

Argomenti trattati nel corso

- Inizieremo parlando di **QUERY CON PREFERENZE**
 - Cosa sono, perché sono importanti, i limiti dei DBMS nel trattarle, cosa succede nel caso di più DBMS e su Web, ecc.
- Passeremo poi ad occuparci di **SORGENTI DATI "NON RELAZIONALI"**
 - Sorgenti testuali e multimediali, concetti di similarità
- Vedremo poi **STRUTTURE (INDICI) PER QUERY SU TALI SORGENTI**
 - Come si trovano gli oggetti più vicini/simili a uno dato
- Inoltre tratteremo aspetti di **QUERY REFINEMENT**
 - Cosa fare se l'utente "non è soddisfatto" dei risultati ottenuti

N.B. I lucidi su tali argomenti sono in **inglese**, disponibili sul sito del corso informato PDF

Sistemi Informativi LS

2



Scenario 1: "good" e-commerce solutions

- Siti Web di e-commerce normalmente mettono a disposizione maschere di ricerca per "selezionare i prodotti/offerte" di interesse
 - Ricerca di auto usate, appartamenti, viaggi, aste elettroniche, configurazione di sistemi (e.g., PC), ...
- In questo caso l'obiettivo
 - del **cliente** è trovare qualcosa che soddisfi al meglio le proprie **preferenze**
 - del **sistema** è fornire al cliente le **soluzioni migliori in un tempo ragionevole**
- Il successo di un sito di e-commerce dipende in buona parte da quanto riesce a fornire **risultati di "qualità" senza perdere in efficienza**, e senza imporre all'utente vincoli rigidi nella specifica delle richieste

We're sorry. There are no Flight + Hotel Combo deals available for your search criteria.

What you can do:

- Please change your search criteria by adding more connections and removing airline or time preferences. For international travel, please consider searching for different dates.
- If you would like help purchasing Flight + Hotel Combo for your search criteria, please call 888-TRAVELOCITY (1-888-872-8356) to speak to a Travelocity Agent.

1 Where are you going?

From:

To:

2 When are you going?

Depart: Dec 20

Return: Dec 28

3 What is your preferred number of stops?

Maximum Stops:

4 How many travelers are there?

Adults (13-64): Seniors (65+): Children (2-12):

Nessun risultato!

Lowest Price Trip - 2 Adults

Total Air, Hotel + Tax **\$2070**

Sun, Dec 21

New York-Kennedy, NY (JFK) to Cancun, Mexico (CUN)

4:25pm to 10:59pm

Sun, Dec 28

Cancun, Mexico (CUN) to New York-Kennedy, NY (JFK)

7:10am to 3:12pm

Continental Airlines

1 stop

Continental Airlines

1 stop

Select Different Flights

Sun, Dec 21

Check In

Sun, Dec 28

Check Out

1 Room

Select Hotel Below

Rilassando un criterio di ricerca...

Current Sort: Travelocity Picks

Sort By: Rating

Sort By: Price

Sol Cabanas Del Caribe

Cozumel,

The Sol Cabanas del Caribe is an affordable vacation destination located on Cozumel's north ... [More](#)

Price Includes: 7-night stay in 1 Deluxe Room (other room types), 2 round-trip tickets, and all taxes.

★★★★☆

Travelocity

TotalTrip

\$2070

\$1035 Per person

Select

Camino Real Cancun

Cancun,

The Camino Real is a 381-room, luxury resort surrounded on three sides by the Caribbean Sea ... [More](#)

Price Includes: 7-night stay in 1 Deluxe Beachview King (other room types), 2 round-trip tickets, and all taxes.

★★★★☆

Travelocity

TotalTrip

\$3462

\$1731 Per person

Select

Sistemi Informativi LS

5

CarPrices.com

2004 LINCOLN LS

CLICK TO BUILD YOUR LINCOLN

Research and Price

Find Your Vehicle

Sell Your Vehicle

Anche qui nessun risultato, ma...

