

Introduction to Recommender Systems

Joseph A. Konstan
University of Minnesota

konstan@cs.umn.edu
<http://www.grouplens.org>

These notes will be posted next week at:
<http://www.umn.edu/~konstan>

While you're waiting, please jot down:

- Where you are from
- Your experience with recommenders
- What you want to get out of this tutorial

A Bit of History

- Ants, Cavemen, and Early Recommender Systems
 - The emergence of critics
- Information Retrieval and Filtering
- Manual Collaborative Filtering
- Automated Collaborative Filtering
- The Commercial Era

A Bit of History

- Ants, Cavemen, and Early Recommender Systems
 - The emergence of critics
- Information Retrieval and Filtering
- Manual Collaborative Filtering
- Automated Collaborative Filtering
- The Commercial Era

Information Retrieval

- Static content base
 - Invest time in indexing content
- Dynamic information need
 - Queries presented in “real time”
- Common approach: TFIDF
 - Rank documents by term overlap
 - Rank terms by frequency

Information Filtering

- Reverse assumptions from IR
 - Static information need
 - Dynamic content base
- Invest effort in modeling user need
 - Hand-created “profile”
 - Machine learned profile
 - Feedback/updates
- Pass new content through filters

A Bit of History

- Ants, Cavemen, and Early Recommender Systems
 - The emergence of critics
- Information Retrieval and Filtering
- **Manual Collaborative Filtering**
- Automated Collaborative Filtering
- The Commercial Era

Collaborative Filtering

- Premise
 - Information needs more complex than keywords or topics: quality and taste
- Small Community: Manual
 - Tapestry – database of content & comments
 - Active CF – easy mechanisms for forwarding content to relevant readers

A Bit of History

- Ants, Cavemen, and Early Recommender Systems
 - The emergence of critics
- Information Retrieval and Filtering
- Manual Collaborative Filtering
- **Automated Collaborative Filtering**
- The Commercial Era

Automated CF

- The GroupLens Project (CSCW '94)
 - ACF for Usenet News
 - users rate items
 - users are correlated with other users
 - personal predictions for unrated items
 - Nearest-Neighbor Approach
 - find people with history of agreement
 - assume stable tastes

The screenshot shows the xrn newsreader interface. At the top, there's a list of news articles with their subject lines and IDs. Below this is a toolbar with buttons for navigating through the news feed. The main area displays a single news article in a detailed view. The article header includes the From, Subject, Organization, and a note about being exported from MasterCook. The body of the article contains a recipe for Crème Brûlée, listing ingredients and preparation steps. At the bottom of the article view, there's a message encouraging users to like similar articles, followed by a row of interaction buttons.

Subject	ID	Author
17107 Cream Brulee	[63]	Art Poe
17108 xSamHaring	[63]	Sam Haring
17109 Banana Dream Pie	[63]	Art Poe
17110 Baked Potato Soup	[63]	Art Poe
17111 xHinespoon	[63]	Sam Haring
17112 xPorkyDuddy	[63]	Sam Haring
17113 xLekkuchen	[64]	Sam Haring
17114 xPlum Chutney	[63]	Sam Haring
17115 xSwedish Flatbread	[453]	Sam Haring

Operations apply to current selection or cursor position

From: "Art Poe" <poet@unicow.net>
Subject: Cream Brulee
Organization: Unicow

MasterCook export: Cream Brulee
* Exported from MasterCook *

Creme Brulee

Recipe Size : Food and Wine - Dec98
Serving Size : 8 Preparation Time 10:00
Categories : Desserts

Amount	Measure	Ingredient	Preparation Method
4	Cups	Heavy cream	
1	Pinch	Vanilla bean	
8	3/4 Cup	Salt	
2	Tablespoons	Egg yolks	
8	Tablespoons	Sugar	
		Brown sugar	

Preheat the oven to 300F. In a heavy medium saucepan, combine the cream, vanilla bean and salt. Warm over moderate heat until the surface begins to shimmer, about 5 minutes. In a large bowl, stir the egg yolks and sugar until thick and pale. Slowly add the warm custard, stirring until forming air bubbles. Strain the custard into a large measuring glass and skim off any surface air bubbles. (Rinse the vanilla bean and reserve for future use.) Place the ramekins in a roasting pan. Pour the custard into the ramekins, filling them up to the top. Place the roasting pan in the oven and pour in enough warm water to reach halfway up the sides of the ramekins. Cover loosely with foil and bake for 1 1/4 hours, or until the custard is set and the top is golden brown (it will be wobbly in the center but it will firm up as it chills.) Remove the ramekins from the

This article is great, I'd like to see more like this one!

Save Reply Forward Followup Followup & Reply Cancel Rot-13 Translate
Toggle header Print artRate1 artRate2 artRate3 artRate4 artRate5

Does it Work?

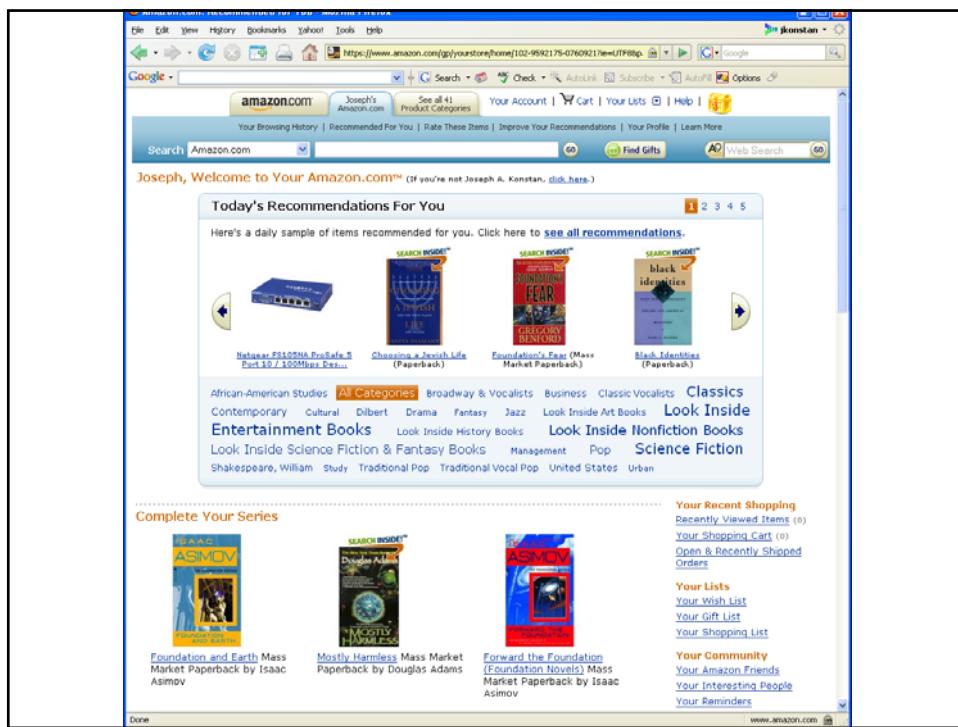
- Yes: The numbers don't lie!
 - Usenet trial: rating/prediction correlation
 - rec.humor: 0.62 (personalized) vs. 0.49 (avg.)
 - comp.os.linux.system: 0.55 (pers.) vs. 0.41 (avg.)
 - rec.food.recipes: 0.33 (pers.) vs. 0.05 (avg.)
 - Significantly more accurate than predicting average or modal rating.
 - Higher accuracy when partitioned by newsgroup

It Works Meaningfully Well!

- Relationship with User Behavior
 - Twice as likely to read 4/5 than 1/2/3
- Users *Like* GroupLens
 - Some users stayed 12 months after the trial!

A Bit of History

- Ants, Cavemen, and Early Recommender Systems
 - The emergence of critics
- Information Retrieval and Filtering
- Manual Collaborative Filtering
- Automated Collaborative Filtering
- The Commercial Era



Historical Challenges

- Collecting Opinion and Experience Data
- Finding the Relevant Data for a Purpose
- Presenting the Data in a Useful Way

A Few Challenges Specifically Relevant to SIGMOD Folks

(we'll come back to these)

- Building database systems that can compute recommendations efficiently
- Scaling up to processing entire streams of transactions

Recommenders

- Tools to help identify worthwhile stuff
 - Filtering interfaces
 - E-mail filters, clipping services
 - Recommendation interfaces
 - Suggestion lists, “top-n,” offers and promotions
 - Prediction interfaces
 - Evaluate candidates, predicted ratings

Scope of Recommenders

- Purely Editorial Recommenders
- Content Filtering Recommenders
- Collaborative Filtering Recommenders
- Hybrid Recommenders

Wide Range of Algorithms

- Simple Keyword Vector Matches
- Pure Nearest-Neighbor Collaborative Filtering
- Machine Learning on Content or Ratings

Tutorial Goals and Outline

Goals

- When you leave, you should ...
 - Understand recommender systems and their application
 - Know enough about recommender systems technology to evaluate application ideas
 - Be familiar with a variety of recommendation algorithms
 - See where recommender systems have been, and where they are going
 - Have seen a large number of recommender system example cases from research and practice

Outline

- Introduction
- Recommender Application Space
 - Dimensions of Analysis
 - Case Examples
- Algorithms
- Issues / Current Research / Advanced Topics
- Discussion and Questions
 - also encouraged throughout
- Bonus content on slides

Introductions

- Recommender systems
- This tutorial
- Me
- You

About Me

- Professor of Computer Science
 - University of Minnesota
- Background: Human-Computer Interaction
- Recommender Systems Experience
 - Started on GroupLens project in late 1994
 - co-founded Net Perceptions
 - co-authored *Word of Mouse*
 - Still actively working on RS research
 - applications to digital libraries
 - also work on online community; e-Public Health

About You

- Name
- What you do
- Who you work for / where you study
- Briefly
 - Your experience with recommender systems
 - One key thing you want to get out of this tutorial

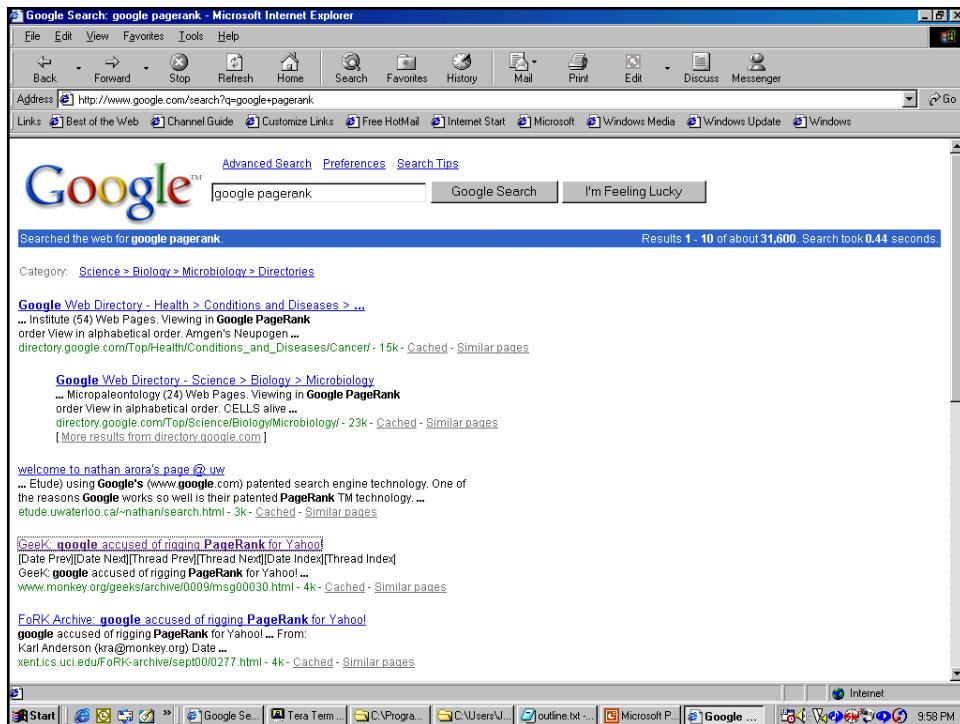
Recommender Application Space

Recommender Application Space

- Dimensions of Analysis
 - Domain
 - Purpose
 - Whose Opinion
 - Personalization Level
 - Privacy and Trustworthiness
 - Interfaces
 - <Algorithms Inside>

Domains of Recommendation

- Content to Commerce
 - News, information, “text”
 - Products, vendors, bundles



Purposes of Recommendation

- The recommendations themselves
 - Sales
 - Information
- Education of user/customer
- Build a community of users/customers around products or content

The screenshot shows a product page for a Lexar 128MB USB Flash Jump Drive on the buy.com website. The page includes the following details:

- Product Name:** LEXAR 128MB USB FLASH JUMP DRIVE
- Price:** \$45.47 (After Rebate)
- Rebate Offer:** \$20 Mail-in Rebate available.
- Smart Savings:** Offers for a Humping USB Mini-Hub and a 32MB Flash Key USB Pen Drive Storage Device.
- Customer Rating:** Based on 16 reviews.
- Related Products:** Viking 128MB USB Flash Drive, 32MB Flash Key USB Pen Drive, and Lexar 128MB Memory.
- Also Bought:** Hawking Tech Humping USB Mini-Hub, CD CYCLONE, and SmartDisk Universal Memory.

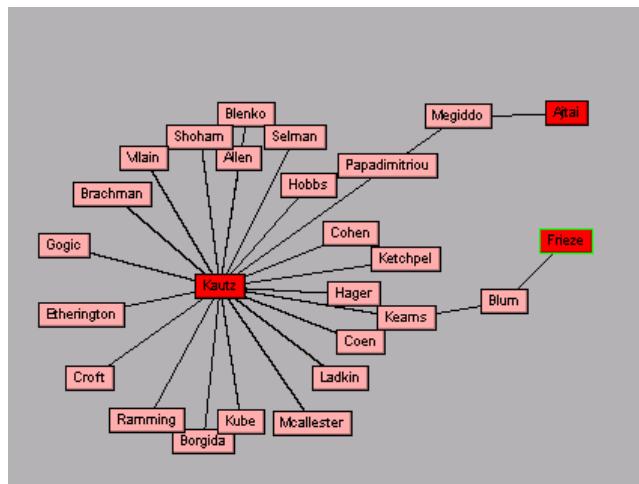
The screenshot shows a Microsoft Internet Explorer window with the title bar "Opinions.com - Reviews of 2000 Toyota Sienna - Microsoft Internet Explorer". The address bar contains the URL "http://www.opinions.com/auto_Make-2000_Toyota_Sienna". The main content area displays a product review page for a 2000 Toyota Sienna. At the top, it says "Average rating: ★★★★ (62 Member Opinions)" and "Recommended 93% of the time". A thumbnail image of the van is shown, along with its price (\$20730-27334), manufacturer (Toyota), class (Van & Minivan), model year (2000), and model pictured (XLE 4 Dr Pass Van). To the right, there's a sidebar with a woman's face and the text "YOU HAVE THAT ON OPINIONS?". Below the main content, there are several sections: "Love It / Hate It", "Get Updates", "Help others decide. Write an opinion!", "Go Shopping!", "Related Items" (listing 2001 and 2000 models), and "Member Opinions" (listing reviews like "Toyota Sienna: Try to Find a better minivan" and "Quiet, comfortable ride"). On the right side, there are banners for "Write an Opinion", "Professional Reviews" (listing "1999 Toyota Sienna XLE"), and "See What's New". The bottom right corner of the window shows the word "Internet".

OWL Tips

OWL Tips			
Date	Word Commands	Comments	Priority
9/30/98	EditFind	Use Find more to search for text in file	88
	EditDeleteWord	Learn the shortcut keys to delete words	87
	FormatUnderline	Try using underlining for formatting text	80
	FileClose	Try different ways to close your file	74
	EditReplace	Use Replace more for finding and replacing text	72
	ViewZoom	Learn how to enlarge or reduce the display	69
	ViewShowAll	FYI-more than average use for Show All comman	67
	FormatBulletsAndNumbering	Learn how to automatically add bullets and num	64
	ToolsWordCount	Use Word Count to look up statistics on files	61
	ViewPage	Use Page Layout to view files before printing	57

User M06375 OWL Version 5.0c [OWL Help](#) [OK](#)

ReferralWeb



Whose Opinion?

