UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO

SWAP researchgroup

Semantic
Web
Access and
Personalization
research group
http://www.di.uniba.it/~swap

Recommender Systems

Giovanni Semeraro

semeraro@di.uniba.it

Marco de Gemmis, Ph.D

degemmis@di.uniba.it

Pasquale Lops, Ph.D

lops@di.uniba.it

Information Overload

Problems…

- Explosion of irrelevant, unclear, inaccurate information
- Users overloaded with a large amount of information impossible to absorb

…and consequences

- Searching is time consuming
- Need for intelligent solutions able to support users in finding documents according to their interests

How much information?

- Print, film, magnetic and optical storage media produced between 3.4 and 5.4 Exabytes of unique information in 2002
  - ✓ 92% stored on magnetic media, mostly hard disks
  - ✓ 500–800 MB per person each year
- Information flow through electronic channels – telephone, radio, TV, and the Internet – contained almost 18 Exabytes of new information in 2002 (3 ½ more than recorded in storage media), 281 Exabytes in 2007 (45 GB per person)

Medium	Terabytes
Radio	3,488
Television	68,955
Telephone	17,300,000
Internet	532,897
TOTAL	17,905,340

Medium	Terabytes
Surface Web	167
Hidden Web	91,850
E-mail	440,606
Inst. Messaging	274
TOTAL	532,897

Lyman, Peter and Hal R. Varian, *How Much Information*, School of Information Management and Systems, University of California at Berkeley, 2003. URL: http://www.sims.berkeley.edu/how-much-info-2003. Last access: May 23rd, 2007.

Big Data

How much information?

Inertial sensors (accelerometers, compasses, …)

- 16M devices sold in 2006 (source: iSuppli 2010)
- 436 M devices (+2700%) sold in 2009
- 6B $ revenues in 2014 (source: IC Insights 2010)

64M tweets, 90M photos on Facebook, 50B SMS per day in 2010

The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide Information Growth Through 2011. IDC White paper, March 2008. URL: http://www.idc.com.

How big is an Exabyte?

Kilobyte (KB)	1,000 bytes or 10^3 bytes. 2 Kilobytes: A Typewritten page. 100 Kilobytes: A low-resolution photograph.
Megabyte (MB)	1,000,000 bytes or 10^6 bytes. 1 Megabyte: A small novel or a 3.5 inch floppy disk. 2 Megabytes: A high-resolution photograph. 5 Megabytes: The complete works of Shakespeare. 10 Megabytes: A minute of high-fidelity sound. 100 Megabytes: 1 meter of shelved books.
Gigabyte (GB)	1,000,000,000 bytes or 10^9 bytes. 1 Gigabyte: a pickup truck filled with books. 20 Gigabytes: A good collection of the works of Beethoven. 100 Gigabytes: A library floor of academic journals.
Terabyte (TB)	1,000,000,000,000 bytes or 10^{12} bytes 1 Terabyte: 50,000 trees made into paper and printed. 2 Terabytes: An academic research library.
Petabyte (PB)	1,000,000,000,000,000 bytes or 10^{15} bytes. 2 Petabytes: All U.S. academic research libraries. 20 Petabytes: Production of hard-disk drives in 1995. 200 Petabytes: All printed material.
Exabyte (EB)	1,000,000,000,000,000,000 bytes or 10^{18} bytes. 2 Exabytes: Total volume of information generated in 1999. 5 Exabytes: All words ever spoken by human beings.

TABLE 1.1: How Big is an Exabyte? Source: Many of these examples were taken from Roy Williams "Data Powers of Ten" web page at Caltech.

IN 60 SECONDS...
1 NEW DEFINITION IS ADDED ON urbandictionary
1,600+ READS ON Scribd
13,000+ HOURS MUSIC STREAMING ON PANDORA
12,000+ NEW ADS POSTED ON craigslist
170,000+ MINUTES VOICE CALLS ON skype
98,000+ TWEETS
20,000+ NEW POSTS ON tumblr.
LARGEST
320+ NEW twitter ACCOUNTS
100+ NEW LinkedIn ACCOUNTS
13,000+ iPhone APPLICATIONS DOWNLOADED
QUESTIONS ASKED ON THE INTERNET...
Answers.com
Yahoo!
600+ NEW VIDEOS
6,600+ NEW flickr
50+ WordPress DOWNLOADS
25+ HOURS TOTAL DURATION
70+ DOMAINS REGISTERED
60+ NEW BLOGS
168 MILLION EMAILS ARE SENT
694,445 SEARCH QUERIES
1,700+ Firefox DOWNLOADS
695,000+ facebook STATUS UPDATES
125+ PLU DOW
1,500+ BLOG POSTS
79,364 WALL POSTS
510,040 COMMENTS
Google
GO-Globe.com

Meno's Paradox

Search is not so simple as it might seem

Meno: And how will you enquire, Socrates, into that which you do not know? What will you put forth as the subject of enquiry? And if you find what you want, how will you ever know that this is the thing which you did not know?

Socrates: I know, Meno, what you mean; but just see what a tiresome dispute you are introducing. You argue that man cannot enquire either about that which he knows, or about that which he does not know; for if he knows, he has no need to enquire; and if not, he cannot; for he does not know the very subject about which he is to enquire.

