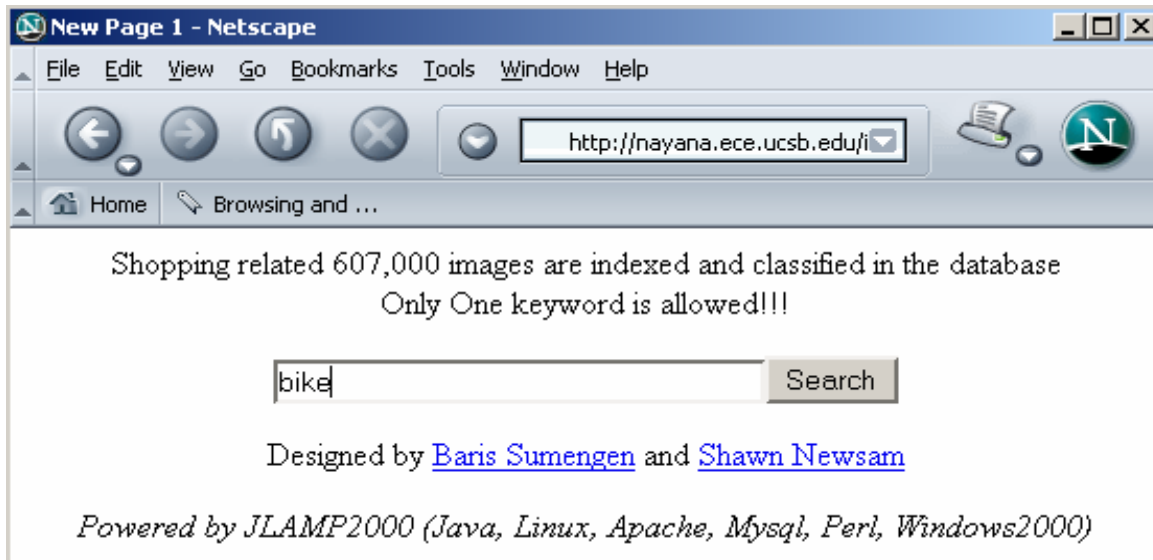


(Δ) Επαφές Χρήσης για Ανατροφοδότηση / Βελτίωση της Αναζήτησης

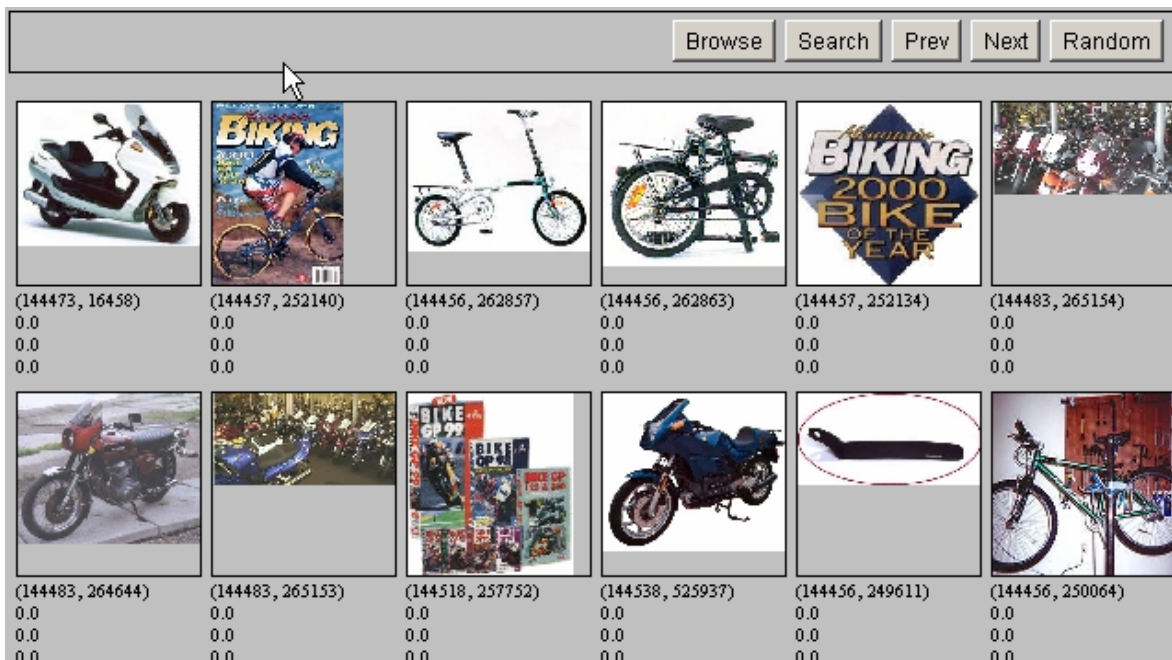
The screenshot shows a web-based search engine interface. At the top, there are search filters and a search bar. Below the search bar, a list of search results is displayed, each with a checkbox and a brief description. One result is selected, and its full text is shown in a large text area below. The text area contains a detailed document snippet, likely a research paper or technical report, discussing various topics related to information retrieval and search engines. The interface includes standard web elements like a search bar, filters, and a list of results.