Inventory Search for: AUDI A4

Matches Found: 4

Vehicle

Body Style

Price Range

AUDI/A4

All

From \$1000 to \$21,000

Year

Mileage

Exterior Color

From 1996 to 2004

40,000 miles

All

Distance

Start Over

100 miles

See similar vehicles

Sort Results by: Reverse Order

We were not able to find a vehicle matching your exact specifications. Here are the closest matches:

Results 1 - 4 of 4

Clear

Type	Year	Make/Model/Trim	Bodystyle	Price	Mileage	Exterior	Distance		
<input type="checkbox"/>	Used	2001	AUDI/A4/ FrontTrak 1.8T 1.8L I-4 EFI Tbo DOHC	SEDAN	\$20,991.00	46,020	SILVER	90 miles	See Details
<input type="checkbox"/>	Used	2001	AUDI/A4/ FrontTrak 1.8T 1.8L I-4 EFI Tbo DOHC	SEDAN	\$22,991.00	22,423	WHITE	90 miles	See Details
<input type="checkbox"/>	Used	2001	AUDI/A4/ FrontTrak 1.8T 1.8L I-4 EFI Tbo DOHC	SEDAN	\$22,991.00	42,703	BLUE	90 miles	See Details
<input type="checkbox"/>	Used	2001	AUDI/A4/ FrontTrak 1.8T 1.8L I-4 EFI Tbo DOHC	SEDAN	\$25,995.00	15,329	Tan	91 miles	See Details

Compare

...il sistema fornisce comunque delle "buone" risposte!

To compare vehicles, use the checkboxes on the left side of the result table. Choose vehicles to compare, then click the "compare" button.

Sistemi Informativi LS

6



Scenario 2: “good” multimedia solutions

- I sistemi DB di tipo multimediale (MM) vengono usati in una varietà di settori applicativi
 - Design, moda, medicina, scienze biologiche, agenzie stampa, video-on-demand, e-learning, ...
- In tutti questi casi è comune che l'utente ricerchi oggetti MM con certe caratteristiche di “**contenuto**”, specificato tramite un esempio (QBE = Query By Example)
- In questo caso l'obiettivo
 - del **cliente** è trovare qualcosa che sia **semanticamente simile** a quanto cercato
 - del **sistema** è fornire al cliente le **soluzioni migliori in un tempo ragionevole**
- Il successo di un DBMS multimediale dipende in buona parte da quanto riesce a fornire risultati di “qualità” senza perdere in efficienza



Content-based Image Retrieval





Scenario 3: Web search engines

- Un “buon” motore di ricerca permette di trovare le pagine Web più “interessanti” rispetto alla query espressa dall’utente
- In questo caso l’obiettivo
 - del **cliente** è trovare qualcosa che sia **rilevante** rispetto a quanto cercato
 - del **sistema** è fornire al cliente le **soluzioni migliori in un tempo ragionevole**
- Oltre a quali pagine sono gestite dal motore di ricerca, altri fattori influenzano quali risultati vengono restituiti (e in che ordine)



The “Pareto set” query

- Consideriamo un problema di ottimizzazione P in cui si vogliono massimizzare contemporaneamente N funzioni obiettivo $F_1, \dots, F_j, \dots, F_N$ (**ottimizzazione multi-obiettivo**)
- Una soluzione A per P è un insieme di valori per le variabili del problema P, che rispettano tutti i vincoli di P (e.g., $X_i \geq 0$)
- A si dice **Pareto optimal** se non esiste un’altra soluzione B che **domina** A, ovvero che è meglio di A per almeno una funzione obiettivo e non peggio di A rispetto alle altre funzioni obiettivo
$$A \text{ è Pareto optimal} \Leftrightarrow \neg \exists B: (\exists k: F_k(B) > F_k(A) \wedge \forall j: F_j(B) \geq F_j(A))$$
- Il **Pareto set** è l’insieme di tutte le soluzioni Pareto optimal
- ... rivedremo il concetto più avanti