- “Experts”
- Ordinary “phoaks”
- People like you

1997 Casa Lapostolle Cabernet Sauvignon, Rapel Valley, Chile

Peter's Tasting Chart

intensity	dry or sweet	delicate	powerful
bone dry	dry or sweet	bone dry	desert
body	body	light body	very full body
acidity	soft, gentle	soft, gentle	very crisp
tannin	none	none	heavy tannins
oak	oak	none	heavy oak
complexity	complexity	direct	very complex

Frequently Mentioned Resources

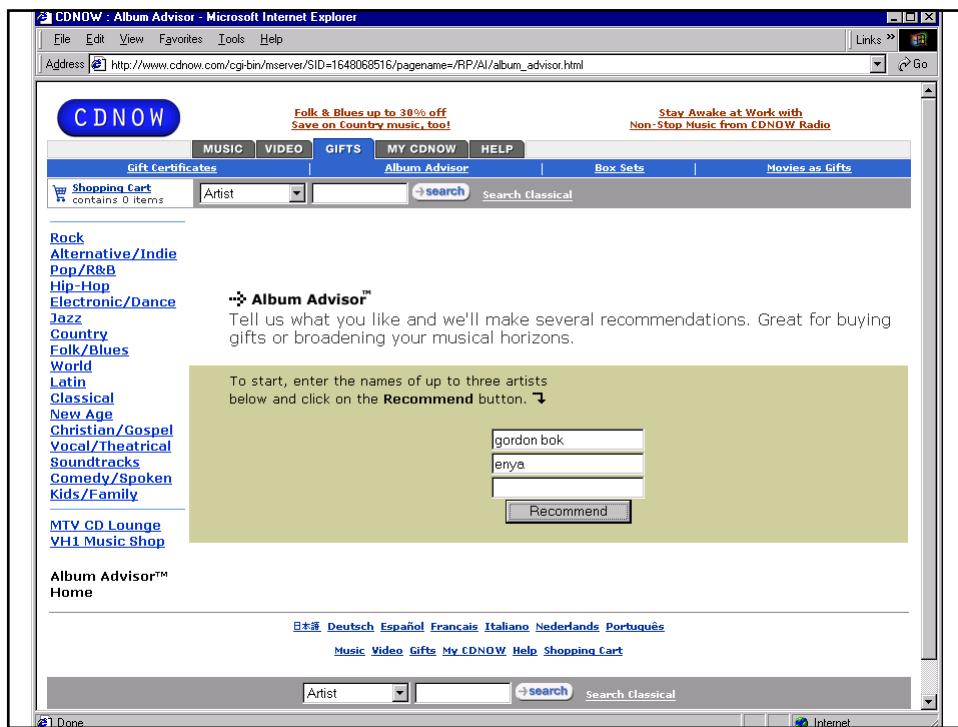
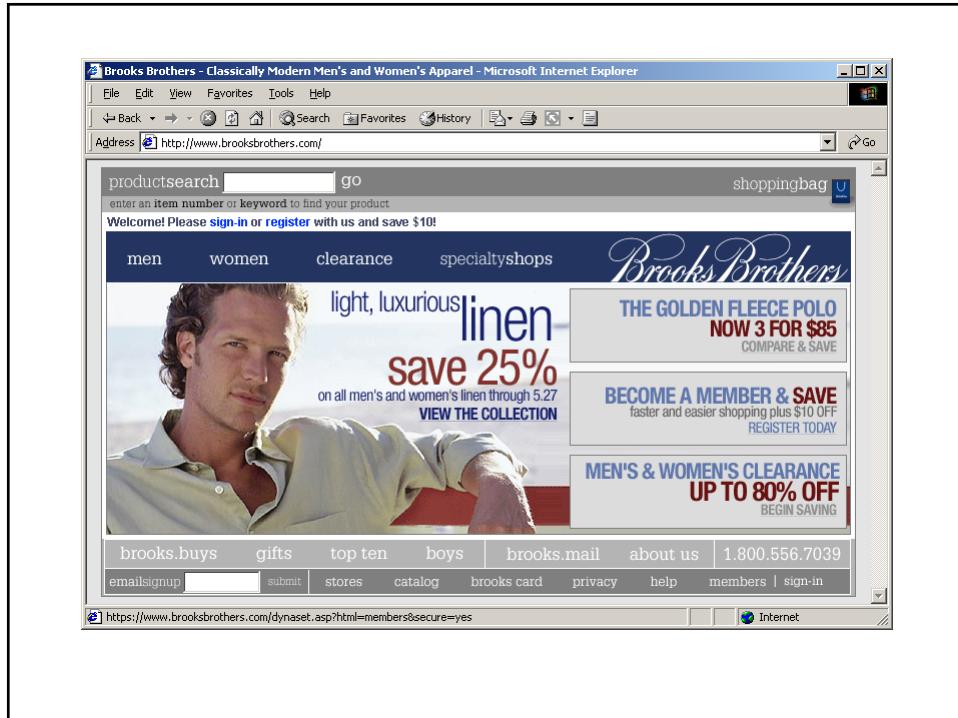
Resource Title	Distinct Posters	Click on Bars for Message Context(s)
1) Bob Dylan - Bob Links	23	...
2) Bob Dylan Chords	9	[Bar]
3) RemarQ - The Internet's Best Collaboration...	8	[Bar]
4) bobdylan.com: Bob Dylan	7	[Bar]
5) CDNOW	6	[Bar]
6) Mailing List WWW Gateway	6	[Bar]
7) Deja.com	4	[Bar]
8) LC Z39.50 Server Soft Reference	4	[Bar]
9) Resource at www.cs.umass.edu	4	[Bar]
10) Sidewalk	4	[Bar]

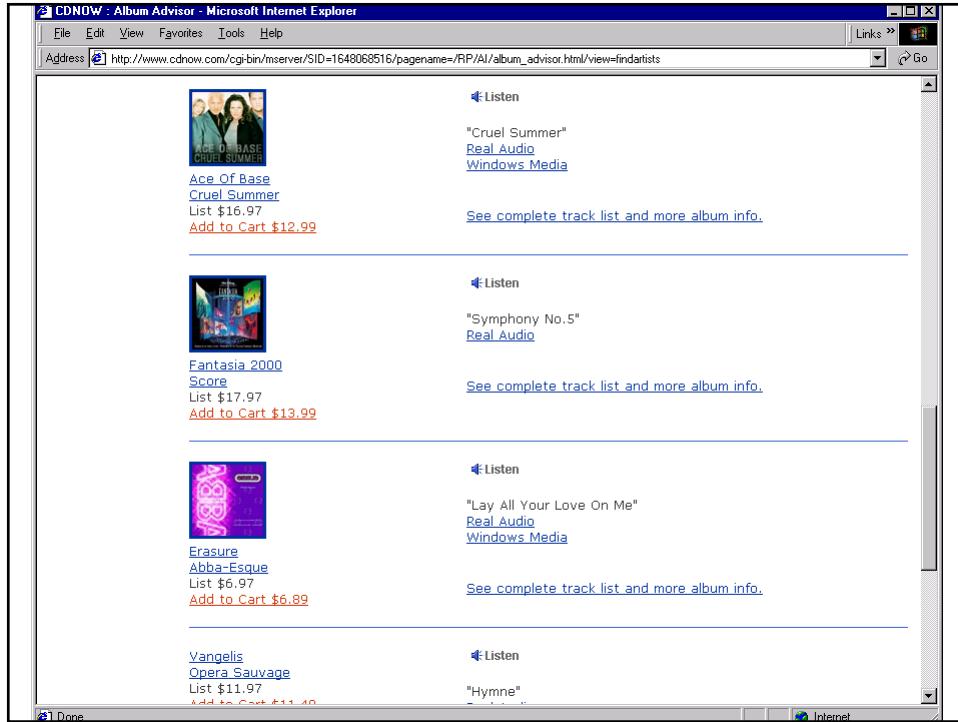
* Note: each square represents the posting of one resource (e.g., URL) by one person. The lighter the square, the more recent the post. Click on a square to view messages where this resource was mentioned. Posting a web resource does not necessarily imply endorsing that resource. Sometimes it may actually mean the opposite. Consult the relevant netnews messages to obtain context.

Personalization Level

- Generic
 - Everyone receives same recommendations
- Demographic
 - Matches a target group
- Ephemeral
 - Matches current activity
- Persistent
 - Matches long-term interests

The screenshot shows a Microsoft Internet Explorer window displaying the Lands' End website. The main headline reads "It makes looking good look easy. The slimming Faile Tankini – just \$58!" Below this, there's a section titled "AS SEEN ON TV!" featuring three women in Faile Tankinis. To the right, there's an "Important Notice" about Lands' End being acquired by Sears. Further down, there's a "Lands' End Custom" section where users can design their own pants, and a "Subscribe!" section for newsletters. On the left sidebar, there are links for shopping bags, checkout, account management, and personal shoppers. The footer includes a "Swim Finder" tool and information about new Lands' End Maternity products.





ORDER HISTORY

Order #14856068 Placed on December 25, 1999 Status: Order Shipped Get additional order information.

Rewards

Current Program: None You are not currently signed up for any rewards program. Start earning rewards for FREE CDs and more right now!

PREFERENCES

Customize your

CDNOW RECOMMENDS

John Coltrane Love Supreme List Price \$17.97 Add to Cart \$13.99 More items recommended for you!

WISH LIST

Want to keep track of items you might like to purchase at a later date? Start your list today!

FAVORITE ARTISTS

Bok*Gordon Muir*Ann Mayo / Trickett*Ed Go to Artist

Modify your options and check for Advance Orders and New Releases from your favorite artists.

ACCOUNT INFORMATION & ADDRESS BOOK

Modify your Account Information, change your Primary Address or set up Express Checkout.

RATE YOUR MUSIC

You have 3 items that you haven't rated yet. Tell us [] about the music you own and help us make better recommendations!

Privacy and Trustworthiness

- Who knows what about me?
 - Personal information revealed
 - Identity
 - Deniability of preferences
- Is the recommendation honest?
 - Biases built-in by operator
 - “business rules”
 - Vulnerability to external manipulation

Interfaces

- Types of Output
 - Predictions
 - Recommendations
 - Filtering
 - *Organic vs. explicit presentation*
 - Agent/Discussion Interface
- Types of Input
 - Explicit
 - Implicit

Collaborative Filtering: Techniques and Issues

Collaborative Filtering Algorithms

- Non-Personalized Summary Statistics
- K-Nearest Neighbor
- Dimensionality Reduction
- Content + Collaborative Filtering
- Graph Techniques
- Clustering
- Classifier Learning

Teaming Up to Find Cheap Travel

- Expedia.com
 - “data it gathers anyway”
 - (Mostly) no cost to helper
 - Valuable information that is otherwise hard to acquire
 - Little processing, lots of collaboration

Build your own trip - MSP to LAS - Mozilla Firefox

File Edit View History Bookmarks Yahoo! Tools Help

http://www.expedia.com/pub/agent.dll?tovr=1294677293

Google sigir 2008

Welcome My Itineraries

Expedia.com

Home Flights Hotels Cars Vacation Packages Cruises Activities Deals & Destinations Maps Bu

Minneapolis, MN (MSP) to Las Vegas, NV (LAS)

Change your search

Departure airport: MSP (Minneapolis) Destination airport: LAS (Las Vegas)

Travel Dates: Specific | Flexible Departing: (mm/dd/yy) 9/23/2007 Returning: (mm/dd/yy) 9/26/2007 Anytime

Airline: [More Info] No Preference Class: Economy / Coach Nonstop flights only Refundable flights

	All Results	Sun Country Airlines	nwa Northwest	US Airways	Midwest Airlines	Frontier Airlines
Nonstop	from \$265 see below	from \$265	from \$265	from \$275	---	---
1 stop	from \$269 see below	---	from \$269	from \$279	from \$283	from \$286
2+ stops	---	---	---	---	---	---

Show more airlines »

Note: The prices shown below are for the **flight only**; they are e-ticket prices and include [all flight taxes and fees](#). If your itinerary requires paper tickets there will be an [additional charge](#). These results cover a metro area with [several airports](#). Review your choices carefully.

1 Choose a departing flight or view complete roundtrips

Booking online is a snap -- or call (800) 434-2370.

Sort by: Price Duration Departure time Arrival time

Find: Next Previous Highlight all Match case javascript:PrevFlexSearch()

The screenshot shows the Expedia.com website for flight search. The URL is <http://www.expedia.com/pub/agent.dll?ovr=1294677293>. The search parameters are set for "Minneapolis, MN (MSP) to Las Vegas, NV (LAS)". The departure date is set to September 2007. A calendar is displayed for September 2007, showing flight prices for each day. The prices are as follows:

Date	Price
1	\$289
2	\$265
3	\$265
4	\$275
5	\$265
6	\$275
7	\$275
8	\$275
9	\$265
10	\$265
11	\$265
12	\$275
13	\$275
14	\$285
15	\$275
16	\$265
17	\$265
18	\$265
19	\$275
20	\$275
21	\$275
22	\$275
23	\$265
24	\$274
25	\$275
26	\$285
27	\$285
28	\$275
29	\$275
30	\$275

Other sections visible include "Start search over", "Questions?", "How this works:", and "Expedia Information".

The screenshot shows the Zagat.com website for restaurant search. The URL is <http://www.zagat.com/Search/Results.aspx?Ne=1118&Ntt=764&VID=8&N=120+4294909891&Ntk=GeoChildID&Ns=Frontmatter+Nun>. The search parameters are set for "Find a Restaurant in Vancouver". The results are listed in Map View, showing five restaurants:

- West**: 2881 Granville St. (W. 13th Ave.), Vancouver, British Columbia, Canada | Map | 604-738-8938 | Reserve Online
- Lumière**: 2551 W. Broadway (bet. Larch & Trafalgar Sts.), Vancouver, British Columbia, Canada | Map | 604-739-8165 | Reserve Online
- Bishop's**: 2183 W. 4th Ave. (bet. Arbutus & Yew Sts.), Vancouver, British Columbia, Canada | Map | 604-738-2025 | Reserve Online
- Pear Tree, The**: 4120 E. Hastings St. (bet. Cambie & Gilmore Aves.), Burnaby, British Columbia, Canada | Map | 604-299-2772 | Reserve Online
- ToJo's**: 1133 W. Broadway (bet. Hastings & Willow Sts.), Vancouver, British Columbia, Canada | Map | 604-372-8050 | Reserve Online

Each restaurant entry includes its address, neighborhood, cuisine, ZAGAT RATED score (e.g., 29, 26, 27, 57), and a link to add a review.

The screenshot shows a web browser displaying the Zagat website at <http://www.zagat.com/Verticals/PropertyDetails.aspx?ID=8&R=99905&AJX=N%253D1118%2526Ntt%253D764%2526VID%253D8%2525>. The page is for a restaurant named Tojo's. On the left, there's a sidebar with 'TOP LISTED' sections for Food Tops, Japanese Tops, and Most Popular. Below that is a 'ToJo's Stats' section with details like Meals Served (Dinner), Payment (Accepts Major Credit Cards), Dress Code (Casual), Year Opened (1988), Good For (Business Dining), Special Features (Outdoor Seating, Takeout Available), Vibe / Atmosphere (Power Scene), and Hours (Mon-Fri: 5:00PM - 9:45PM). The main content area features a 'ZAGAT Ratings & Review' section with a rating of 27, 23, and \$73. It includes a detailed description of the restaurant, a review form, and a map showing its location on W 8th Ave. To the right, there's an 'EDIT YOUR REVIEW' section, a sidebar for finding places near Tojo's, and advertisements for Tropicana Pure Valencia juice and BOSE.

Zagat: Is Non-Personalized Good Enough?

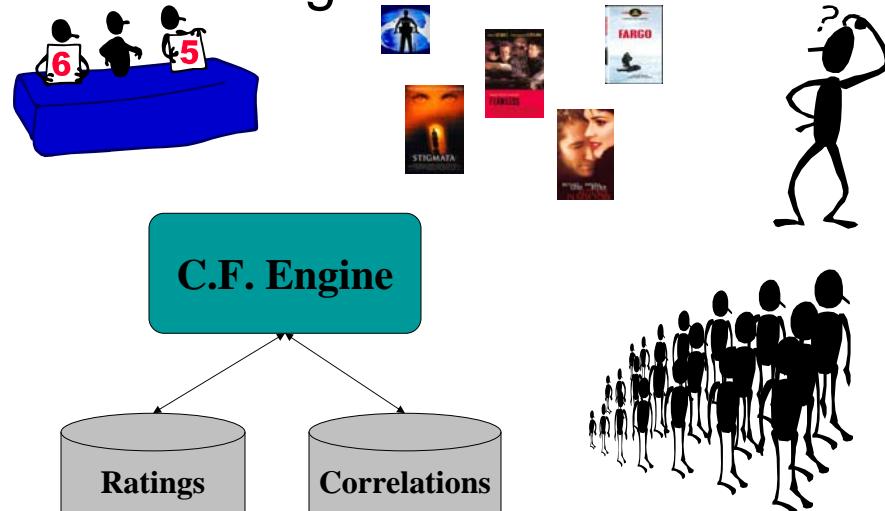
- What happened to my favorite guide?
– They let *you* rate the restaurants!