Plato Meno 80d-81a

http://www.gutenberg.org/etext/1643

Meno's question at our times

The Dictionary Problem

With Information Retrieval & Filtering systems, aka «modern *Delphic Oracle*»:

"How to discover the *magic words* that will connect us to the precious boxes the information we are seeking is hidden into?"

The focus of many studies has been to point out methods for text processing to enable a system to make connections between the terminology of the user request (implicit/explicit) and related terminology in the information needed

⇩

Need for semantics for text interpretation

⇩

From basic Retrieval & Filtering systems
towards
Intelligent Information Access by means of NLP techniques

Intelligent Information Access

Infusing knowledge/semantics into (traditional) software and services

"technology should adapt to people rather than vice versa"
(M. Weiser & J. Seely Brown, Xerox PARC, 1990s)

Information Access Strategies

Process	Information Need	Information Sources
Information Retrieval	Dynamic-Specific	Stable-Unstructured
Information Filtering	Stable-Specific	Dynamic-Unstructured
Text Mining	Stable-Specific	Stable
Database Access	Dynamic-Specific	Stable-Structured
Exploration	Broad	Varied

Information Retrieval [Baeza-Yates and Ribeiro-Neto 1999]

- "Information Retrieval (IR) deals with the representation, storage, organization of, and access to information items"
- "…the user must first translate this information need into a *query* which can be processed by a search engine (or IR system)".
- "Given the user query, the key goal of an IR system is to retrieve information which might be useful or relevant to the user. The emphasis is on the retrieval of *information* as opposed to the retrieval of *data*".

Information Access Strategies

Process	Information Need	Information Sources
Information Retrieval	Dynamic-Specific	Stable-Unstructured
Information Filtering	Stable-Specific	Dynamic-Unstructured
Text Mining	Stable-Specific	Stable
Database Access	Dynamic-Specific	Stable-Structured
Exploration	Broad	Varied

Information Filtering [Hanani et al. 2001]

- "The aim of Information Filtering is to expose users to only the information that is relevant to them. Some examples of filtering applications are: filters for search results on the internet,… e-mail filters based on *personal profiles*, … filters for e-commerce applications that address products and promotions to potential customers only…"
- "There are many systems of widely varying philosophies, but all share the goal of automatically directing the most valuable information to users in accordance with their user model…"

Comparing IR and IF

Common Mechanisms

- **Representation:** Both the user's information need – query or profile – and the document set must be represented for *comparison*
- **Comparison:** String matching? Concept matching? User–User Correlation? Item–Item correlation?
- **Feedback:** To improve the performance of the IR/IF system, a feedback mechanism is usually incorporated.

Comparing IR and IF

- Where IR is concerned with the *collection and organization* of texts, IF is concerned with the *distribution* of texts to groups or individuals.

❷ Where IR is typically concerned with the selection of texts from a relatively *static database*, IF is mainly concerned with the selection or elimination of texts from a *dynamic datastream*.

❸ Where IR is concerned with responding to the user's interaction with texts within a *single information-seeking episode*, IF is concerned with long-term changes over *a series of information-seeking episodes*.

[Belkin and Croft 1992]

Comparing IR and IF

Parameters	Information Retrieval	Information Filtering
Representation of Information Needs	queries	profiles
Goal	selection of relevant items for query	filtering out irrelevant items or collecting items
Frequency of use	ad hoc use one time user	repetitive use long term users
Type of Users	Not known to the system	"Profiled"
Database	(relatively) static	very large dynamic

Table adapted from [Hanani et al. 2001]

Inside Recommender Systems: Definitions & Paradigms

Some Definitions

- Recommender Systems have the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options [Burke, 2002]
- Recommender Systems provide personalized suggestions about items that the user might find interesting, by matching items to user profiles or groups [Kangas, 2001]
- Recommendation process: Given a large set of items and a description of the user's needs, the goal is to present a small set of the items that are suited to the user needs

[Burke 2002] R. Burke. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction* 12(4):331–370, 2002.

[Kangas 2001] Sonja Kangas. *Collaborative Filtering and Recommendation Systems*. Technical Report TTE4-2001-35. VTT Information Technology, Espoo, Finland, 2001.

IF systems architecture

A classification of IF systems

[Hanani et al. 2001]

A classification of IF systems

[Hanani et al. 2001]

What do RS do exactly?

1. Predict how much you may like a certain product/service
2. Compose a list of **N** best items for you
3. Compose a list of **N** best users for a certain product/service
4. Explain why these items are recommended to you
5. Adjust the prediction and recommendation based on your feedback and other people

	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8	I_9
U_1				1		5			4
U_2	4			2		5			
U_3					4			5	
U_4	5			2			4		
A	1	3	1	3	1	4	5	8	

Collaborative / Social Filtering (CF)

1. *Word of Mouth / "Wisdom of the Crowds"*
2. Makes use of a database of user preferences in order to:
   - ✓ find users with *similar* interests (*neighborhood* computation)
   - ✓ predict whether an *unseen* information item is likely to be of interest for a user (based on how *other* users rated that item)
3. Widely adopted in Recommender Systems [Resnick and Varian 1997, Linden et al. 2003]

[Resnick and Varian 1997] Resnick, P. and H. Varian. Recommender Systems. *Communications of the ACM* 40(3):56–58, 1997.
[Linden et al. 2003] Linden, G., B. Smith, and J. York. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing* 7(1):76–80, 2003.