The "Pareto set" query on Google



Google ha cercato **Pareto set** nell'intera rete mondiale. Risultati 1 - 10 di circa 90.0

Pareto set - [Traduci questa pagina]

... **Pareto set**, also **pareto front**, a set of states of objective parameters satisfying the criterion of **Pareto** optimality (see also multi-criteria optimization) ...
ls11-www.cs.uni-dortmund.de/people/beyer/EA-glossary/node103.html - 3k - Copia cache - Pagine simili

VISUALIZATION OF PARETO SET IN FINITE CHOICE PROBLEMS - [Traduci questa pagina]
VISUALIZATION OF **PARETO SET** IN FINITE CHOICE PROBLEMS. Bushenkov VA ... Method for visualization of the **Pareto set** is proposed. Usually, visualization ...
www.ccas.ru/mmes/mmeda/papers/finite.htm - 2k - Copia cache - Pagine simili

Reduced Pareto Set Genetic Algorithm (RPSGAe): Comparative Study

Formato file: PDF/Adobe Acrobat - Versione HTML
Page 1. Reduced **Pareto Set** Genetic Algorithm (RPSGAe): Comparative Study A
Gaspar-Cunha Dept. of Polymer Engineering, University of Minho, Portugal. ...
india.ulb.ac.be/~jknowles/MPSN-II/Gaspar-Cunha.pdf - Pagine simili

Next: Performance Up: P Previous: Pareto optimality

Approximating the Pareto Set: Concepts, Diversity Issues, e

Formato file: PDF/Adobe Acrobat - Versione HTML
Approximating the **Pareto Set**: Concepts, Diversity Issues, and Performance
Marco Laumanns G'unter Rudolph Hans-Paul Schwefel March 15, 1999 Abst
eldorado.uni-dortmund.de/8080/FB4/ stb531/1999/reiheC172/c172.pdf - Pagine

Pareto set

also **pareto front**, a set of states of objective parameters satisfying the criterion of **Pareto** optimality (see also multi-criteria optimization)

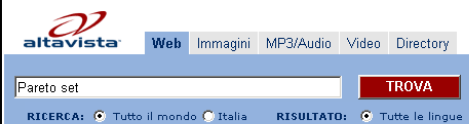
Nice result!

Sistemi Informativi LS

11



The "Pareto set" query on Altavista



Altavista ha trovato 41.074 risultati. Informazioni su

Vilfredo Pareto

... affair broke in France, **Pareto** put his poison to work against the anti-Semitic authorities. **Pareto** also set himself to work, producing a three-volume edition of lecture notes, Cours d'économie ...
cepa-newschool.edu/het/profiles/pareto.htm • Aggiornato nelle ultime 48 ore • E correlate • Traduci
Ulteriori pagine in cepa-newschool.edu

Zipf, Power-law, Pareto - a ranking tutorial

... $a = 1+k$ (where k is the **Pareto** distribution shape parameter). See Appendix discussion of ... Fig. 3 shows such a plot for the same data set of AOL users' ... The relationship is ...
www.hpl.hp.com/shi/papers/ranking/ranking.html • Aggiornato nelle ultime 48 ore • Pagine correlate • Traduci
Ulteriori pagine in www.hpl.hp.com

VISUAL BASIC DATA MINING .NET CONSULTANT

... mining, data warehousing, relational databases, **pareto** analysis, software programming, business ... which returns string data as well or XML record set ...
www.paretoanalysts.com • Aggiornato nelle ultime 48 ore • Pagine correlate •

Vilfredo Pareto, 1848-1923



This is why!

Not exactly what we are looking for!

The Italian economist Vilfredo Pareto was one of the leaders of the Lausanne School and an illustrious member of the "second generation" of the Neoclassical revolution. Although only mildly influential during his lifetime, his "tastes-and-obstacles" approach to general equilibrium theory were resurrected during the great "Paretian Revival" of the 1930s and have guided much of economics since.

Sistemi Informativi LS

12



Scenario 4: new products arrive...