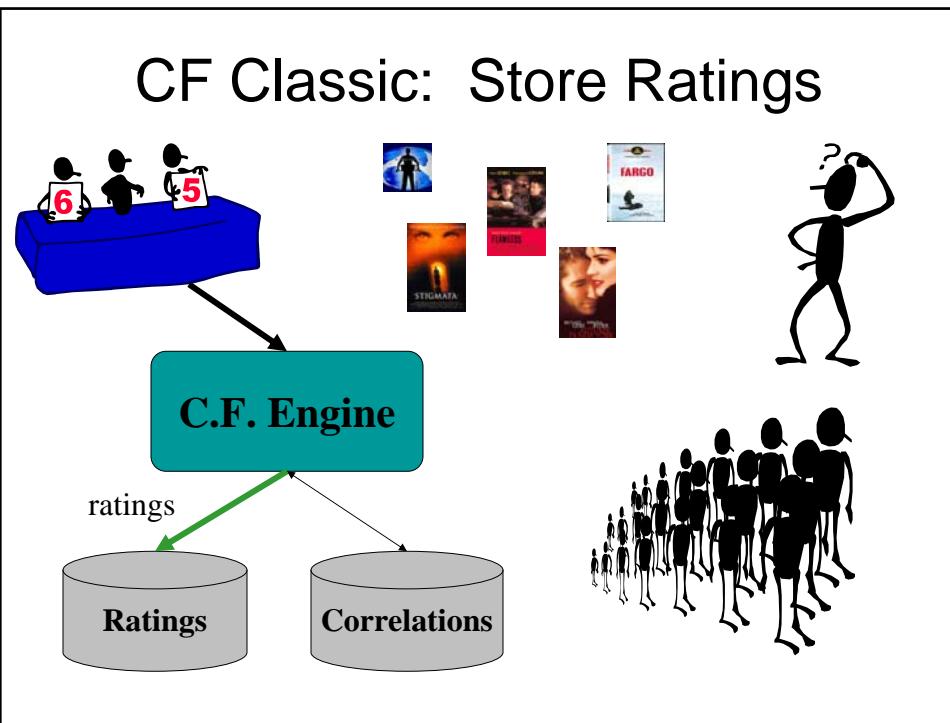
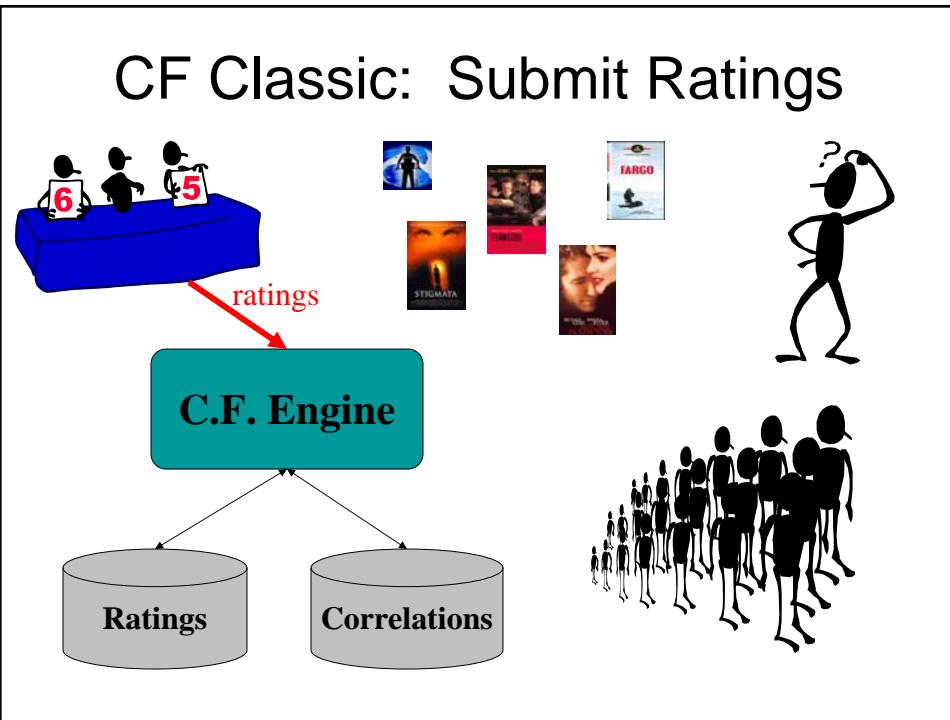
- What should be done?
– Personalized guides, from the people who “know good restaurants!”

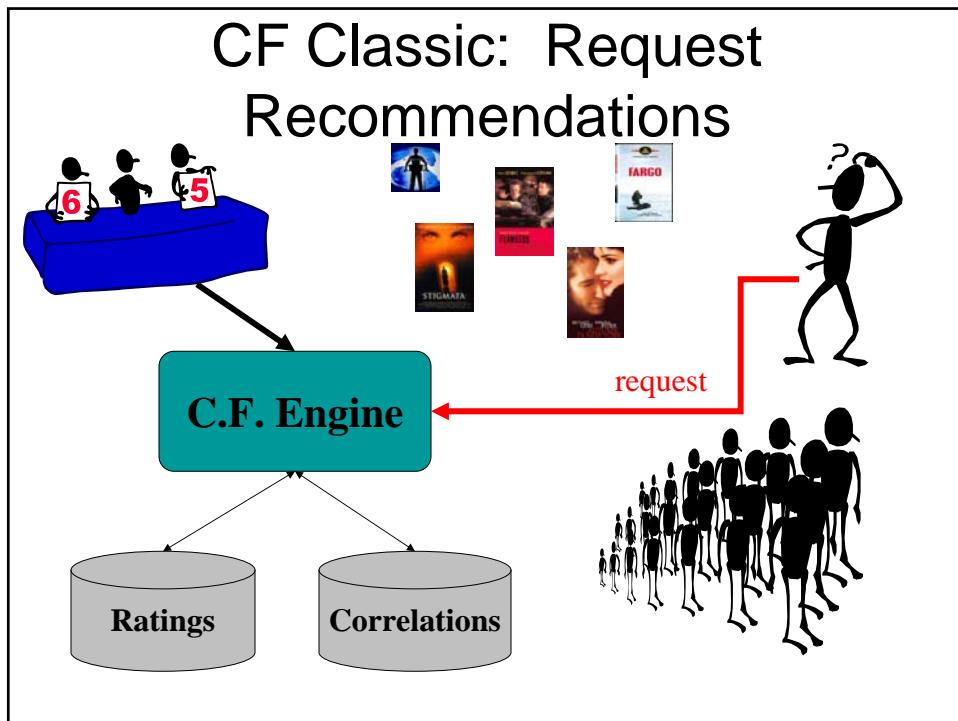
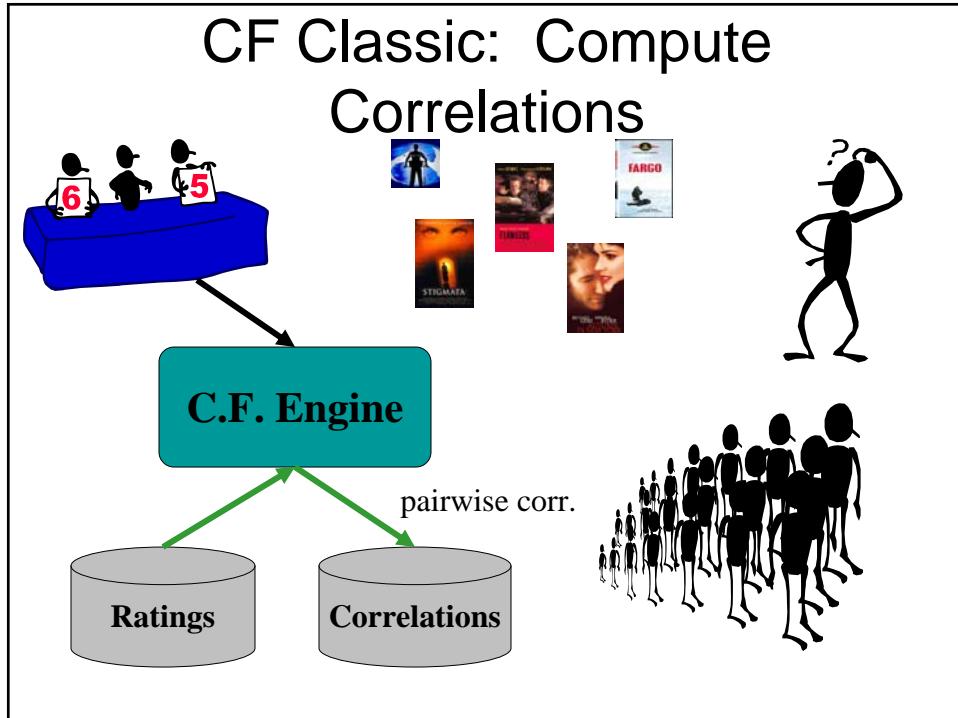
Collaborative Filtering Algorithms

- Non-Personalized Summary Statistics
- K-Nearest Neighbor
 - user-user
 - item-item
- Dimensionality Reduction
- Content + Collaborative Filtering
- Graph Techniques
- Clustering
- Classifier Learning

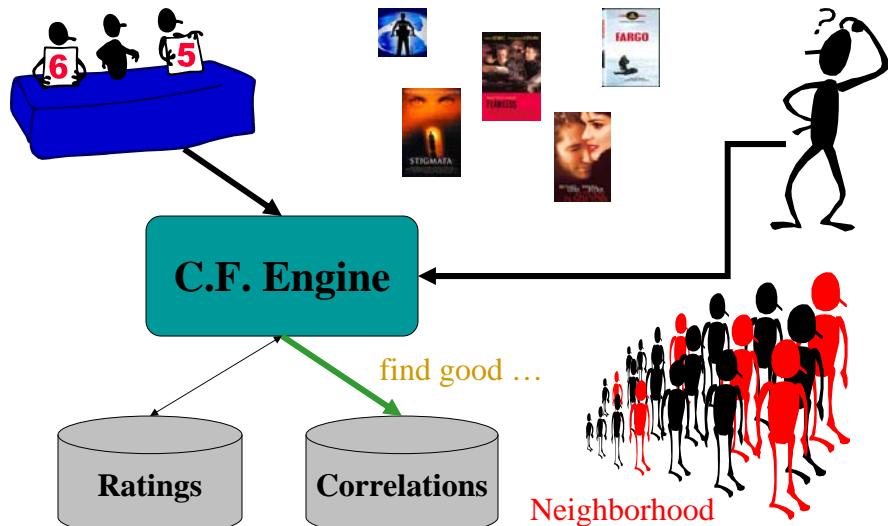
CF Classic: K-Nearest Neighbor User-User



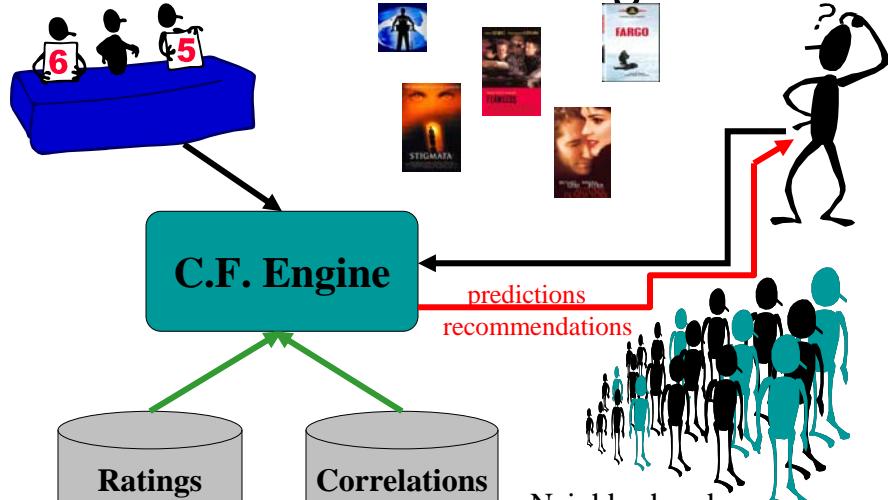




CF Classic: Identify Neighbors



CF Classic: Select Items; Predict Ratings



Understanding the Computation

	Hoop Dreams	Star Wars	Pretty Woman	Titanic	Blimp	Rocky XV
Joe	D	A	B	D	?	?
John	A	F	D		F	
Susan	A	A	A	A	A	A
Pat	D	A		C		
Jean	A	C	A	C		A
Ben	F	A				F
Nathan	D		A		A	

Understanding the Computation

	Hoop Dreams	Star Wars	Pretty Woman	Titanic	Blimp	Rocky XV
Joe	D	A	B	D	?	?
John	A	F	D		F	
Susan	A	A	A	A	A	A
Pat	D	A		C		
Jean	A	C	A	C		A
Ben	F	A				F
Nathan	D		A		A	

Understanding the Computation

	Hoop Dreams	Star Wars	Pretty Woman	Titanic	Blimp	Rocky XV
Joe	D	A	B	D	?	?
John	A	F	D		F	
Susan	A	A	A	A	A	A
Pat	D	A		C		
Jean	A	C	A	C		A
Ben	F	A				F
Nathan	D		A		A	

Understanding the Computation

	Hoop Dreams	Star Wars	Pretty Woman	Titanic	Blimp	Rocky XV
Joe	D	A	B	D	?	?
John	A	F	D		F	
Susan	A	A	A	A	A	A
Pat	D	A		C		
Jean	A	C	A	C		A
Ben	F	A				F
Nathan	D		A		A	

Understanding the Computation

	Hoop Dreams	Star Wars	Pretty Woman	Titanic	Blimp	Rocky XV
Joe	D	A	B	D	?	?
John	A	F	D		F	
Susan	A	A	A	A	A	A
Pat	D	A		C		
Jean	A	C	A	C		A
Ben	F	A				F
Nathan	D		A		A	

Understanding the Computation

	Hoop Dreams	Star Wars	Pretty Woman	Titanic	Blimp	Rocky XV
Joe	D	A	B	D	?	?
John	A	F	D		F	
Susan	A	A	A	A	A	A
Pat	D	A		C		
Jean	A	C	A	C		A
Ben	F	A				F
Nathan	D		A		A	

Understanding the Computation

	Hoop Dreams	Star Wars	Pretty Woman	Titanic	Blimp	Rocky XV
Joe	D	A	B	D	?	?
John	A	F	D		F	
Susan	A	A	A	A	A	A
Pat	D	A		C		
Jean	A	C	A	C		A
Ben	F	A				F
Nathan	D		A		A	

The screenshot shows the MovieLens website interface. At the top, there's a navigation bar with links like File, Edit, View, Go, Bookmarks, Tools, Help, Home, Manage Buddies, Your Preferences, Help, Publish, and Logout. Below the navigation is a banner for 'Welcome riedl@cs.umn.edu' with information about movie counts and visitor history. A sidebar on the left lists various rating and search options. The main content area features a 'Welcome to MovieLens!' message, an 'Advanced Search' section, and a 'Did you know...' statistic. At the bottom, there are two columns of movie recommendations: 'New movies' and 'New DVDs', each with a list of titles and their ratings.

The screenshot shows the MovieLens homepage in Mozilla Firefox. The title bar reads "movielens - Mozilla Firefox". The main content area features the MovieLens logo and the tagline "helping you find the right movies". A welcome message for "riedl@cs.umn.edu" indicates they have rated 205 movies and are the 31st visitor in the past hour. To the right, a legend defines rating stars: 5 stars = Must See, 4 stars = Will Enjoy, 3 stars = It's OK, 2 stars = Fairly Bad, 1 star = Awful. Below the legend, there are links for Home, Manage Buddies, Your Preferences, Help, Publish, and Logout. On the left, there are three search boxes: "Search Titles" (with a "Go!" button), "Search Genres" (with dropdowns for Sci-Fi, 2000s, and All movies, plus a "Search Genres!" button), and "Advanced Search". The central content area is titled "Welcome to MovieLens!" and contains a section about the Advanced Search feature. It also mentions the "publish" feature and links to archived announcements. A "Did you know..." section notes that 4909 people joined MovieLens on the same day. Below this are two lists: "New movies" and "New DVDs", each with five movie entries and their ratings.