Item-to-Item CF

- Adopted by Amazon.com [Linden et al. 2003]
  - ✓ Amazon.com has more that 30 million customers and several million catalog items
  - ✓ scales to massive datasets and produces high-quality real-time recommendations since #items << #users
- Rather than starting by matching the user to similar customers (impossible for a new user), exploits a similar-items table to find items that customers tend to purchase together
  - ✓ Acceptable performance by creating the (expensive) similar-items table offline

The algorithm:

- ✓ finds items similar to each of the active user's purchases and ratings
- ✓ aggregates those items and recommends the *most popular* or *correlated* items

Item-to-Item CF: similarity computation

	I_1	I_2	I_3	I_4	I_5	I_6
I_1	–	w_{12}	w_{13}	…		w_{16}
I_2	w_{21}	–	…	…		
I_3			–			
I_4		…		–		
I_5			…		–	
I_6		…				–

w_{ij} computed in different ways:

- ✓ #users co-rated I_i and I_j
- ✓ cosine similarity between item vectors I_i and I_i from user-item matrix
- ✓ …

Amazon Recommender System@work

amazon.com | Marcol...'s Amazon.com | See All 41 Product Categories | Your Account | Cart | Your Lists | Help | NEW

Your Browsing History | **Recommended For You** | Rate These Items | Improve Your Recommendations | Your Profile | Learn More

Search Amazon.com GO | Find Gifts | A9 Web Search GO

Marcolino's Amazon.com™ › Recommended for You
(If you're not Marcolino, click here.)

Recommendations by Category

- Apparel & Accessories
- Baby
- Beauty
- Books
- Camera & Photo
- Computers & PC Hardware
- DVD
- Electronics
- Gourmet Food
- Grocery
- Health & Personal Care
- Home Improvement
- Industrial & Scientific
- Jewelry & Watches
- Kitchen & Housewares
- Magazine Subscriptions
- Music
- Outdoor Living
- Software

These recommendations are based on items you own and more.

view: **All** | New Releases | Coming Soon

More results

1.

Pirates of the Caribbean - Dead Man's Chest (Two-Disc Collector's Edition)
DVD ~ Orlando Bloom (Dec 5, 2006)
Average Customer Review: ★★★★☆ (672)
In Stock
List Price: ~~$34.99~~
Price: $24.99
Add to cart | Add to Wish List
121 used & new from $9.99
Rate it
ted **Cars (Widescreen Edition)** and more (Fix this)

) Mindstorms NXT: The Mayan Adventure (Technology in Action)
n Kelly (Dec 11, 2006)
age Customer Review: ★★★★½ (14)
ock
Price: ~~$24.99~~
: $15.74
Add to cart | Add to Wish List
ed & new from $11.02
Rate it

Amazon Recommender System@work

Marcolino, Welcome to Your Amazon.com™ (If you're not Marcolino, click here.)

Today's Recommendations For You 1 2 3

Here's a daily sample of items recommended for you. Click here to **see all recommendations**.

The Da Vinci Code (DVD)

Star Wars Episode VI - Return of the Jedi (DVD)

X-Men - The Last Stand (Widescreen Edition) (DVD)

Underworld - Evolution (DVD)

Star Wars - Episode III, Revenge of the Sith (DVD)

Maximum Lego NXT: Building Robots with Java Brains
Paperback by Brian Bagnall

LEGO Mindstorms NXT

Your Recent Shopping
Recently Viewed Items (2)
Your Shopping Cart (2)
Open & Recently Shipped Orders

Your Lists
Your Wish List
Your Gift List
Your Shopping List

Your Community
Your Amazon Friends
Your Interesting People
Your Reminders
Your Profile

Amazon Recommender System@work

amazon.com | Marcol...'s Amazon.com | Books | See All 41 Product Categories | Your Account | Cart | Your Lists | Help

Advanced Search | Browse Subjects | Bestsellers | The New York Times® Best Sellers | New & Future Releases | Libros En Español | Magazines | Sell Your Stuff | Bargain Books | Textbooks

Search Books GO | Find Gifts | Web Search GO

Prime
Sign up for Amazon Prime today

To get this item by **Tuesday**, May 22 order within 3hr 52min.
Buy now and get FREE Two-Day Shipping
(Learn more)

LEGO Mindstorms NXT: The Mayan Adventure (Technology in Action) (Paperback)
by Jim Kelly (Author)
(14 customer reviews)

List Price: ~~$24.99~~
Price: **$15.74** & eligible for **FREE Super Saver Shipping** on orders over $25. Details
You Save: $9.25 (37%)

Quantity: 1
Add to Shopping Cart
or
Sign in to turn on 1-Click ordering.