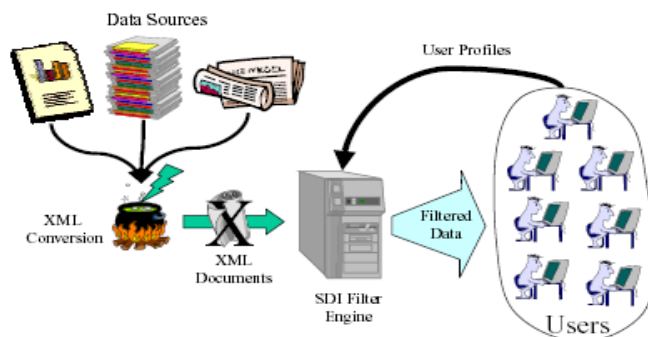
- Sistemi che pubblicizzano i propri prodotti devono evitare di “inondare” gli utenti con informazioni per essi irrilevanti (fenomeno di **information overload**)
- 2 scenari distinti ma con molti aspetti in comune:
 - Sistemi **publish & subscribe**: “*delivering published events to subscribers who expressed interests in such events*” (e.g., sistemi di news)
 - Sistemi di **raccomandazione**: “suggest products to customers” (e.g., electronic catalogs)
- In entrambi i casi l’obiettivo
 - del **cliente** è ricevere qualcosa che soddisfi al meglio le proprie **preferenze** (o “**profilo**”)
 - del **sistema** è fornire al cliente le **soluzioni migliori in un tempo ragionevole**



Selective Dissemination of Information

Information Searching: *selecting (relevant) objects from a “universe”*
1 query at a time, **large universe**

Information Filtering: *removing (unrelevant) objects from a “universe”*
N “queries” (profiles), **medium-small universe of new objects**





Amazon.com recommendations

Shop in Sporting Goods (Beta-What is this?) **amazon.com.** VIEW CART | WISH LIST | YOUR ACCOUNT |

SPORTING GOODS WELCOME PAOLO'S STORE BOOKS APPAREL & ACCESSORIES ELECTRONICS TOYS & GAMES MUSIC HOME & GARDEN SEE STORE

YOUR FAVORITE STORES YOUR RECOMMENDATIONS NEW FOR YOU THE PAGE YOU MADE FRIENDS & FAVORITES

Search All Products: Browse:

SEARCH

All Products or rate items you own

RECOMMENDATIONS

All Products
All Bargains
All Used
[Add Favorites](#)
Books
Camera & Photo
Computer & Video

WELCOME TO Your Recommendations

Hello, Paolo Ciaccia. Explore today's featured recommendations. (If you're not Paolo Ciaccia, [click here.](#))

Book Recommendations

Refactoring

Amazon.com

Your class library works, but could it be better? *Refactoring: Improving the Design of Existing Code* shows how refactoring can make object-oriented code simpler and easier to maintain. Today refactoring requires considerable design know-how, but once tools become available, all programmers should...

[Read more](#) [\(Why was I recommended this?\)](#) [More Book Recommendations](#)

Personalizzate!

Perché?

Sistemi Informativi LS

15



Amazon.com recommendations: why?

Why was I recommended this?

[Close window](#)

We recommended...

Refactoring
by Martin Fowler (Author), et al
List Price: \$49.99
Price: **\$49.99**
Used & new from \$36.99
[I own it](#) [Not interested](#)

Because you purchased or rated...

...una delle possibili ragioni

Purchased or Rated Items	Not Rated	Dislike it <	love it!
Transaction Processing by Jim Gray, Andreas Reuter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	?	1	2 3 4 5
<input checked="" type="checkbox"/> Use for Recommendations			

[Save changes](#)

Customers who bought the item above also bought:

...un'altra possibile ragione

Transactional Information Systems
by Gerhard Weikum, Gottfried Vossen

[More like this](#)

Principles of Transaction Processing
by Philip A. Bernstein, Eric Newcomer (Contributor)

[More like this](#)

Are recommended items not quite on target? [Improve your Recommendations.](#)

16