The screenshot shows the search results page in Mozilla Firefox. The title bar reads "movielens - Mozilla Firefox". The main content area features the MovieLens logo and the tagline "helping you find the right movies". A welcome message for "riedl@cs.umn.edu" indicates they have rated 206 movies and are the 31st visitor in the past hour. To the right, a legend defines rating stars: 5 stars = Must See, 4 stars = Will Enjoy, 3 stars = It's OK, 2 stars = Fairly Bad, 1 star = Awful. Below the legend, there are links for Home, Manage Buddies, Your Preferences, Help, Publish, and Logout. On the left, there are three search boxes: "Search Titles" (with a "Go!" button), "Search Genres" (with dropdowns for Sci-Fi, 2000s, and All movies, plus a "Search Genres!" button), and "Advanced Search". The central content area displays search results for "Sci-Fi" movies from the "2000s" genre. The results table has columns for Predictions for you, Your Ratings, Movie Information, and Wish List. The results include "Eternal Sunshine of the Spotless Mind" (2004), "X-Men 2 (a.k.a. X2: X-Men United)" (2003), "Donnie Darko" (2001), "Dune (miniseries)" (2000), and "Spider-Man (a.k.a. Spiderman)". Each result row includes a link to the movie's info page and its IMDB page.

The screenshot shows the MovieLens homepage in Mozilla Firefox. The title bar reads "movielens - Mozilla Firefox". The menu bar includes File, Edit, View, Go, Bookmarks, Tools, and Help. The address bar shows the URL "http://movielens.umn.edu/search?searchPhrase=&action=newSearch&hiddenParam=1&genre=". Below the address bar is a toolbar with links to "Mozilla Firebird Help", "User Support Forum", "Plug-in FAQ", "Kayak", "LAKAWA- Lakes Area...", "Yahoo! Calendar - jtr...", and "Slashdot: News for n...". The main content area features the MovieLens logo ("m o v i e l e n s" in red) and the tagline "helping you find the right movies". To the right, a welcome message says "Welcome riedl@cs.umn.edu", "You've rated 206 movies.", and "You're the 31st visitor in the past hour.". A rating scale from 1 to 5 stars is shown with descriptions: ★★★★★ = Must See, ★★★★☆ = Will Enjoy, ★★★☆☆ = It's OK, ★★★☆☆ = Fairly Bad, and ★★★☆☆ = Awful. Below the welcome message are links for Home, Manage Buddies, Your Preferences, Help, Publish, and Logout. The left sidebar contains "Shortcuts" and "Search" sections. The "Search" section includes a "Search Titles" input field with a "Go!" button, a checkbox for "Use selected buddies!", and "Search Genres" dropdowns for Sci-Fi (selected), 2000s, and Domain (All movies). There is also a "Search Genres!" button and a "Use selected buddies!" checkbox. The main content area displays a search result for "Sci-Fi" movies from 2000s. It shows 71 movies sorted by prediction, with genres Sci-Fi, Exclude Genres: None, Dates: 2000s, Domain: All, Format: All, and Language: All. Buttons for Show Printer-Friendly Page, Download Results, and Suggest a Title are available. The results are paginated at Page 1 of 5, with a link to page 2. The results table has columns for Predictions for you, Your Ratings, Movie Information, and Wish List. Each row represents a movie with its title, release year, genres, and a "Wish List" checkbox. The first few rows are: "Eternal Sunshine of the Spotless Mind (2004)" (Sci-Fi, Comedy, Drama, Romance), "X-Men 2 (a.k.a. X2: X-Men United) (2003)" (Action, Adventure, Sci-Fi), "Donnie Darko (2001)" (Drama, Mystery, Romance, Sci-Fi), "Dune (miniseries) (2000)" (Drama, Fantasy, Sci-Fi), and "Spider-Man (a.k.a. Spiderman)". The "Your Ratings" column shows ratings from Not seen to 5 stars.

movieLens - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

mozilla Firebird Help User Support Forum Plug-in FAQ Kayak LAKAWA-Lakes Area... Yahoo! Calendar - jtrn... Slashdot: News for n...

<http://movielens.umn.edu/search?searchPhrase=matrix&titleSearchBtn=Go%21&action=newse>

movieLens
helping you find the *right* movies

Welcome riedl@cs.umn.edu
You've rated 206 movies.
You're the 31st visitor in the past hour.

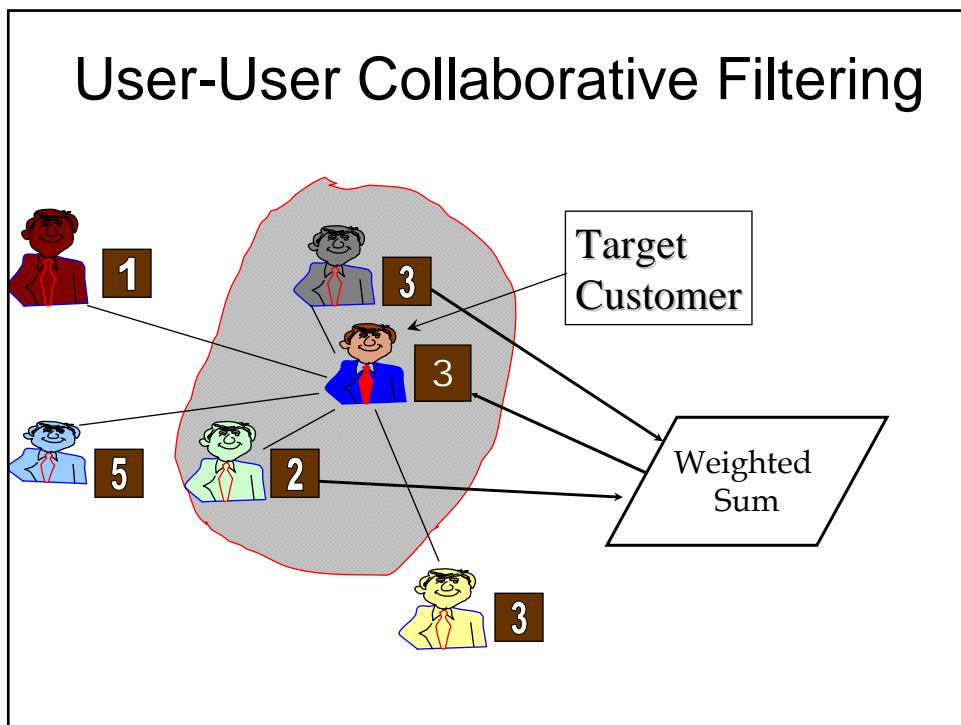
★★★★★ = Must See
★★★★☆ = Will Enjoy
★★★★☆ = It's OK
★★☆☆☆ = Fairly Bad
☆☆☆☆☆ = Awful

[Home](#) | [Manage Buddies](#) | [Your Preferences](#) | [Help](#) | [Publish](#) | [Logout](#)

Shortcuts	Search																
<p>Search Titles <input type="text" value="matrix"/> <input type="button" value="Go!"/> <input type="checkbox"/> Use selected buddies!</p> <p>Search Genres <input type="button" value="All Genres"/> <input type="button" value="All Dates"/> Domain: <input type="button" value="All movies"/> <input type="checkbox"/> Use selected buddies! <input type="button" value="Search Genres"/></p> <p>Advanced Search</p>	<p>You've searched for titles matching: matrix Show Printer-Friendly Page Download Results Suggest a Title</p> <p>Page 1 of 1</p> <table border="1"> <thead> <tr> <th>Predictions for you</th> <th>Your Ratings</th> <th>Movie Information</th> <th>Wish List</th> </tr> </thead> <tbody> <tr> <td></td> <td>4.5 stars <input type="button" value="▼"/></td> <td>Matrix Reloaded, The (2003) DVD, VHS, info imdb Action, Sci-Fi, Thriller</td> <td><input type="checkbox"/></td> </tr> <tr> <td></td> <td>4.0 stars <input type="button" value="▼"/></td> <td>Matrix Revolutions, The (2003) DVD, VHS, info imdb Action, Sci-Fi, Thriller</td> <td><input type="checkbox"/></td> </tr> <tr> <td></td> <td>5.0 stars <input type="button" value="▼"/></td> <td>Matrix, The (1999) DVD, VHS, info imdb Action, Sci-Fi, Thriller</td> <td><input type="checkbox"/></td> </tr> </tbody> </table> <p>Page 1 of 1</p> <p>Create a shortcut to this search! Enter a name then press "Create!" <input type="text"/> <input type="button" value="Create!"/> What are shortcuts?</p>	Predictions for you	Your Ratings	Movie Information	Wish List		4.5 stars <input type="button" value="▼"/>	Matrix Reloaded, The (2003) DVD , VHS , info imdb Action, Sci-Fi, Thriller	<input type="checkbox"/>		4.0 stars <input type="button" value="▼"/>	Matrix Revolutions, The (2003) DVD , VHS , info imdb Action, Sci-Fi, Thriller	<input type="checkbox"/>		5.0 stars <input type="button" value="▼"/>	Matrix, The (1999) DVD , VHS , info imdb Action, Sci-Fi, Thriller	<input type="checkbox"/>
Predictions for you	Your Ratings	Movie Information	Wish List														
	4.5 stars <input type="button" value="▼"/>	Matrix Reloaded, The (2003) DVD , VHS , info imdb Action, Sci-Fi, Thriller	<input type="checkbox"/>														
	4.0 stars <input type="button" value="▼"/>	Matrix Revolutions, The (2003) DVD , VHS , info imdb Action, Sci-Fi, Thriller	<input type="checkbox"/>														
	5.0 stars <input type="button" value="▼"/>	Matrix, The (1999) DVD , VHS , info imdb Action, Sci-Fi, Thriller	<input type="checkbox"/>														

The screenshot shows a Mozilla Firefox browser window displaying the MovieLens website (<http://movielens.umn.edu/search?genreSearch=1&genre>All&format>All&exGenre=None&date>). The search bar at the top contains the query "Joe, Maureen". The results page is titled "Buddy Search with: Joe, Maureen" and indicates "Results restricted to movies that have predictions for each buddy". It shows a table of movie predictions for the user "You" (the current user) and their buddy "Joe" and "Maureen". The table includes columns for Combined Preds, You, Joe, Maureen, Your Ratings, Movie Information, and Wish List.

Combined Preds	You	Joe	Maureen	Your Ratings	Movie Information	Wish List
★★★★★	4.5	4.5	4.5	Not seen	Scrooge (1951) <input type="checkbox"/> DVD, VHS, info imdb Drama, Fantasy	
★★★★★	4.5	4.5	4.5	Not seen	Decade Under the Influence, A (2003) <input type="checkbox"/> DVD, VHS, info imdb Documentary	
★★★★★	4.5	4.5	4.5	Not seen	Miracle Worker, The (1962) <input type="checkbox"/> DVD, VHS, info imdb Drama	
★★★★★	4.5	4.5	4.5	Not seen	Enter the Dragon (1973) <input type="checkbox"/> DVD, VHS, info imdb Action, Crime, Drama	



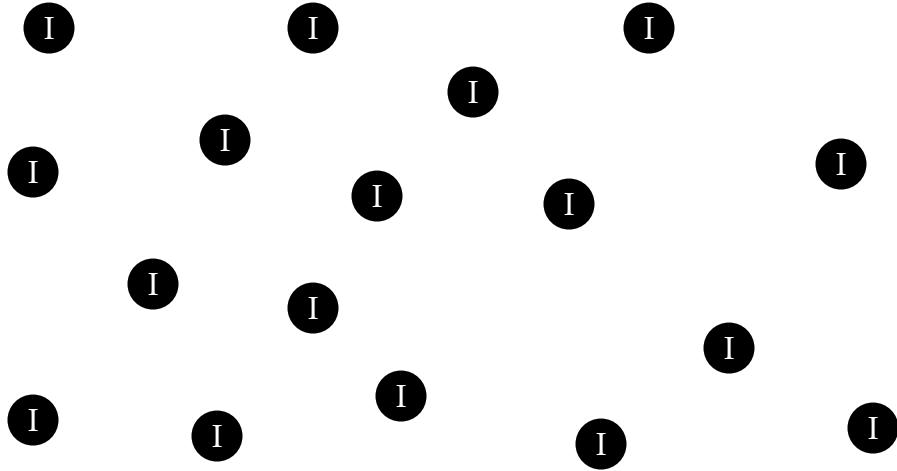
A Challenge: Sparsity

- Many E-commerce and content applications have many more customers than products
- Many customers have no relationship
- Most products have some relationship

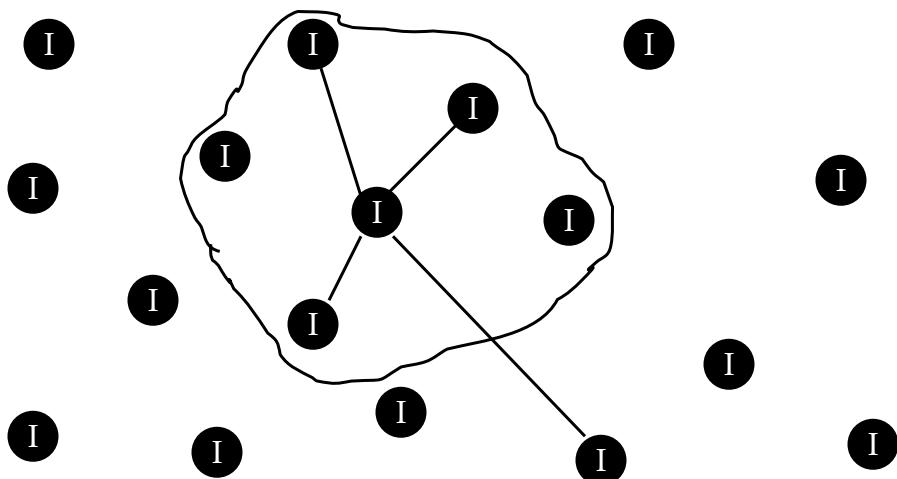
Another challenge: Synonymy

- Similar products treated differently
 - Have skim milk? Want whole milk too?
- Increases apparent sparsity
- Results in poor quality

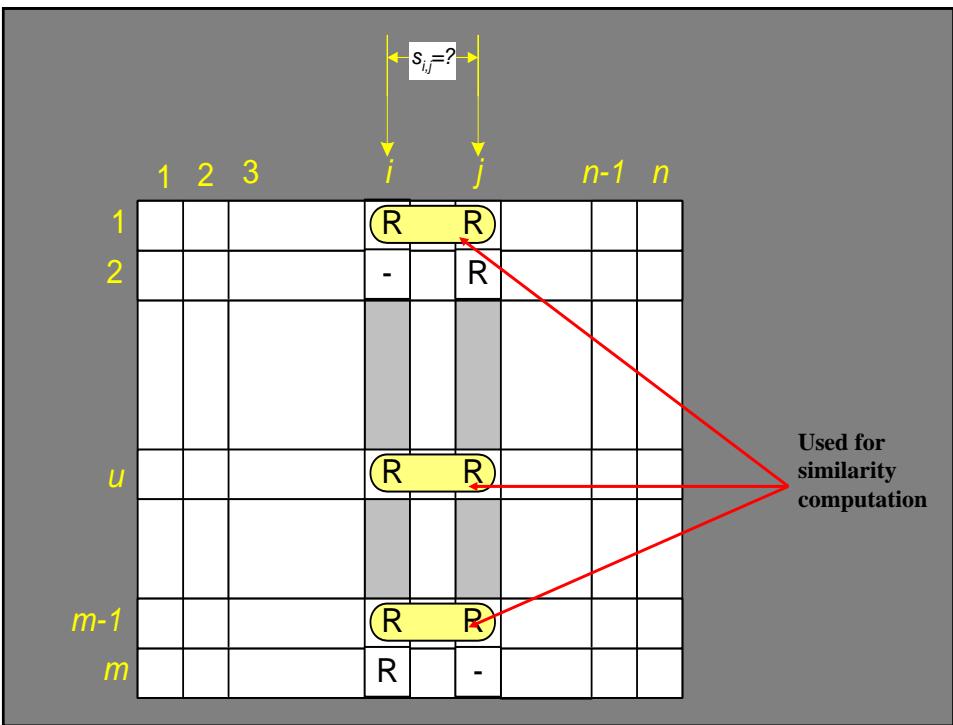
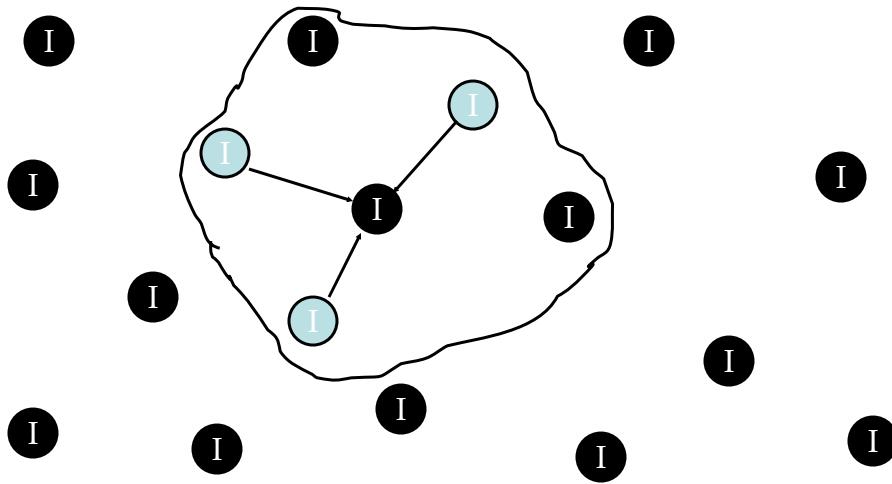
Item-Item Collaborative Filtering