More Buying Choices
50 used & new from $11.02

Better Together

Buy this book with LEGO MINDSTORMS NXT Hacker's Guide by Dave Prochnow today!

+

Buy Together Today: $32.21

Buy both now!

Recomm. based on item similarity

Customers who bought this item also bought

- LEGO MINDSTORMS NXT Hacker's Guide by Dave Prochnow
- LEGO Mindstorms NXT by LEGO
- Maximum Lego NXT: Building Robots with Java Brains by Brian Bagnall
- Extreme NXT: Extending the LEGO Mindstorms NXT to the Next Level (Technology in Action) by Michael Gasperi
- Building Robots With Lego Mindstorms : The Ultimate Tool for Mindstorms Maniacs by Mario Ferrari

Explore similar items : Books (43) Toys (3) Electronics (1) Software (1) Magazines (1)

Collaborative Filtering Issues

- Pros: 👍
  - ✓ well-understood,
  - ✓ works well in some domains
  - ✓ no knowledge engineering required
- Cons: 👎
  - ✓ requires user community
  - ✓ sparsity problems
  - ✓ no integration of other knowledge sources
  - ✓ no explanation of results

Content-based Filtering

Information Source

User Profile

User profile compared against items for *relevance* computation

Items recommended to the user

Target User

Content-based Recommender Systems (CBRS)

- Recommend an item to a user based upon a description of the item and a profile of the user's interests
- Implement strategies for:
  - ✓ representing the items that may be recommended
  - ✓ creating a profile of the user that describes the types of items the user likes/dislikes
  - ✓ comparing the user profile to some reference characteristics with the aim to predict whether the user is interested in an unseen item

[Lops11] P. Lops, M. de Gemmis, and G. Semeraro. Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor (Eds.), *Recommender Systems Handbook,* Springer, 73-105, 2011.

[Pazzani07] Pazzani, M. J., & Billsus, D. Content-Based Recommendation Systems. *The Adaptive Web*. Lecture Notes in Computer Science vol. 4321, 325-341, 2007.

Content-based Filtering: characteristics

1. Each user is assumed to operate independently
2. Items are represented by some features
   - ✓ Movies: actors, director, plot, …
3. The profile is often created and updated automatically in response to feedback on the desirability of items that have been presented to the user
   - ✓ Machine Learning for automated inference
   - ✓ Relevance judgment on items, e.g. ratings
   - ✓ Training on rated items → user profile
4. Filtering based on the comparison between the *content (features)* of the items and the user preferences as defined in the user profile
   - ✓ Keyword-based representation for content and profiles → string matching or text similarity

A LinkedIn profile

www.linkedin.com/home

SWAP - Semantic W... | Google Traduttore | CSA Archiviazione e ... | Progetto Casa Apert... | JCR-Web 4.5 Welco... | chair options | Co

LinkedIn Account Type: Basic | Upgrade

Home | Profile | Contacts | Groups | Jobs | Inbox 36 | Companies | News | More | People

Wake up, invest in Brazil - The Next Leading Investment Market. Brazil's Fores

Share an update

Attach a link | Share

LinkedIn Today: See all Top Headlines for You

There Is No Invisible Hand

Best and Worst Jobs 2012 - Table

Job Title
Software Engineer
Actuary
Human Resources Manager
Dental Hygienist
Financial Planner
Audiologist
Occupational Therapist

4 Reasons Pinterest Wins with Women (And Facebook Loses)

All Updates • Shares • Groups • More

Recent • Top

Cosimo Palmisano is now connected to Davide Angotta, Marketing Director at RIM BlackBerry and Alessandro Bonaita, Head of CRM Analytics at RCS MediaGroup

Send a message • 3 minutes ago

Gianni Sebastiano has an updated profile (Expertise)

Like • Comment • Send a message • 22 minutes ago

A LinkedIn profile

Pasquale Lops

Assistant Professor at University of Bari

Bari Area, Italy | Higher Education

Current	**Assistant Professor** at **University of Bari "Aldo Moro"**
Past	Phd student at University of Bari "Aldo Moro" Research student at Fraunhofer Integrated Publications and Informations Systems Institute, Darmstadt, Germany Owner and employee at Sword ICT Srl see all
Education	Università degli Studi di Bari
Connections	179 connections
Websites	Personal Website
Twitter	pasqualelops
Public Profile	http://it.linkedin.com/in/pasqualelops

Summary

Specialties

Recommender Systems,
Information Filtering,
Personalization,
User Modeling,
Social and Semantic Web,
Social Network Analysis,
Linked Data,
Machine Learning.