(Δ) Επαφές Χρήσης για Ανατροφοδότηση / Βελτίωση της Αναζήτησης



(Δ) Επαφές Χρήσης για Ανατροφοδότηση / Βελτίωση της Αναζήτησης





(Δ) Επαφές Χρήσης για Ανατροφοδότηση / Βελτίωση της Αναζήτησης

YOU ARE HERE > [Home](#) > [My InfoSpace](#) > [Meta-Search](#) > Web Search Results

Web Search Results

Your Search

Select:

[Yellow Pages](#) [White Pages](#) [Classifieds](#)

Are you looking for?

[Jacksonville Jaguars](#)

[Jaguar Car](#)

[Black Jaguar](#)

[Jaguar Xk8](#)

[Wild Jaguars](#)

[Jaguare](#)

[Jaguar Accessories](#)

[Jaguar Automobile](#)



Faceted Browsing

- Faceted indexing and searching is being increasingly studied in the literature, and global-scoped applications based on this paradigm have started to emerge. Roughly, the faceted paradigm can bridge the gap between querying and browsing. In general, query services are either too simplistic (e.g. free text queries in Information Retrieval systems or Web search engines), or too sophisticated (e.g. SQL queries, or Semantic Web Queries). Browsing is either too simplistic (e.g. plain Web links) or very application specific (dynamic pages derived by specific application programs). Faceted indexing and searching can bridge this gap and provide effective and efficient general purpose semantic access services.



Faceted Browsing

- The key point, is that within this very short period and with a few clicks efforts, the user not only he finds the desired (or most preferred) choice, but at the same time he gets acquainted with the entire information space, i.e. with the combinations of criteria values that are available and the associated tradeoffs. The efficiency obtained, in terms of user effort, is rather impossible to achieve with the classical querying or plain browsing services



Faceted Browsing

References

- Yannis Tzitzikas, Nicolas Spyrtos, Panos Constantopoulos and Anastasia Analyti, Extended Faceted Taxonomies for Web Catalogs, ERCIM News, Volume 51, October 2002 (special issue: Semantic Web)
- Yannis Tzitzikas, Anastasia Analyti, Nicolas Spyrtos and Panos Constantopoulos, An Algebra for Specifying Compound Terms for Faceted Taxonomies, Procs of the 13th European-Japanese Conference on Information Modelling and Knowledge Bases, EJC'2003,
- Yannis Tzitzikas, R. Launonen, M. Hakkarainen, P. Kohonen, T. Leppanen, E. Simpanen, H. Tornroos, P. Uusitalo and P. Vanska, FASTAXON: A system for FAST (and Faceted) TAXONomy design, Procs of the 23rd International. Conference. on Conceptual Modeling, ER'2004, Shanghai, China, November 2004
- For more see: <http://www.ics.forth.gr/~tzitzik/publications.htm#CTCA>

Projects and Systems

- <http://www.geocities.com/xtopicmaps/FacetedTaxonomy.html>
- FACTS <http://www.ucl.ac.uk/fatks/database.htm>
- FLAMENGO <http://bailando.sims.berkeley.edu/flamenco.html>



Faceted Browsing

Example: Find Hotel (example from www.venere.com)

Accommodation type

- Hotels
- Bed and Breakfasts
- Apartments
- Town House Suites
- Residences
- Farm Houses
- Villas
- Other accommodation

Star rating

- 1 star Hotels
- 2 star Hotels
- 3 star Hotels
- 4 star Hotels
- 5 star Hotels
- 5L star Hotels

Price (EUR|USD|GBP)

per room per night

- from 0 to 75 €
- from 75 € to 125 €
- from 125 € to 225 €
- over 225 €

Area of Rome

- Rome - Historic Center
- Colosseum
- Navona
- Pantheon
- Porta Portese
- Spanish Steps
- Trastevere

Accommodation type

- Hotels
- Bed and Breakfasts
- Apartments
- Town House Suites
- Residences
- Farm Houses
- Villas
- Other accommodation

Star rating

- 1 star Hotels
- 2 star Hotels
- 3 star Hotels
- 4 star Hotels
- 5 star Hotels
- 5L star Hotels

Price (EUR|USD|GBP)

per room per night

- from 0 to 75 €
- from 75 € to 125 €
- from 125 € to 225 €
- over 225 €

Area of Rome

- Rome - Historic Center
- Colosseum
- Navona
- Pantheon
- Porta Portese
- Spanish Steps
- Trastevere

Accommodation type

- Hotels
- Bed and Breakfasts
- Apartments
- Town House Suites
- Residences
- Farm Houses
- Villas
- Other accommodation

Star rating

- 1 star Hotels
- 2 star Hotels
- 3 star Hotels
- 4 star Hotels
- 5 star Hotels
- 5L star Hotels

Price (EUR|USD|GBP)

per room per night

- from 0 to 75 €
- from 75 € to 125 €
- from 125 € to 225 €
- over 225 €

Area of Rome

- Rome - Historic Center
- Colosseum
- Navona
- Pantheon
- Porta Portese
- Spanish Steps
- Trastevere

User selects two price ranges

The categories from stars and areas that do not match fade.