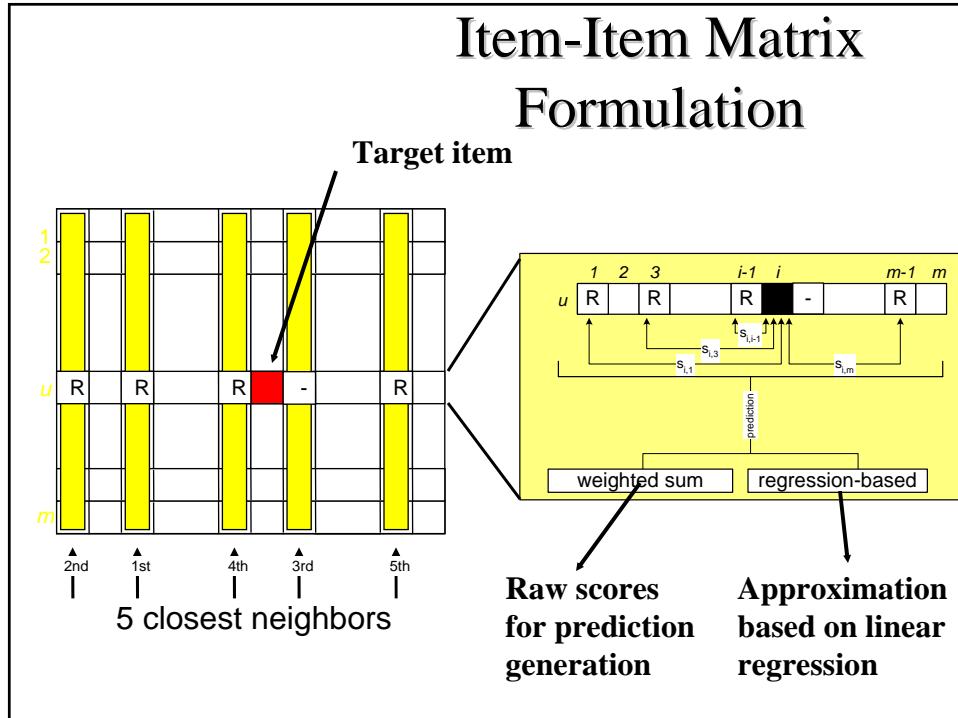


Item-Item Collaborative Filtering



Item-Item Collaborative Filtering





Item-Item Discussion

- Good quality, in sparse situations
- Promising for incremental model building
 - Small quality degradation
 - *Nature* of recommendations changes
 - Big performance gain

Collaborative Filtering Algorithms

- Non-Personalized Summary Statistics
- K-Nearest Neighbor
- Dimensionality Reduction
 - Singular Value Decomposition
 - Factor Analysis
- Content + Collaborative Filtering
- Graph Techniques
- Clustering
- Classifier Learning

Dimensionality Reduction

- Latent Semantic Indexing
 - Used by the IR community
 - Worked well with the vector space model
 - Used Singular Value Decomposition (SVD)
- Main Idea
 - Term-document matching in feature space
 - Captures latent association
 - Reduced space is less noisy

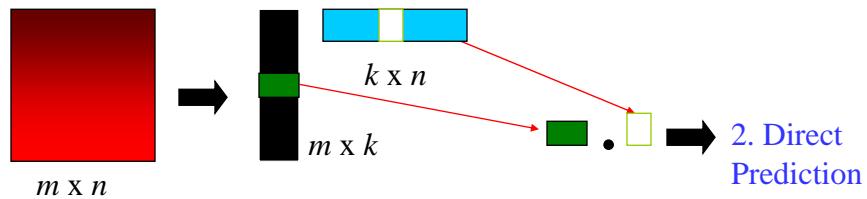
SVD: Mathematical Background

$$\begin{array}{c}
 \boxed{R_k} = \boxed{U U_k} \quad \boxed{S} \quad \boxed{V^* V_k'} \\
 m \times n \qquad m \times k \qquad k \times k \qquad k \times n
 \end{array}$$

The reconstructed matrix $R_k = U_k S_k V_k'$ is the closest *rank-k* matrix to the original matrix R .

SVD for Collaborative Filtering

1. Low dimensional representation
 $O(m+n)$ storage requirement



Singular Value Decomposition

Reduce dimensionality of problem

- Results in small, fast model
- Richer Neighbor Network

Incremental Update

- Folding in
- Model Update

Trend

- Towards use of probabilistic LSI

Collaborative Filtering Algorithms

- Non-Personalized Summary Statistics
- K-Nearest Neighbor
- Dimensionality Reduction
- Content + Collaborative Filtering
- Graph Techniques
 - Horting: Navigate Similarity Graph
- Clustering
- Classifier Learning
 - Rule-Induction Learning
 - Bayesian Belief Networks

Issues in Collaborative Filtering

Research Issues in Collaborative Filtering

- Confidence and Explanation
- Social Navigation
- Recommending for Groups
- Reducing Rating Effort
- New Items and Users
- Evaluation
- Some Challenges

Trust/Acceptance

- Part of the problem is external – eliciting trust in the customer
- Part of the problem is that people don't understand the basis for a recommendation

Some Stories

*Where do users think
MovieLens recommendations
come from?*

Confidence

- Why would someone distrust a recommendation?
 - Can I trust the provider?
 - How does this work, anyway?
 - Does the system know enough about me?
 - Does the system know enough about the item it is recommending?
 - How sure is it?

Trusting the Provider

- Concern about ulterior motives
 - Amazon's pricing experiments
- Concern about external tampering
 - Easier with pseudonyms and rapid sign-up
- Possible approaches
 - Codes of conduct / disclosed policies
 - External auditing of recommenders
 - “Recommender in your pocket”

The Confidence Challenge

- Why should users believe recommendations?
- When should users believe them?
- Approaches
 - Confidence indicators
 - Explain the recommendations
 - Reveal data, process
 - Corroborating data, track record
 - Offer opportunity to correct mistaken data

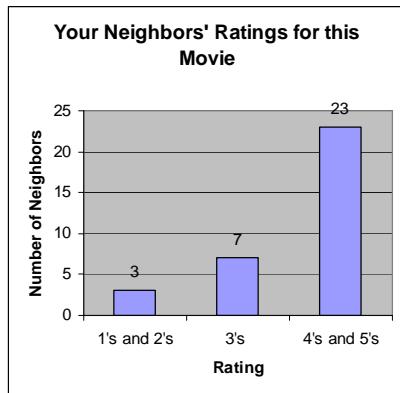
A Simple Confidence Display

Your Top 5 Recommendations	
Recent Box Office	
1. Star Wars: Episode II - Attack of the Clones (2002)	★★★
2. Spider-Man (a.k.a. Spiderman) (2002)	★★★
3. Minority Report (2002)	★★★
4. Insomnia (2002)	★★★
5. Sum of All Fears, The (2002)	★★★
Recent DVDs	
1. Beautiful Mind, A (2001)	★★★★
2. Red Beard (Akahige) (1965)	★★★★½
3. From Hell (2001)	★★★★
4. Traffic (2000)	★★★★
5. Horse's Mouth, The (1958)	★★★★½
Recent Home Videos	

Explanation Studies

- Pilot study of explanation feature
 - Users liked explain
 - Unclear whether they become more effective decision-makers
- Comprehensive study of different explanation approaches
 - Wide variation of effectiveness
 - Some explanations hurt decision-making

Most Compelling Interfaces

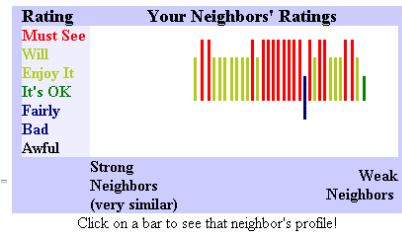


- *Simple visual representations of neighbors ratings*
- Statement of strong previous performance
“MovieLens has predicted correctly 80% of the time for you”

Explanations

Less Compelling Interfaces

Ratings for *Sixth Sense, The* (1999) by your Neighbors



- Anything with even minimal complexity
 - More than two dimensions
- Any use of statistical terminology
 - Correlation, variance, etc.

Explanations

Explanation: Key Lessons

- Persuasion vs. Information
- Complex explanations often don't work
- Users often have wrong mental models
 - We don't sit in a room picking movie recommendations!

Amazon.com Improve Your Recommendations - Mozilla (Build ID: 2008031104)

File Edit View Search Go Bookmarks Tools Help Debug QA

Back Forward Reload Stop http://www.amazon.com/exec/obidos/rg/reccollection-edit/1/104-2814376-0529007?cont_page=recsigned-in=continuous&cont_type=collection

Search All Products: Search | WISHLIST | YOUR ACCOUNT | HELP

amazon.com

WELCOME **YOUR FAVORITE STORES** BOOKS ELECTRONICS DVD & BEAUTY VIDEO COMPUTERS SEE MORE STORES

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

Search All Products: Books

Your Recommendations > Improve Your Recommendations

To exclude an item from being used for your recommendations, uncheck the "Use to make recommendations" option. Remember to save any changes below when you are done making your selections.

EDIT YOUR INFO

Items you own (210)
Items in your wishlist (20)
Items you've rated (122)
Items you've marked not interested (1)

Need Help? Learn more about improving your recommendations.

Show: Items not rated | All items In: All Products

		Not Rated	Your Rating:
1.	<i>Stupid White Men... and Other Sorry Excuses for the State of the Nation</i> by Michael Moore	<input checked="" type="radio"/> ?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
Amazon.com purchase <input type="checkbox"/> Use to make recommendations			
2.	<i>The Nanny Diaries</i> by Emma McLaughlin, Nicola Kraus	<input checked="" type="radio"/> ?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
Amazon.com purchase <input type="checkbox"/> Use to make recommendations			
3.	<i>Atonement</i> by Ian McEwan	<input checked="" type="radio"/> ?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
Amazon.com purchase <input type="checkbox"/> Use to make recommendations			
4.	<i>The Ultimate French Review and Practice</i> by David M. Stillman, Ronni L. Gordon	<input checked="" type="radio"/> ?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
Amazon.com purchase <input type="checkbox"/> Use to make recommendations			
5.	<i>Irish Heartbeat</i> ~ Van Morrison & The Chieftains	<input checked="" type="radio"/> ?	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5
Amazon.com purchase <input type="checkbox"/> Use to make recommendations			

Amazon.com - Recommendations: Why? - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop http://www.amazon.com/exec/obidos/tg/recs/r-why-dvd/0/regular/none/B00003CXR4/0/1/pqrref=pd_ir_d_why/102-5987387-8866549

DVD Recommendations > Why was I recommended this?

We recommended...

[Crouching Tiger, Hidden Dragon](#) DVD
~ Yun-Fat Chow
Average Customer Rating: ★★★★☆
Our Price: \$20.97
[I own it](#) [Not interested](#)

Because you purchased or rated...

Purchased or Rated Items	Not Rated	Your Rating: Don't like it < > I love it!					Exclude Item
The New Joy of Cooking by Irma S. Rombauer, et al	<input checked="" type="radio"/> ?	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5	<input type="checkbox"/>
Cookwise by Shirley O. Corriher	<input checked="" type="radio"/> ?	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5	<input type="checkbox"/>
Essential Talmud by Adin Steinsaltz	<input checked="" type="radio"/> ?	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5	<input type="checkbox"/>
Still Pumped from Using the Mouse by Scott Adams	<input checked="" type="radio"/> ?	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5	<input type="checkbox"/>
Where Wizards Stay Up Late by Katie Hafner, Matthew Lyon	<input checked="" type="radio"/> ?	<input type="radio"/> 1	<input type="radio"/> 2	<input type="radio"/> 3	<input type="radio"/> 4	<input type="radio"/> 5	<input type="checkbox"/>



Issues in Collaborative Filtering

- Confidence and Explanation
- Social Navigation
- Recommending for Groups
- Reducing Rating Effort
- New Items and Users
- Evaluation
- Some Challenges

Social Navigation

- Assumptions:
 - Awareness of others (current or past) helps user find relevant path
 - Paths/location of others is distinctive enough for user to recognize
- Approach
 - Make history or present visible
- Applications
 - Real-World (baggage claim)
 - On-Line (footprints)
 - Virtual (GeoNotes)

[MODIFY this page](#) [To the Top](#) [Recently Modified](#) [Recently Accessed](#)

Recent Changes

14 January 1999

- [Groundhog Day checklist](#)   ...anx3-11.dialup.emory.edu
- [Groundhog Day Logo and T-shirt Design](#)   ...juggler2.cc.emory.edu
- [What are these funny markers for](#)   ...juggler2.cc.emory.edu
- [News](#)   ...170.140.164.100

13 January 1999

- [Groundhog Day 1999](#)   ...209.86.62.75
- [Yo Yo Letter](#)   ...209.86.62.75
- [People needing crash space for GHD99](#)   ...juggler2.cc.emory.edu

12 January 1999

- [Rumors exchange](#)   ...anx52-130.dialup.emory.edu
- [Juggler's Pages](#)   ...205.219.180.52
- [AJA address list](#)   ...pc253.netzip.com
- [Ideas and suggestions for the festival](#)   ...pc253.netzip.com
- [Formatting Rules](#)   ...170.140.37.49
- [What is a Swiki server](#)   ...170.140.37.49

Issues in Collaborative Filtering

- Confidence and Explanation
- Social Navigation
- **Recommending for Groups**
- Reducing Rating Effort
- New Items and Users
- Evaluation
- Some Challenges

Recommending for Groups

- Problem: People watch movies together
- Solution: A recommender for groups
- Issues
 - Group formation, rules, composition
 - Recommender algorithm for groups
 - User interface

Goals

- Explore group recommender design space
- See if users would want and use a group recommender, at least for movies
- Study behavior changes in group members
 - group vs. other users
 - new users via groups vs. other new users
- Learn lessons about group recommenders

Design Issues

- Characteristics of groups
 - public or private
 - many or few
 - permanent or ephemeral
- Formation and evolution of groups
 - joining policy
 - administration and rights

Design Issues

- What is a group recommendation?
 - group user vs. combined individuals
 - social good functions
- Privacy and interface issues
 - control over joining groups
 - withholding and recommendations
 - balancing between info overload and support

PolyLens

- Design choices
 - private, small, administered, invited groups
 - combine individual recs with minimum misery
 - high-information interface with opt-out

Group: Dantest		Back To Individual Recommendations				
TITLE	GENRE	REVIEWS	GROUP	YOUR	cosley@cs.umn.edu	cosley@quasar
Pixote (1981)	Drama	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
Wrong Trousers, The (1993)	Animation, Comedy	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
After Life (1998)	Drama	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★
King of Masks, The (Bian Lian) (1996)	Drama	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★

- External invitations added by popular demand

Field Test Results and Lessons

- Users like and use group recommenders
 - groups have value for all members
 - groups can help with outreach to new members
- Users trade privacy for utility
- Groups are both permanent and ephemeral
- Users must be able to find each other

Subsequent Redesign

- Move from groups to buddies
 - Hybrid permanent/ephemeral
 - Lower overhead for subgroups
- Support e-mail based invitation
 - Viral marketing
 - No need for ML logins
- Remove most privacy settings
 - Buddy model supports person-to-person rather than person-to-group

The screenshot shows a Mozilla Firefox browser window displaying the MovieLens website. The URL in the address bar is <http://movielens.umn.edu/search?genreSearch=1&genre>All&format>All&eGenre=None&date>. The page title is "movielens - Mozilla Firefox".

The interface includes a sidebar with "Shortcuts", "Search", "Search Titles", "Search Genres" (with dropdowns for "All Genres", "All Dates", "Domain: DVD releases", and a checked "Use selected buddies" checkbox), "Advanced Search", and "Select Buddies" (listing "Joe" and "Maureen" with a "What are buddies?" link). The main content area displays a search result titled "Buddy Search with: Joe, Maureen" showing "Results restricted to movies that have predictions for each buddy". It indicates "Page 1 of 301" and provides links to "page 2>" and "1...60...120...180...240...300...last".