A LinkedIn profile

Pasquale Lops

Assistant Professor at University of Bari

Bari Area, Italy | Higher Education

Groups and Associations:

- Alumni Università degli Studi di Bari
- Artificial Intelligence
- Associazione Italiana per l'Intelligenza Artifici
- DIB - Dipartimento di Informatica di Bari
- Framework Programme Seven (FP7)
- Hybrid Smart Cities - Technology Controlling
- Mining Complex Patterns 2011
- RecSys 2011
- RecSys'08

Experience

Assistant Professor
University of Bari "Aldo Moro"
January 2004 – Present (8 years 4 months)

Phd student
University of Bari "Aldo Moro"
November 2001 – March 2005 (3 years 5 months)
PhD in Computer Science, University of Bari, Italy.
Thesis: Hybrid Recommendation Techniques based on User Profiles

Research student
Fraunhofer Integrated Publications and Informations Systems Institute, Darmstadt, Germany
July 2004 – September 2004 (3 months)

Owner and employee
Sword ICT Srl
October 1998 – January 2003 (4 years 4 months)

Consultant
OMAR SpA
1999 – 2001 (2 years)

Skills & Expertise

Recommender Systems | Machine Learning | Personalization | User Modeling | Social Networking | Semantic Web | RDF | Information Retrieval | Linked Data | Natural Language Processing | Collaborative Filtering | Semantic Technologies | Artificial Intelligence

Recommendations in LinkedIn

People You May Know

Elisa Chiabrando, Ph.D.Student at Doctoral School of Sciences and
Connect

Francesco Tangari, M.Sc student in Computer Science at University of
Connect

Francesco Sciacovelli, Planer manutenzione impianti e fabbricati
Connect

See more »

Companies You May Want To Follow

L3S
ontoprise
Numenta
SALFORD SYSTEMS
MONDECA

Feedback | See more »

Jobs You May Be Interested In *beta*

Internship Pumps, Valves and Systems...
GE - Bari

Tenured Faculty Positions
IMT Lucca Institute for Advanced Studies - Prato Area, Italy

Coordinatore/Consulente Formazione...
Michelangelo Formazione - Somma Vesuviana

Feedback | See more »

Get hired faster with Job Seeker Premium

Groups You May Like

SWEL: Ontologies, Semantic Web and Social Semantic Web for E-Learning
Join - Conference Group

Semanticweb.com
Join - Professional Group

GATE
Join - Professional Group

Feedback | See more »

Viewers of this profile also viewed...

Nicola Fanizzi
Assistant Professor

Corrado Mencar
Assistant Professor

Pierpaolo Basile
Co-founder at QuestionCube

Gianni Semeraro
Assoc. Prof. at Univ. of Bari

Nicola Di Mauro
Assistant Professor at University of...

Filippo Lanubile
Associate Professor at University of...

Michelangelo Ceci
Assistant Professor at Università degli...

Oriana Licchelli
Associate Professor of Computer Science...

Content-based Filtering Issues

- Pros: 👍
  - ✓ works well in some domains
  - ✓ no community required
  - ✓ comparison between items possible
  - ✓ Transparency: explanation is possible
- Cons: 👎
  - ✓ Limited content analysis: keywords alone may not be sufficient to judge quality/relevance of a document
  - ✓ overspecialization
  - ✓ Content-descriptions necessary
  - ✓ cold start for new users

Drawbacks of CBRS
Limited Content Analytics

- No suitable suggestions if the analyzed content does not contain enough information to discriminate items the user likes from items the user does not like
- Content must be encoded as meaningful **features**
  - ✓ automatic/manually assignment of features to items might be insufficient to define distinguishing aspects of items (necessary for the elicitation of user interests)
  - ✓ keywords not appropriate for representing content, due to polysemy, synonymy, multi-word concepts (homography, homophony) – "Sator arepo eccetera" (Eco, 2007)

Limited Content Analytics

Keyword-based Profiles

doc1

AI is a branch of computer science

doc2

the 2011 International Joint Conference on **Artificial Intelligence** will be held in Spain

doc3

apple launches a new product…

USER PROFILE

artificial	0.02
intelligence	0.01
apple	0.13
AI	0.15
…	

MULTI-WORD CONCEPTS

Limited Content Analytics

Keyword-based Profiles

doc1

AI is a branch of computer science

doc2

the 2011 International Joint Conference on **Artificial Intelligence** will be held in Spain

doc3

apple launches a new product…

USER PROFILE

artificial	0.02
intelligence	0.01
apple	0.13
AI	0.15
…	

SYNONYMY

Limited Content Analytics
Keyword-based Profiles

doc1
AI is a branch of computer science

doc2
the 2011 International Joint Conference on **Artificial Intelligence** will be held in Spain

doc3
apple launches a new product…

USER PROFILE

artificial	0.02
intelligence	0.01
apple	0.13
AI	0.15
…	

POLYSEMY

Advanced methods are needed for the elicitation of user interests

Recommender Systems that you probably use without your knowledge aka “a tua insaputa”

Twitter

Home | Connect | Discover | Search

Pasquale Lops

@pasqualelops

Italy · http://www.di.uniba.it/~lops/lops.html

Edit your profile

5 TWEETS

8 FOLLOWING

47 FOLLOWERS

- Tweets
- **Following**
- Followers
- Favorites
- Lists

Similar to you

- **Cataldo Musto** @cataldomusto Following
- **ContextAware IR & RS** @CaRRws Follow
- **SasWeb2011** @sasWeb2011 Follow

twitter

© 2012 Twitter About Help Terms Privacy

Following

Ernesto W. De Luca @ernestowdeluca — Following

Francesco Ricci @fmr59 — Following

Hyptertext 2010 @HT2010 — Following

Hypertext 2010: The 21st ACM Conference on Hypertext and Hypermedia Toronto, June 13-16, 2010