User selects the area Colosseum

The categories from accommodation type, stars and prices that do not match fade.



Faceted Browsing

Example: Faceted DBLP (http://dblp.l3s.de/)

Publication years (Num. hits)

- 1996-2004 (20)
- 2005-2007 (12)

Publication types (Num. hits)

- article(5)
- inproceedings(26)
- proceeding(1)
- www(1)

Venues (Conferences, Journals, ...)

- ER(3)
- SETN(3)
- CAISE(2)
- CoopIS/DOA/ODBASE(2)
- CoopIS/DOA/ODBASE (2)(2)
- AICCSA(1)
- AVI(1)
- CIA(1)
- Data Knowl. Eng.(1)
- Description Logics(1)
- EC-TEL(1)
- ECIR(1)
- EDBT Workshops(1)
- EJC(1)
- FLAIRS Conference(1)
- FoIKS(1)
- More (+10 of total 25)

Authors

- Yannis Tzitzikas(33)
- Nicolas Spyratos(17)
- Carlo Meghini(10)
- Anastasia Analyti(8)
- Panos Constantopoulos(8)
- Jean-Luc Hainaut(2)
- Vassilis Christophides(2)
- Athena Vakali(1)
- Can Türker(1)
- Dimitris Kotzinos(1)
- Dimitris Plexousakis(1)
- Esko Simpanen(1)
- Giorgos Flouris(1)
- Hannu Markkanen(1)
- Hannu Törnroos(1)
- Marco Mesiti(1)
- More (+10 of total 24)

Results

Found 33 entries matching the search term. Showing 33 according to the selection in the

Matches	Authors	Title	Venue	Year	Link
1	Yannis Theoharis, Yannis Tzitzikas	Naming Functions for the Vector Space Model.	ECIR	2007	DB DC Bit
1	Anastasia Analyti, Nicolas Spyratos, Panos Constantopoulos, Yannis Tzitzikas	An algebra for specifying valid compound terms in faceted taxonomies.	Data Knowl. Eng.	2007	DB Bit
1	Yannis Tzitzikas, Jean-Luc Hainaut	On the visualization of large-sized ontologies.	AVI	2006	DB DC Bit
1	Dimitris Kotzinos, Dimitris Plexousakis, Giorgos Flouris, Hannu Markkanen, Nicolas Spyratos, Vassilis Christophides, Yannis Tzitzikas	Triological E-Learning and Emergent Knowledge Artifacts.	EC-TEL	2006	DB DC Bit
1	Carlo Meghini, Nicolas Spyratos, Yannis Tzitzikas	Abduction for Extending Incomplete Information Sources.	SETN	2006	DB DC Bit
1	Yannis Tzitzikas	Revising Faceted Taxonomies and CTCA Expressions.	SETN	2006	DB DC Bit
1	Yannis Tzitzikas	An algebraic method for compressing symbolic data tables.	Intell. Data Anal.	2006	DB Bit



Faceted Browsing

Example: Flamenco Fine Arts Search

Flamenco Fine Arts Search

Images from the Collections of the Fine Arts Museums of San Francisco, Legion of Honor and de Young Museums, <http://www.thinker.org>

Powered by Flame

Save Search History and Settings Return to Search New Search Log

all items in current results

These terms define your current search. Click the to remove a term.

BUILT_PLACES: Bridge

Items 1 to 40 of 431 results

Group by: [Built_Places](#)

1 41 81 121 161 201 241 281 321 361 401

Refine your search within these categories:

MEDIA (group results)

[Book](#) (3) [Painting](#) (1)
[Ceramic](#) (3) [Photograph](#) (9)
[Drawing](#) (26) [Print](#) (384)
[Objects](#) (7)

LOCATION (group results)

[Asia](#) (22) [North America](#) (77)
[Central America](#) (1) [South America](#) (5)
[Europe](#) (324)

OBJECTS (group results)

[Clothing](#) (95) [Musical Instruments](#) (3)
[Containers](#) (10) [Timepieces](#) (4)
[Food and Meals](#) (111) [Vehicles](#) (161)
[Fuel](#) (3) [Weapons](#) (20)
[Lighting](#) (4) [Writing Tools](#) (27)

BUILT_PLACES: [all](#) > [Bridge](#)

ANIMALS AND PLANTS (group results)

[Birds](#) (25) [Mammals, Hoofed](#) (76)



[Untitled](#)
19th century



[The Port of Middel...](#)
18th century



[Landscape with a b...](#)
1793



[\(Landscape with fl...](#)
1763



[\(Landscape with tw...](#)
18th - 19th century



[\(Landscape at sunr...](#)
1763



[Opening of Waterlo...](#)
19th century



[De Roowaensche Kaay](#)
17th century