A table lists movie recommendations for the user "You" (the current user) and their buddies "Joe" and "Maureen". The columns are "Combined Preds" (rating scale from 1 to 5 stars), "You", "Joe", "Maureen", "Your Ratings" (dropdown menu set to "Not seen"), "Movie Information", and "Wish List" (checkboxes). The table contains four rows of data:

Combined Preds	You	Joe	Maureen	Your Ratings	Movie Information	Wish List
★★★★★	4.5	4.5	4.5	Not seen	Scrooge (1951) DVD, VHS, info imdb Drama, Fantasy	<input type="checkbox"/>
★★★★★	4.5	4.5	4.5	Not seen	Decade Under the Influence, A (2003) DVD, VHS, info imdb Documentary	<input type="checkbox"/>
★★★★★	4.5	4.5	4.5	Not seen	Miracle Worker, The (1962) DVD, VHS, info imdb Drama	<input type="checkbox"/>
★★★★★	4.5	4.5	4.5	Not seen	Enter the Dragon (1973) DVD, VHS, info imdb Action, Crime, Drama	<input type="checkbox"/>

Issues in Collaborative Filtering

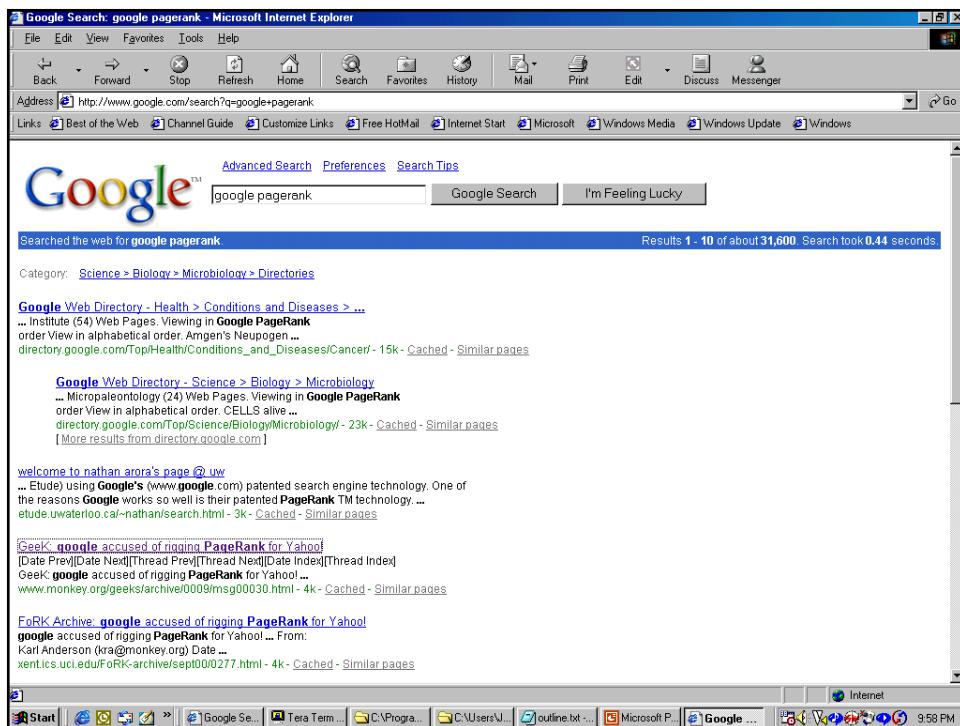
- Confidence and Explanation
- Social Navigation
- Recommending for Groups
- Reducing Rating Effort
- New Items and Users
- Evaluation
- Some Challenges

Overview

- Gather “Work Product” Data
- Less User Effort
- Faster Start-Up
- Potentially More Accurate

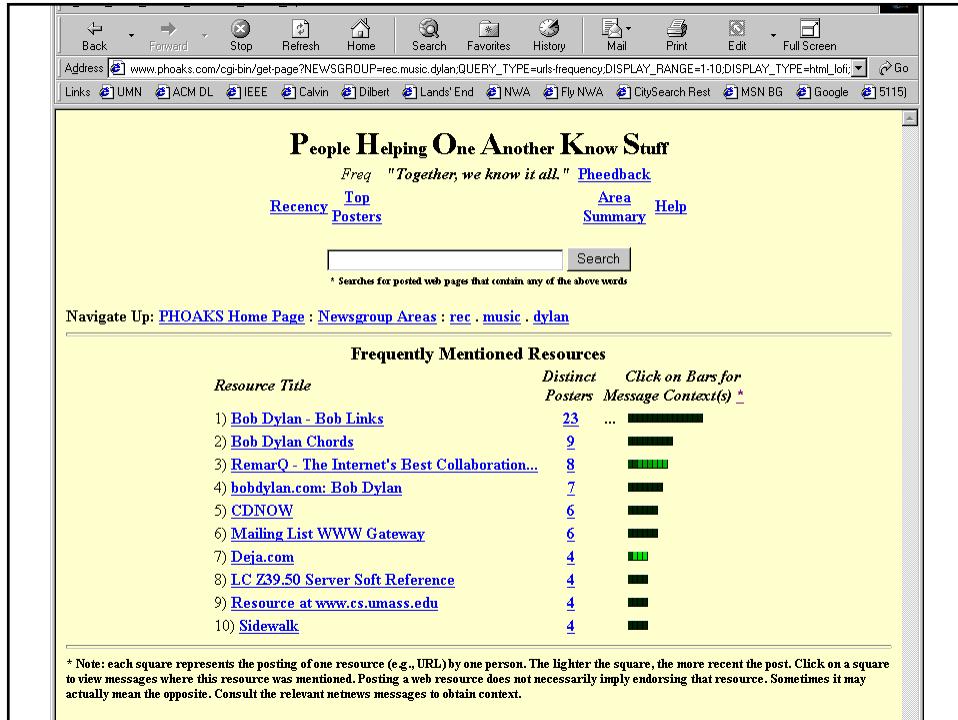
Google PageRank

- Ranks pages based on incoming links
- Links from higher ranked pages matter more
- Combines text analysis with importance to decide which pages to show you
- Runs on network of thousands of PCs!
- Works to be hard to trick (e.g., citation trading)



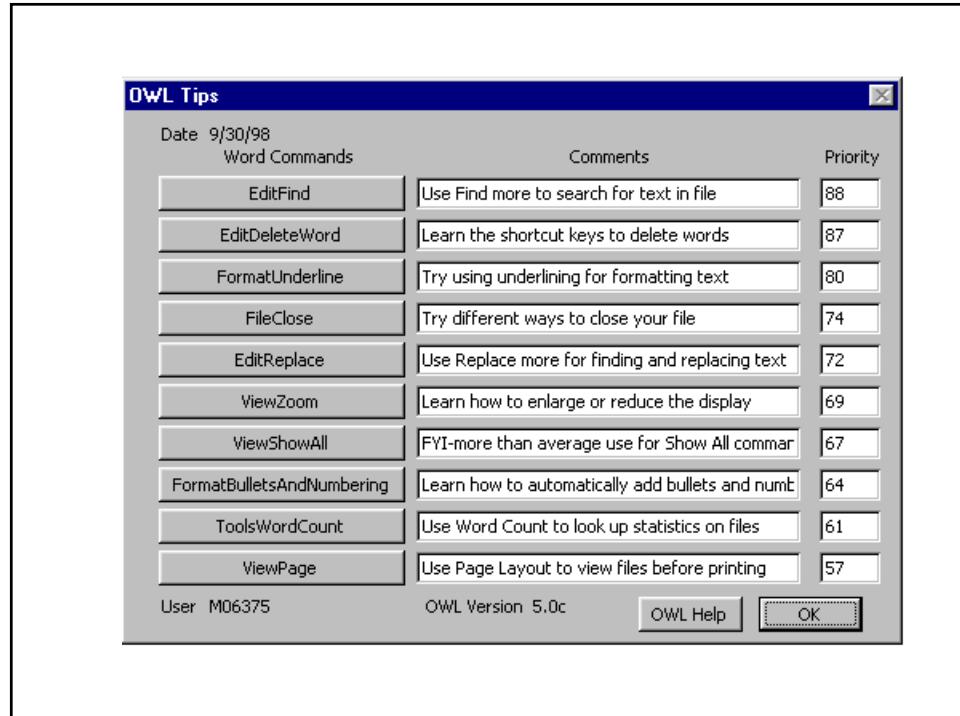
PHOAKS

- Read Usenet news to find web sites!
 - Implicit ratings
 - Filter URLs to find endorsements
 - Create top-n lists of web sites for a Usenet newsgroup community
- Links to endorsements (with age shown)
- Tested against hand-maintained FAQ lists



MITRE's OWL

- Recommender for word processors
 - Monitored word processor command use
 - Identified common patterns of use
 - Recommended commands to learn



Issues in Collaborative Filtering

- Confidence and Explanation
- Social Navigation
- Recommending for Groups
- Reducing Rating Effort
- **New Items and Users**
- Evaluation
- Some Challenges

Recommending New Items?

Collaborative filtering cannot recommend new items: no one has rated them

- Random
- Content analysis
- Filterbots

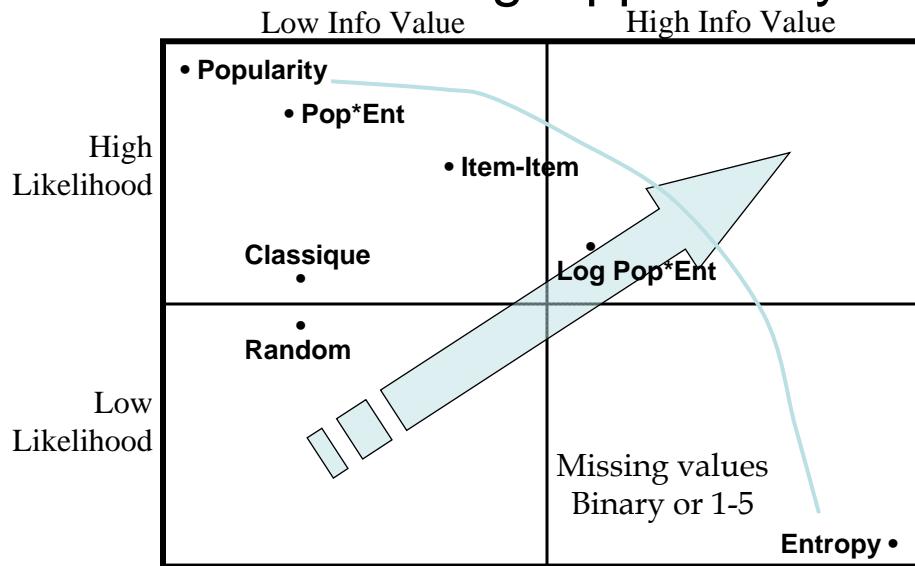
What About New Users?

- Collaborative filtering cannot match new users: they have rated nothing
 - Provide average ratings
 - User agents collect implicit ratings [Wasfi]
 - Put users in categories [Haddaway et al.]
 - Carefully select items for users to rate [Pennock & Horvitz, Kohrs & Merialdo]

Goals for New Users

- User effort
- User satisfaction
- Recommendation accuracy
- System utility

Value of Rating Opportunity

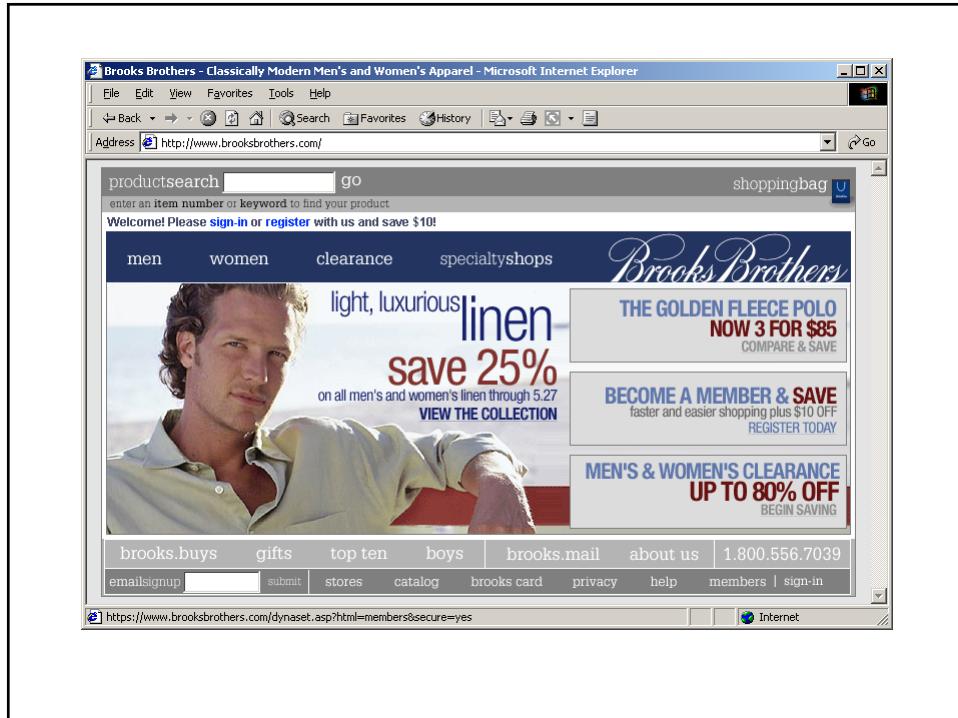


Selling Demographics: DoubleClick

- Ad delivery service
 - Inventory management problem
 - “value” of a page view
- Collect demographic information across sites
 - Use to choose ads to show on new site
 - Cookies on browser
- Consumer resistance
 - Ignored
 - But eventually too expensive to continue

Selling Demographics: Angara

- Service to sell demographics of “new” user to web sites
- Personalize to first-time user
- Claimed substantial improvements in purchase rates
- Out of business in Web bust (merged with Personify)



Selling Demographics: Lessons Learned

- Customers hate cross-site tracking
- Sharing data about millions of customers is expensive
- Effective solutions will require privacy protection and data reduction

Issues in Collaborative Filtering

- Confidence and Explanation
- Social Navigation
- Recommending for Groups
- Reducing Rating Effort
- New Items and Users
- Evaluation
- Some Challenges

Evaluating Recommenders

How do we Evaluate Recommenders -- today?

- Industry outcome
 - Add-on sales
 - Click-through rates

Real-world Experience

- Large international catalog retailer
 - 17% hit rate, 23% acceptance rate in call center
- Medium European outbound call center
 - 17% hit rate, 6.7% acceptance rate from an outbound telemarketing call
 - \$350.00 price of average item sold
 - Items were in an electronics over-stocked category and were sold-out within 3 weeks
- Medium American online toy store (e-mail campaign)
 - 19% click-thru rate vs. 10% industry average
 - 14.3% conversion to sale vs. 2.5% industry average

How do we Evaluate Recommenders -- today?

- Industry outcome
 - Add-on sales
 - Click-through rates
- Research measures
 - User satisfaction
- Metrics
 - To anticipate the above beforehand (offline)

Evaluating Recommendations

- Prediction Accuracy
 - MAE, MSE,
- Decision-Support Accuracy
 - Reversals, ROC
- Recommendation Quality
 - Top-n measures (e.g., Breese score)
- Item-Set Coverage



J. Herlocker et al. Evaluating Collaborative Filtering Recommender Systems. *ACM Transactions on Information Systems* 22(1), Jan. 2004.

What's Wrong with This Approach?

- What is the purpose of recommenders?
 - to help people find things they don't already know – and that they'll like/value/use
 - to serve as a useful advisor
- What are we measuring, mostly?
 - how well the recommenders perform at finding things the users already know
 - performance on individual recommendations

There are Alternatives!

- The “easy” alternative
 - test on real users, real situation
 - have them consume and evaluate
- The “hard” alternative
 - extend our knowledge and understanding about metrics

Extending our Knowledge ...

From Items to Lists

- Do users really experience recommendations in isolation?