ECML PKDD 2010 @ecmlpkdd10 — Following

News for the ECML PKDD 2010 conference, which will take place in Barcelona, Sep 20-24 2010

Giovanni Semeraro @semeraro_g — Following

Recommendations in Twitter

Home | Connect | Discover | Enter a #hashtag or keyword

Stories
Activity
Who to follow
Find friends
Browse categories

Italy trends · Change
#cambia
#maroni
Rosy Mauro
#bergamo
Sumatra
Oregon
Matt Groening
Bossi
Luna Rossa
Belsito

twitter
© 2012 Twitter About Help Terms Privacy
Blog Status Apps Resources Jobs
Advertisers Businesses Media Developers

Who to follow

Twitter accounts suggested for you based on who you follow and more.

Search using a person's full name or @username | Search Twitter

josek_net @josek_net
#research #startups #tech ; CTO at @bsins & @wipley also professor at @canal_uem & @icemd. Follow @bsins & @gamisfaction : #eCommerce and #gamification tweets
Followed by recsys2011 and ACM Recsys 2010 .
Follow

Eugenio Di Sciascio @edisciascio
Followed by Cataldo Musto and Ec-Web 2010 .
Follow

Pearl Pu @pearlhzpu
Followed by recsys2011 and Ernesto W. De Luca .
Follow

barrysmyth @barrysmyth
Professor of Computer Science at University College Dublin. My research covers recommender systems, social search, and more. My current startup is heystaks.com.
Followed by ACM Recsys 2010 and Ernesto W. De Luca .
Follow

BigML @bigmlcom
Machine Learning, made simple. Offering
Follow

Recommendations in Facebook

Genius

File Modifica Vista Controlli Store Avanzate ?

iTunes

LIBRERIA
- Musica
- Film
- Programmi TV
- Libri
- App
- Radio

STORE
- iTunes Store
- iTunes Match
- Ping
- Acquisti

CONDIVISE
- Condivisione in famiglia

GENIUS
- Genius

PLAYLIST
- iTunes DJ
- Aggiunti di recente
- Anni '90
- Classifica
- I 25 più ascoltati
- Musica classica
- Recenti
- Videoclip

Genius

Genius makes playlists and mixes from songs in your library that go great together.
Genius also introduces you to music, movies, and TV shows you'll love, but don't already have in your library.

Genius Playlists.
To make a Genius playlist, select a song in your library and click the Genius button at the bottom of the iTunes window.

Genius Mixes.
Genius Mixes are created for you by iTunes and play continuously with one click. They will appear on the left after you turn on Genius.

iTunes Sidebar.
See recommendations for music, movies, and TV shows you don't already have each time you make a selection in your library.

Genius Recommendations.
Get recommendations right on the iTunes Store home page. Genius makes its picks by knowing what you already have.

Turning on Genius will send information about your iTunes library to Apple. Learn more.

Turn On Genius

Copyright © 2012 Apple Inc. All rights reserved. | Privacy Policy | Terms of Service | Terms of Sale

Recommendations by Genius

Google News

Netflix Prize: Home | Netflix - Watch TV Sh | Google News

news.google.it/?pog=false

presepe | CSA Archiviazio... | Oxford Paravia i... | JCR-Web 4.5 W... | II Facoltà di Scie... | Online convert T...

+Tu Ricerca Immagini Maps Play YouTube News Gmail Documenti Calendar Altro

Google

degemmis74@gmail.com

News

Edizione Italia | Moderno

Prima pagina
- Esodati
- Gervasoni
- Hollande
- Imu
- Dodge Dart
- Kim Jong-il
- Calciopoli
- X Factor
- Nba
- Seat Pagine Gialle

Puglia
Esteri
Italia
Economia
Scienza e tecnologia
Spettacoli
Sport
Salute
In evidenza

Prima pagina

Esodati, sfila a Roma il corteo dei sindacati. "Dal governo ...
Adnkronos/IGN - 12 minuti fa
Roma - (Adnkronos) - Camusso (Cgil): "Se non fosse un problema tragico sarebbe un gioco di prestigio. Il messaggio che arriva dopo il calcolo fatto dal governo è che gli esodati non esistono, sono dei fantasmi. Ma nessuno si può permettere di giocare ...

Sindacati in piazza per gli esodati. Camusso: «Soluzione, o ... Corriere della Sera
Esodati, Camusso: se i conti non tornano il presidente dell'Inps ... Il Sole 24 Ore

Correlati
Esodati »
Confederazione Generale Italiana del Lavoro »
Confederazione Italiana Sindacati Lavoratori »

Provenienza: Egitto: Esodati, sono 65 mila, il ministro Fornero parla di risorse adeguate BeFan (Blog)
Approfondimento: maxi protesta contro il governo TGCOM
Visualizza tutte le 193 fonti »

TM News | L'Unità | ParmaOggi | il Giornale | Sky.it | Panorama... | Quotidian...

PUNTO 1-Italia, produzione febbraio cala oltre attese, -0,7% su mese
Reuters Italia - 48 minuti fa
MILANO/ROMA, 13 aprile (Reuters) - Cala oltre le attese la produzione industriale italiana nel mese di febbraio, scontando oltre alla congiuntura negativa anche l'ondata di gelo che ha colpito il paese: sono dati che confermano una volta di più come ...

Nordcorea: Usa e alleati Pacifico condannano lancio missile
AGI - Agenzia Giornalistica Italia - 24 minuti fa
(AGI) - Washington, 13 apr. - Gli Stati Uniti ei loro principali alleati della regione del Pacifico hanno condannato il lancio del missile balistico operato stamane dalla Corea del Nord, pur se l'operazione e' fallita, definendolo "un atto ...

Alcol: i binge drinkers sono 4 milioni. Il numero in crescita tra ...
Il Sole 24 Ore - 21 minuti fa
Oltre 8 milioni sono gli italiani che consumano alcolici "a rischio" e oltre 4 milioni i "binge drinkers", ovvero quelli che almeno una volta nel corso dell'ultimo anno hanno consumato più di 6 bevande alcoliche in un'unica occasione.

Eni: guerriglieri attaccano installazioni in Nigeria
Agenzia di Stampa Asca - 52 minuti fa
(ASCA-AFP) - Lagos, 13 apr - Un gruppo di guerriglieri del Movimentio per l'Emancipazione del Niger Delta ha attaccato alcune installazioni petrolifere dell'Eni in Nigeria. La notizia, annunciata

Personalizza Google News

Esteri
Italia
Economia
Scienza e tecnologia
Spettacoli
Sport
Salute

Aggiungi qualsiasi argomento di notizie
Esempi: Astronomia, Parlamento, Calcio
Avanzate »

Modifica le fonti
Motori.it

Impostazioni | Reimposta | Centro assistenza

Salva

Recenti

Guzzanti-Marcegaglia al corteo dei sindacati
La Repubblica - 4 minuti fa

Borse giù, Milano e Madrid nel mirino
La Stampa - 30 minuti fa

Puglia » - Modifica

Sfida Campania-Puglia per Nec
Il Sole 24 Ore - 5 ore fa

Sanita': Puglia; Cisl,piano rientro e' a carico povera gente

IT 13:27 13/04/2012

Coming Soon: NetFlix

Netflix Prize: Home | Netflix - Watch TV Sh

https://signup.netflix.com/global

presepe | CSA Archiviazio... | Oxford Paravia i... | JCR-Web 4.5 W... | II Facoltà di Scie... | Online convert T...

NETFLIX

Member Sign In

Watch unlimited movies & TV episodes over the Internet!
For one low monthly price.

- Watch right on your TV or computer
- Cancel online anytime

Sorry, Netflix is not available in your country... yet

Enter your name and email address below and we'll email you when Netflix is available.

First Name
Last Name
Email
Confirm Email

Let me know

Secure Server
We will not sell or rent your email address.
See our Privacy Policy.

Press play and watch in seconds via the Internet.
Watch what you want, when you want. Unlimited!

Watch online on your computer

Watch online on your TV; connect your laptop directly to your TV

Watch online on your TV; connect your game consoles to your Netflix account

About Us | Affiliates | Blog | Contact Us | Developers | Investor Relations | Jobs | Media Center

I have a US mailing address

Use of the Netflix service and this Web site constitutes acceptance of our Terms of Use and Privacy Policy.
© 1997-2012 Netflix, Inc. All rights reserved. U.S. Patent Nos. 6,584,450; 7,024,381; 7,631,323; 7,403,910; and 7,617,127. (eu-west-1 ea8997a3 IT) [ ]

IT 13:24 13/04/2012

Thanks…

"Data scientists will be the rock stars of the Big Data era" (www.greenplum.com)

"La dignità non consiste nel possedere onori ma nella coscienza di meritarli " (Aristotele)

…Questions?

SWAP research group

Semantic
Web
Access and
Personalization
research group

http://www.di.uniba.it/~swap

References

- [Baeza-Yates and Ribeiro-Neto 1999] Baeza-Yates, R.A., Ribeiro-Neto, B.A., Modern Information Retrieval, ACM Press/Addison-Wesley, 1999.
- [Burke 2002] Burke, R. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction* 12(4):331-370, 2002.
- [Hanani et al. 2001] Hanani, U., B. Shapira, and P. Shoval. Information Filtering: Overview of Issues, Research and Systems. *User Modeling and User-Adapted Interaction* 11(3):203-259, 2001.
- [Kangas 2001] Sonja Kangas. *Collaborative Filtering and Recommendation Systems*. Technical Report TTE4-2001-35. VTT Information Technology, Espoo, Finland, 2001.
- [Lauw10] Lauw, H.W., Schafer, J.C., Agrawal, R., & A. Ntoulas. Homophily in the Digital World: A LiveJournal Case Study. *IEEE Internet Computing*14(2):15-23, March-April 2010.
- [Linden et al. 2003] Linden, G., B. Smith, and J. York. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing* 7(1):76-80, 2003.
- [Lops11] P. Lops, M. de Gemmis, and G. Semeraro. Content-based recommender systems: State of the art and trends. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor (Eds.), *Recommender Systems Handbook,* Springer, 73-105, 2011.
- [McNee06] S.M. McNee, J. Riedl, and J. Konstan. Accurate is not always good: How accuracy metrics have hurt recommender systems. In *Extended Abstracts of the 2006 ACM Conference on Human Factors in Computing Systems*, pages 1-5, Canada, 2006.
- [Pazzani07a] Pazzani, M. J., & D. Billsus. Content-Based Recommendation Systems. *The Adaptive Web*. Lecture Notes in Computer Science vol. 4321, 325-341, 2007.
- Plato Meno 80d-81a http://www.gutenberg.org/etext/1643
- [Resnick and Varian 1997] Resnick, P. and H. Varian. Recommender Systems. *Communications of the ACM* 40(3):56-58, 1997.
- [Zuckerman08] E. Zuckerman. Homophily, serendipity, xenophilia. April 25, 2008. www.ethanzuckerman.com/blog/2008/04/25/homophily-serendipity-xenophilia/