C. Ziegler et al. "Improving Recommendation Lists through Topic Diversification", in *Proc. WWW 2005*.

Making Good Lists

- Individually good recommendations do not equal a good recommendation list
- Other factors are important
 - Diversity
 - Affirmation
 - Appropriateness
- Called the “Portfolio Effect”
[Ali and van Stam, 2004]

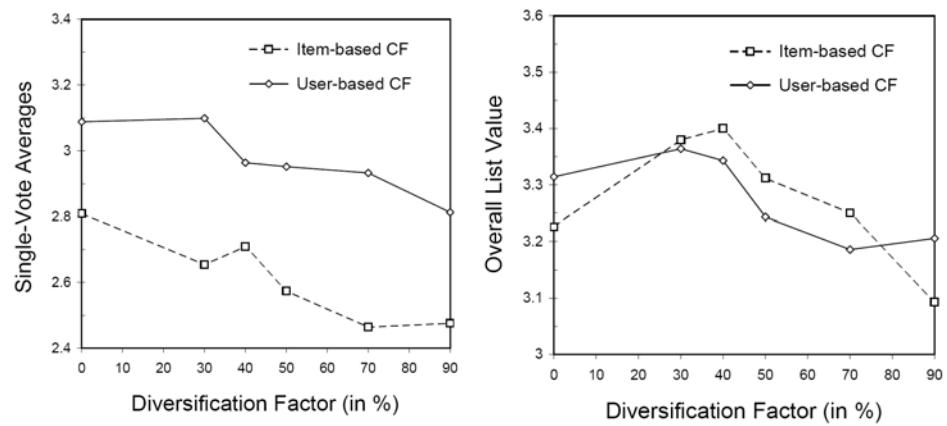
Topic Diversification

- Re-order results in a rec list
- Add item with *least* similarity to all items already on list
- Weight with a ‘diversification factor’
- Ran experiments to test effects

Experimental Design

- Books from BookCrossing.com
- Algorithms
 - Item-based CF
 - User-based CF
- Experiments
 - On-line user surveys
 - 2125 users each saw one list of 10 recommendations

Online Results



Diversity is Important

- User satisfaction more complicated than only accuracy
- List makeup is important to users
- 30% change enough to alter user opinion
- Change not equal across algorithms

Next Steps ...

WARNING: Work in Progress

Human-Recommender Interaction

- Three premises:
 - Users perceive recommendation quality in context; users evaluate lists
 - Users develop opinions of recommenders based on interactions over time
 - Users have an information need and come to a recommender as a part of their information seeking behavior



S. McNee et al. "Making Recommendations Better: An Analytic Model for Human-Recommender Interaction" in *Ext. Abs. CHI 2006*

HRI

- A language for communicating user expectations and system behavior
- A process model for customizing recommenders to user needs
- An analytic theory to help designers focus on user needs

HRI Pillars and Aspects



Recommendation Dialog

- The individual recommender interaction
- Historical Aspects
 - Correctness, Quantity, Spread
- New Aspects
 - Transparency
 - Saliency
 - Serendipity
 - Usefulness
 - Usability

Recommendation Personality

- Experience over repeated interactions
- Nature of recommendations
 - Personalization, Boldness, Freshness, Risk
- Progression over time
 - Adaptability, Pigeonholing
- Relationship
 - Affirmation, Trust

Information-Seeking Task

- One of the current limits of HRI
- Concreteness
- Compromise
- Appropriateness of Recommender
- Role of Recommender
- Expectation of Usefulness

HRI Process Model



- Makes HRI Constructive
 - Links Users/Tasks to Algorithms
- But, Needs New Metrics

Developing New Metrics

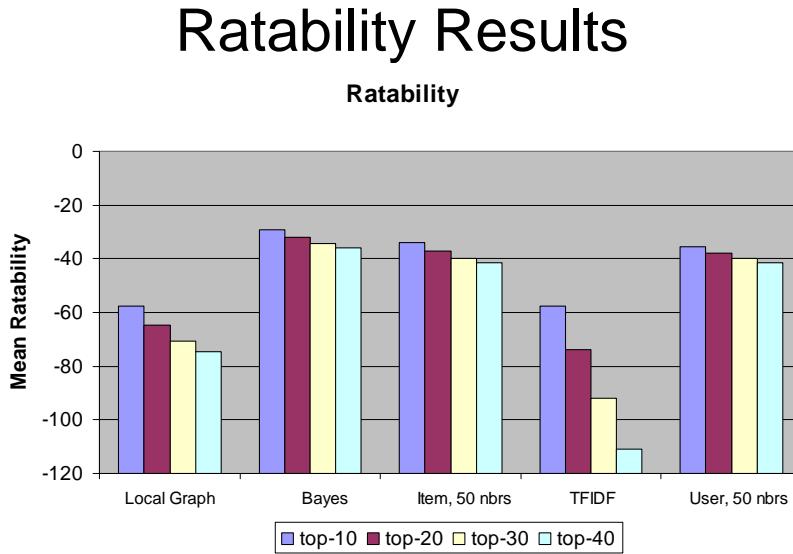
- Identify candidate metrics
- Benchmark a variety of algorithms
 - and datasets?
 - establish that metric can distinguish algorithms
- Establish link to HRI aspects
 - definitional links; user studies
- Detailed Examples:
 - Ratability, Boldness, Adaptability

Metric Experimental Design

- ACM DL Dataset
 - Thanks to ACM!
 - 24,000 papers
 - Have citations, titles, authors, & abstracts
 - High quality
- Algorithms
 - User-based CF
 - Item-based CF
 - Naïve Bayes Classifier
 - TF/IDF Content-based
 - Co-citation
 - Local Graph Search
 - Hybrid variants

Ratability

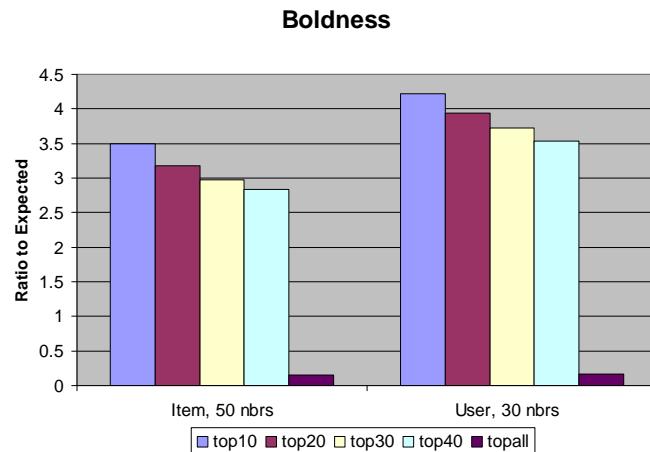
- Probability a user will rate a given item
 - “Obviousness”
 - Based on current user model
 - Independent of liking the item
- Many possible implementations
 - Naïve Bayes Classifier



Boldness

- Measure of “Extreme Predictions”
 - Only defined on explicit rating scale
 - Choose “extreme values”
 - Count appearance of “extremes” and normalize
- For example, MovieLens movie recommender
 - 0.5 to 5.0 star scale, half-star increments
 - Choose 0.5 and 5.0 as “extreme”

Boldness Results

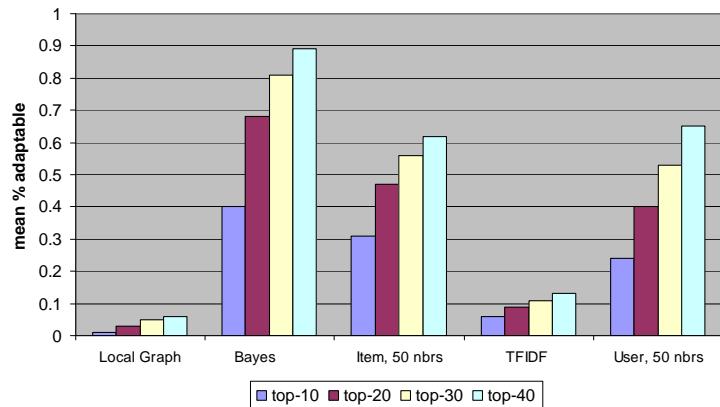


Adaptability

- Measure of how algorithm changes in response to changes in user model
 - How do users grow in the system?
- Perturb a user model with a model from another random user
 - 50% each
 - See quality of new recommendation lists

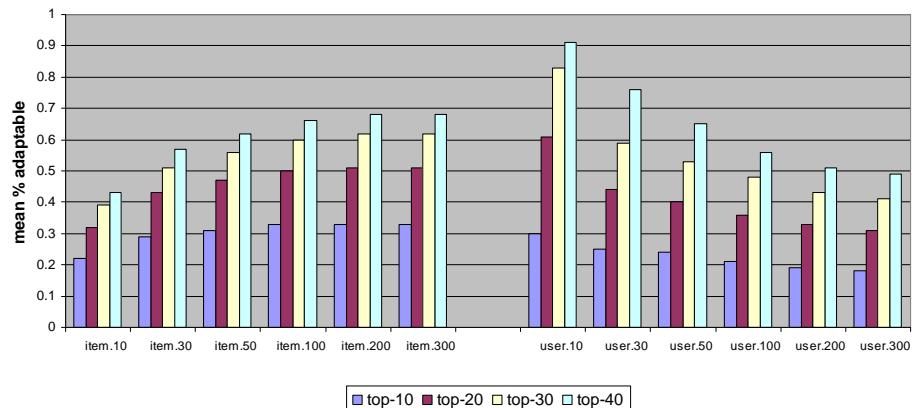
Adaptability Results

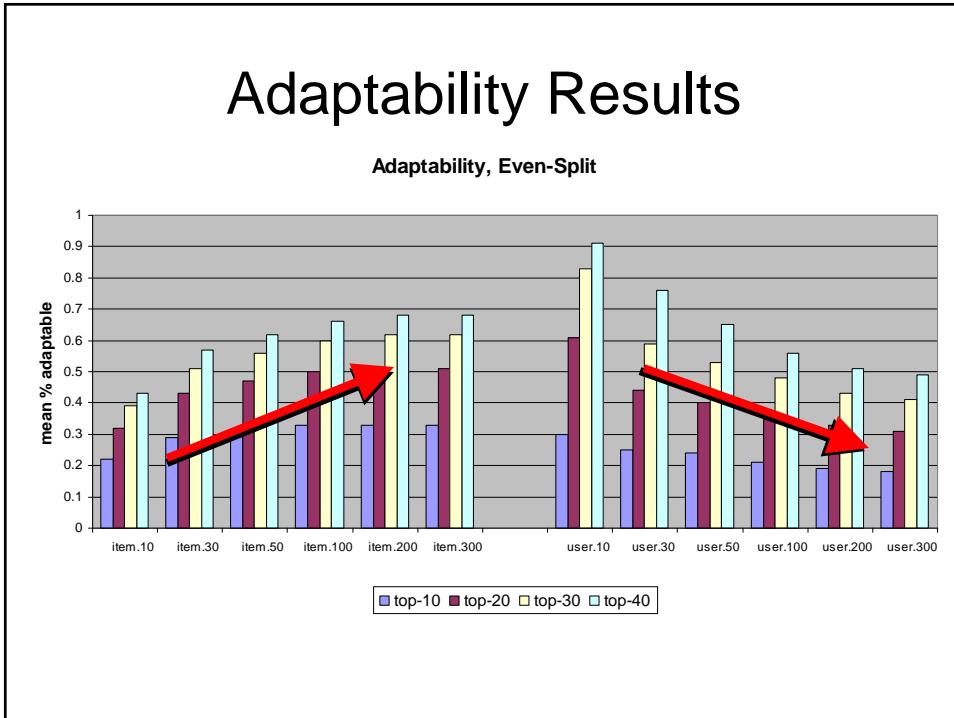
Adaptability, Even-Split



Adaptability Results

Adaptability, Even-Split





Issues in Collaborative Filtering

- Confidence and Explanation
- Social Navigation
- Recommending for Groups
- Reducing Rating Effort
- New Items and Users
- Evaluation
- Some Challenges

Some Challenges

- Diminishing Marginal Returns
 - Why today's recommenders won't be enough ten years from now
- Temporal Recommendation
 - Beyond reacting ...
- Processing Transaction Streams

CF Under Diminishing Returns

- Original goal of CF was to help people sift through the junk to find the good stuff.
- Today, there may be so much good stuff that you need to sift even more.
- Certain types of content yield diminishing returns, even with high quality

Portfolios of Content

- What if my recommender knows which articles I've read, and can identify articles by topic?
- What if it sees that I experience marginal returns from reading similar articles on a topic?
- Could we downgrade some articles based on "lack of new content?" Could we discover which articles using collaborative filtering?

Deeper Criteria than Accuracy

- Understanding user tasks and needs in context
 - Reassurance
 - Novelty
 - Risk Tolerance
 - Set Usage

Temporal Collaborative Filtering

- Today's CF systems may expire or degrade ratings, but do little to detect or predict changes in preference.
- Ripe area with lots of commercial applications ...

Wine for the Time

- Evolving taste – can we help a wine newcomer build her palate? Could we identify wines that take her a step or two beyond her current ones? Can we do so by augmenting regular collaborative filtering with temporal models?

Processing Transaction Streams

- Leaving privacy aside for a minute ...
 - Imagine you're a large credit-card processor (or other large transaction processor), who is trying to provide enhanced service through recommendation ...
 - What can you do to each transaction when it comes through to build more effective user models, product models, etc.?

Conclusions, Resources,
Discussion

Conclusions

- Broad area of research and practice
 - across HCI, IR, Machine Learning, Data Mining
- Simple techniques work fairly well
 - the key challenges are often scale and sparsity
- Complex design issues

Resources

- Where are papers published?
 - CSCW, SIGIR, CHI, AAAI, UM/AH, ICDM, and many other venues
 - ACM Recommender Systems 2008!! (Lausanne)
 - recsys.acm.org
- Bibliography of relevant work
 - A fairly good one (though 3-4 years old) by Vaclav Petricek (University College London); search Google for ‘recommender systems annotated bibliography’
 - Start of one at www.grouplens.org

Resources

- Software to try this yourself
 - COFE - <http://eecs.oregonstate.edu/iis/CoFE/>
 - SUGGEST -
<http://glaros.dtc.umn.edu/gkhome/suggest/overview>
 - Taste - <http://taste.sourceforge.net/>
 - WikiLens – <http://www.wikilens.org/>
- Data Sets (www.grouplens.org)
 - MovieLens
 - Jester
 - BookCrossing

Discussion

Extra Slides as Time Permits

Nine Principles for
Recommender Application
Design

Be a Product Expert

The screenshot shows a Microsoft Internet Explorer window with the following details:

- Title Bar:** ticketmaster.com - Microsoft Internet Explorer
- Address Bar:** Address [E] <http://ticketing.ticketmaster.com/cgi-bin/weblix/weblix.drl?ReserveSeats>
- Content Area:**
 - You Have Five Minutes To Complete This Order.**
 - BEST AVAILABLE TICKETS**
 - Event:** MINNESOTA_WILD
XCEL ENERGY CENTER
Friday, September 29, 2000 7:00PM
 - Area:** END OF ICE THAT OPPONENT SHOOTS AT TWICE
PRICE LEVEL 4
LOWER LEVEL SEATING
 - Selected Tickets:** SECTION 103, ROW 26, SEATS 8 TO 9
 - Seating Chart or Map:** [Click Here](#) (Chart Will Open In A New Browser)
 - Selected Pricing:**

Tic Qty	Type	Fare Value per ticket	Ticketmaster Convenience Charge per ticket	Building Facility Charge	Fare Value Total
2	A	64.00	3.75	0.00	67.75
Subtotal		128.00	7.50	0.00	135.00
Ticketmaster Handling Charge					3.00
Total Charge					138.00

 - Delivery Method:** Will-Call
 - Question:** Do you want to purchase these tickets? Yes No
 - BILLING INFORMATION**

Select the areas in Minneapolis - St. Paul where you'd be willing to stay. priceline will search for a hotel at your price in the area(s) you select. Remember, you'll always stay in a nationally recognized name-brand or well-known independent hotel.

- Airport (MSP) - Mall of America [Details](#)
- Downtown Minneapolis [Details](#)
- Eagan [Details](#)
- Minneapolis North [Details](#)
- Minneapolis West [Details](#)
- Roseville [Details](#)
- Southwest Minneapolis [Details](#)
- St. Paul [Details](#)
- University - Metrodome [Details](#)
- Woodbury [Details](#)
- Select All Zones

See full area map below

Improve your chances choose 2 or more areas

NEXT >

The areas in Minneapolis - St. Paul are highlighted in the map below. You can click on an area for a closer look.

The following price guidelines show you what you might pay if you did not book your hotel room through priceline. The range of rates shown below should be used as a general guideline only. Actual available rates may be higher or lower than the ranges stated. Factors such as holidays, major events, and the day of the week can all influence rates. Remember, at priceline we want you to save by getting your hotel room(s) at the best price of all - yours!