References

- [And83] J. R. Anderson. A Spreading Activation Theory of Memory. *Journal of Verbal Learning and Verbal Behavior* 22:261–295, 1983.
- [Anderson06] C. Anderson. *The Long Tail: Why the Future of Business Is Selling Less of More*. New York: Hyperion, 2006.
- [Basile07] P. Basile, M. Degemmis, A. Gentile, P. Lops, and G. Semeraro. UNIBA: JIGSAW algorithm for Word Sense Disambiguation. In Proceedings of the 4th ACL 2007 International Workshop on Semantic Evaluations (SemEval-2007), Prague, Czech Republic, pages 398–401. Association for Computational Linguistics, June 23–24, 2007.
- [Degemmis07] M. Degemmis, P. Lops, and G. Semeraro. A Content-collaborative Recommender that Exploits WordNet-based User Profiles for Neighborhood Formation. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research (UMUAI)* 17(3):217–255, Springer Science + Business Media B.V., 2007.
- [Egozi08] O. Egozi, E. Gabrilovich, S. Markovitch: Concept-Based Feature Generation and Selection for Information Retrieval. AAAI 2008: 1132–1137, 2008.
- [Egozi11] O. Egozi, S. Markovitch, E. Gabrilovich. Concept-Based Information Retrieval using Explicit Semantic Analysis. *ACM Transactions on Information Systems* 29(2), April 2011.
- [Gabri06] E. Gabrilovich and S. Markovitch. Overcoming the Brittleness Bottleneck using Wikipedia: Enhancing Text Categorization with Encyclopedic Knowledge. In Proceedings of the 21th National Conf. on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence Conference, pages 1301–1306. AAAI Press, 2006.
- [Gabri07] E. Gabrilovich and S. Markovitch. Harnessing the expertise of 70,000 human editors: Knowledge-based feature generation for text categorization. Journal of Machine Learning Research, 8, 2297–2345, (2007)
- [Gabri09] E. Gabrilovich and S. Markovitch. Wikipedia-based Semantic Interpretation for Natural Language Processing. Journal of Artificial Intelligence Research, 34 (2009) 443–498
- [Giles05] J. Giles. Internet Encyclopaedias Go Head to Head. *Nature*, 438:900–901, 2005.
- [Huvila10] Huvila, I. Where does the information come from? Information source use patterns in Wikipedia. *Information Research* 15, 3, 2010.
- [Goel10] S. Goel, A. Z. Broder, E. Gabrilovich, and B. Pang, “Anatomy of the long tail: ordinary people with extraordinary tastes.” in WSDM, B. D. Davison, T. Suel, N. Craswell, and B. Liu, Eds. ACM, 2010, pp. 201–210. [Online]. Available: http://dblp.unitrier.de/db/conf/wsdm/wsdm2010.html
- [Mura08] T. Murakami, K. Mori and R. Orihara. Metrics for Evaluating the Serendipity of Recommendation Lists . Lecture Notes in Computer Science, Vol. 4914/2008, pp. 40–46, 2008.

References

- [Pedersen04] Pedersen, Ted and Patwardhan, Siddharth, and Michelizzi, Jason. WordNet:Similarity – Measuring the Relatedness of Concepts. In Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-2004), pp. 1024-1025, San Jose, CA, July, 2004.
- [Semeraro07] G. Semeraro, M. Degemmis, P. Lops, and P. Basile. Combining Learning and Word Sense Disambiguation for Intelligent User Profiling. In M. M. Veloso, editor, *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007)*, Hyderabad, India, January 6-12, 2007, pages 2856-2861. Morgan Kaufmann, 2007.
- [Sem09] G. Semeraro, P. Lops, P. Basile, and M. de Gemmis. On the Tip of my Thought: Playing the Guillotine Game. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2009)*, Pasadena, CA, USA, July 11-17, 2009, pages 1543-1548, Morgan Kaufmann, 2009.
- [Sem12] G. Semeraro, M. de Gemmis, P. Lops, P. Basile. Knowledge Infusion from Open Knowledge Sources: an Artificial Player for a Language Game, *IEEE Intelligent Systems*. In Press.