Star Level	Rating	Average rates that you would expect to pay in Minneapolis without using priceline
1 Star	Economy	\$58 - \$72
2 Star	Moderate	\$86 - \$105
3 Star	Upscale	\$122 - \$145
4 Star	Deluxe	\$121 - \$151

Here are the 4 easy steps we recommend to name your own price:

1. Shop around for the lowest available hotel rates before you name your price.
2. If the lowest rate you find fits your budget in a hotel you want to stay in, we recommend that you reserve those rooms, since rates and availability change constantly.
3. If not, let priceline try to find available high quality hotel rooms for you, at the price you want to pay.
4. Remember, the more reasonable your price, the better your chance of getting hotel rooms through priceline.com

Priceline.com Incorporated is a seller of travel in the State of California (CST 2040500-50). Registration as a seller of travel does not constitute approval by the State of California.

1 STAR Hotels: Economy
Priceline's 1 Star hotels are national name-brand and well known independent hotels. These hotels are generally 2-4 stories high and are located near major attractions or convenient intersections, often in close vicinity to restaurants and business/shopping centers. [Read](#)

2 STAR Hotels: Moderate
Priceline's 2 Star hotels are national name-brand and well known independent hotels with a reputation for offering a consistent level of quality and amenities. Participation hotels are well-regarded, with high

The screenshot shows a Microsoft Internet Explorer window displaying the Sees Candies website. The URL in the address bar is <http://www.sees.com/Prod.chm?CatGroup=00070058>. The page title is "See's Candies, Inc. - Microsoft Internet Explorer". The main navigation menu includes File, Edit, View, Favorites, Tools, Help, Home, Shop Locations, Free Catalog, Quantity Discount, Fund Raising, Frequent Questions, and About See's.

The page content is for "Milk Chocolates". It features a large image of a 3 lb. box of Milk Chocolates with two layers. A descriptive text block states: "A classic selection of light and delicate milk chocolate coatings over cream, chew, and nut centers." Below this, a caption reads: "Pictured above is our 3 lb. box (two layers)".

A table lists three gift options:

Gift#	Description	Price	Qty	Action
326	1 lb. Milk Chocolates	\$11.70	<input type="text" value="0"/>	Add To Bag
327	2 lb. Milk Chocolates	\$23.40	<input type="text" value="0"/>	
328	3 lb. Milk Chocolates	\$35.10	<input type="text" value="0"/>	

On the left side of the page, there are two columns of candy categories:

- Summer Favorites:** Peanut Brittle, Summer Variety, Almond Almond, Extra Fancy Mixed Salted Nuts, Fudge, Gourmet Lollipops, Lollipop Tins, Lollipop Gift Box, Little Poppet, Coffee Candy.
- Chocolate Assortments:** Assorted Chocolates, Nut & Chees, Milk Chocolates, Dark Chocolates, Chocolate & Variety, Summer Variety, Soft Centers, Milk Chocolate Soft Centers, Pecan Buds, Milk Chocolate Praline [1lb], Dark Chocolate Praline [1lb], Milk Chocolate Nut & Chees, Milk Chocolate Butterscotch, Milk Chocolate Peanuts, Almond Butter, Milk Chocolate Butter, California Brittle.

Fancy Assortments: Gift of Books, Gold & Silver Truffles, 1 Pound Gold Fancy Box, 2 Pound Gold Fancy Box, Blue & Gold Tin.

Gourmet Lollipops: Gourmet Lollipops, Lollipop Tins, Lollipop Gift, Little Poppet.

Be a Customer Agent

The screenshot shows a Microsoft Internet Explorer window displaying a Clinique survey titled 'which products are right for you?'. The page includes a navigation bar with links to CLINIQUE, shop online, world of clinique, for men only, site map, and sign in/my profile. A sub-header reads 'We can recommend skin care and colour just for you. First please tell us a little about yourself. So we can determine your skin type. If this is your first consultation, your answers will be remembered and used to provide you with personalized information on products custom fit for you.' Below this, instructions say 'Please select the choice that describes you. Not Sure? Click the ? buttons for help.' and 'If you are a registered user and you have already taken a personal consultation and would like to call up your information before you have another consultation please [click here to sign in](#)'. A note at the bottom says 'Please select an answer for each question and then click the "tell me my skin type" button.'

The survey consists of a grid of questions with radio button options. The columns represent different skin characteristics:

- COLOUR OF EYES:
 - BLUE
 - SOFT GOLDEN
 - HAZEL
 - BROWN OR BLACK
- NATURAL COLOUR OF HAIR:
 - BLOND
 - RED
 - LIGHT BROWN
 - DARK BROWN OR BLACK
- RESPONSE TO SUN WITHOUT SUN PROTECTION:
 - ALWAYS BURNS
 - THEN TANGS
 - TANS
 - TANS EASILY
- SKIN COLOUR WITHOUT SUNTAN:
 - VERY FAIR
 - MILDLY FAIR
 - MEDIUM
 - DARK OR OLIVE
- PORE SIZE:
 - INVISIBLE
 - SMALL
 - MEDIUM
 - LARGE
- TENDENCY TO BREAKOUT:
 - NEVER
 - RARELY
 - OCCASIONALLY
 - FREQUENTLY
- OILONESS VERSUS DRYNESS:
 - DRY OR QUITE DRY
 - ONLY IN AREAS
 - OILY
 - VERY OILY
- SURFACE FACIAL LINES:
 - MANY
 - SEVERAL
 - FEW
 - VERY FEW

Many Touchpoints, One Business

NORTHWEST AIRLINES

Travel Planner

View Travel Reservations

You may view your travel reservations prior to travel dates and up to 24 hours after a trip is complete.

Simply enter your Reservation Confirmation Number and the passenger's last name, then select 'Travel Reservations' to view the details of your current travel plans.

Confirmation Number (6 alpha numeric characters):

Passenger's Last Name:

Select 'Travel Reservations' to view the details of your existing travel plans.

Travel Reservations

PROMOTIONAL EMAIL SUBSCRIBE UNSUBSCRIBE

©2000 NORTHWEST AIRLINES

Hertz Rates & Reservations

Select one of the following Secure options:

General Rates & Reservations
 Hertz #1 Club Members
 Hertz #1 Club Gold Members
 Modify/Cancel Existing Reservation

Submit

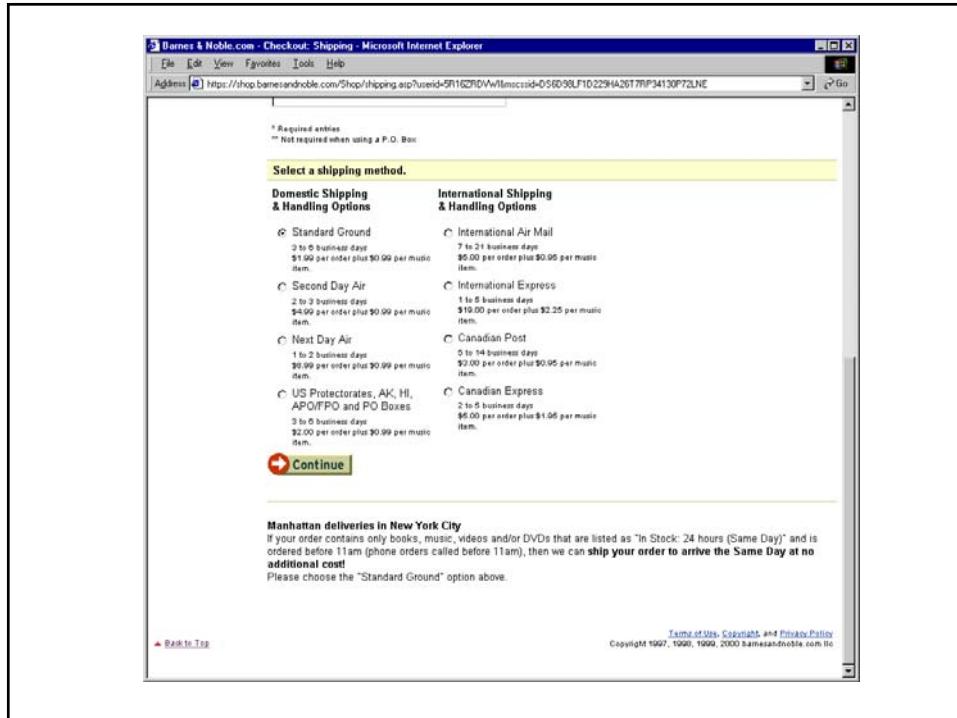
A valid major credit card number is required with all reservations. Debit Cards are not acceptable for confirming reservations.

Please refer to our policy information in the Policies & Procedures section for details on:

- rental qualifications and requirements
- acceptable forms of payment
- airport fees, facility fees, vehicle licensing fees, etc.
- age requirements, especially if you are under 25 years of age

If you have questions regarding online reservations, policies and procedures, or programs and services, please review our list of "Frequently Asked Questions."

Hertz strongly discourages fictitious or duplicate reservations. If Hertz establishes that an individual user has submitted and confirmed multiple

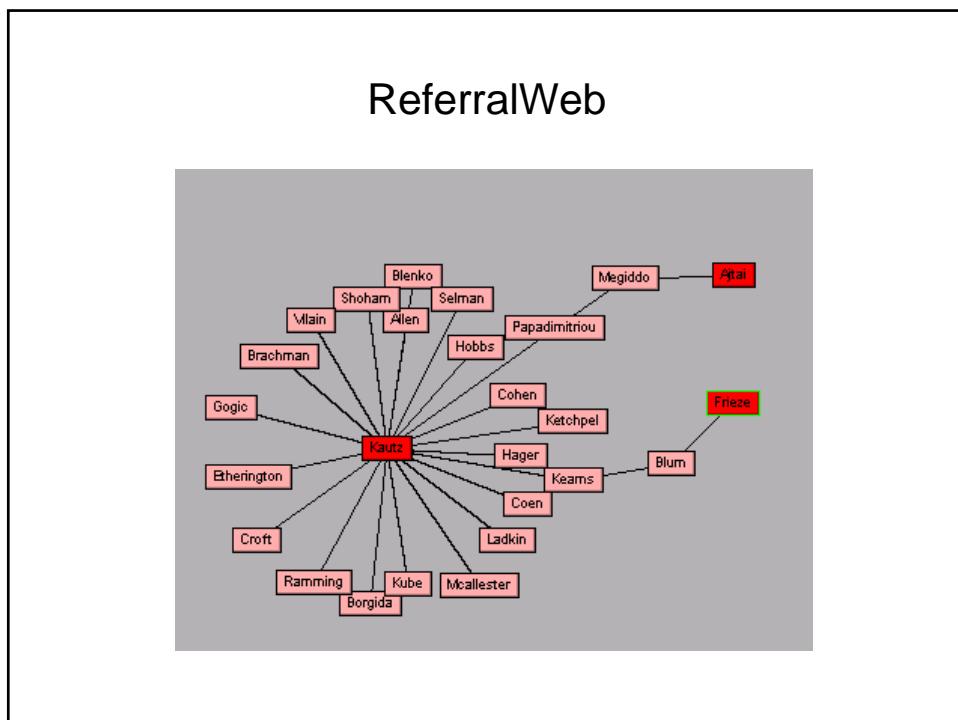


Box Products, Not People



Watch What I Do, Not What I Say

A screenshot of a Microsoft Internet Explorer browser window displaying the Tacit Knowledge Systems website. The page features a blue header bar with the company name and a navigation menu. Below the header is a logo consisting of a grid of colored squares. The main content area includes a banner image showing a computer screen with multiple windows open, labeled 'KnowledgeMail™ Fully Automated Knowledge Discovery and Exchange'. To the right of the banner are sections for 'Top Stories' (with links to news items), 'Quotes' (with a quote from Esther Dyson), and 'Events' (listing web seminars and conferences). At the bottom of the page is a footer with contact information and legal disclaimers.



Click-by-Click Marketing



Customers who bought this item also bought

Information Retrieval: Algorithms and Heuristics (The Information Retrieval Series) (2nd Edition) by David A. Grossman	Information Retrieval: Data Structures and Algorithms by William B. Frakes	Managing Gigabytes: Compressing and Indexing Documents and Images (The Morgan Kaufmann Series in Multimedia and Information Systems) by Ian H. Witten	Mining the Web: Discovering Knowledge from Hypertext Data by Soumen Chakrabarti	The Geometry of Information Retrieval by C. J. van Rijsbergen
\$25.48	\$69.66	\$46.96	\$65.95	\$45.60

[Explore similar items: Books \(22\)](#)

Editorial Reviews

David D. Lewis, AT&T Labs
 This is a marvelous collection. The papers have been selected with care, and provide an unprecedented concentration of knowledge about information retrieval. The introductory material, by two of the leading researchers in IR, is itself a valuable reference to the history and status of the field and the important ideas in it. This book will instantly become the most important reference in the field.

Review
 "This is a marvelous collection. The papers have been selected with care, and provide an unprecedented concentration of knowledge about information retrieval. The introductory material, by two of the leading researchers in IR, is itself a valuable reference to the history and status of the field and the important ideas in it. This book will instantly become the most important reference in the field."
 --David D. Lewis, AT&T Labs

[See all Editorial Reviews](#)

J_konstan Front Page

Handy Tools

Yahoo Search

Portfolios

- DIA 10765.52 +77.60
- NASDAQ 3828.87 -68.57
- S&P 1449.03 -2.29
- NYSE 13 3/8 +5/8
- BRKA 55400 -700

Scoreboard

TODAY

MLB	White Sox vs. Minnesota	0	top 3rd
-----	-------------------------	---	---------

YESTERDAY

MLB	Texas vs. Minnesota	6	12th
-----	---------------------	---	------

Weather

St. Paul, MN 39...57 F

click on city for extended forecast

Top Stories from AP

- Gore, Bush Tangle Over Releasing Oil
- Brit Cops Find Grenade Launcher
- Accuser Sues Clinton, First Lady
- Veterans Affairs Computers At Risk
- Fordstone, Ford Had Wrong Pressure

US News from AP

- Gore Urges Clinton on Oil Reserve
- Tornado Hits Okla. Town and Kills 1
- Crimers Keep Many Blacks From Voting
- Veterans Affairs Computers At Risk
- Alaska Oil Payout: \$2K Per Person

Politics from NPR (Requires Real Player)

- Gore, Clinton Note Hispanic Gains
- Bush Attacks Gore on Character
- Prescription Drug Ads
- Presidential Campaigns Turn Ugly
- Whitewater Counsel Clears Clintons of Wrongdoing

World from Reuters

- Yugoslav PM Deals Blow to Serb Opposition Hopes
- London MI6 Spy HQ Hit by Missile
- Lone Air France Concorde Flies Home
- Russia Police Prepare Hostage Rescue if Talks Fail

Community Filtering

The screenshot shows a Microsoft Internet Explorer window displaying the Expedia.com website. The title bar reads "Fare List - Microsoft Internet Explorer". The main content area shows a list of flight deals between Minneapolis, MN (MSP) and Amsterdam, Netherlands (AMS-Schiphol). The deals are sorted by price. Each entry includes the departure/return date, length of trip, airline, and fare amount. A legend indicates that a yellow star next to a deal means it's the best deal found by other customers for that route.

Departure/Return	Length	Airline	Lowest fare
Wed 1-Nov / Wed 8-Nov	7 days	Icelandair	\$463 ★
Wed 15-Nov / Wed 29-Nov	14 days	Icelandair	\$463 ★
Wed 13-Dec / Thu 21-Dec	8 days	Icelandair	\$463 ★
Thu 14-Dec / Thu 21-Dec	7 days	Icelandair	\$463 ★
Wed 18-Oct / Mon 23-Oct	5 days	Northwest	\$472 ★
Thu 19-Oct / Mon 23-Oct	4 days	Northwest	\$472 ★
Tue 24-Oct / Tue 31-Oct	7 days	Northwest	\$472 ★
Tue 31-Oct / Thu 9-Nov	9 days	Northwest	\$473 ★
Wed 22-Nov / Tue 28-Nov	6 days	Northwest	\$473 ★
Mon 11-Dec / Mon 18-Dec	7 days	Northwest	\$473 ★
Tue 12-Dec / Mon 19-Dec	6 days	Northwest	\$473 ★

The screenshot shows the eBay User Feedback page for user ipc2001. The top section displays the overall profile makeup, showing 16 positives, 15 unique users, and a rating of 4 stars. Below this is the eBay ID card, which includes a summary of most recent comments and a table of feedback counts for the past 7 days, month, and 6 months. The table shows:

	Past 7 days	Past month	Past 6 mo.
Positive	0	1	8
Neutral	0	0	0
Negative	0	0	0
Total	0	1	8
Bad Retractions	0	0	0

Below the feedback table, there is a note about leaving feedback and a link to the Feedback Forum. It also mentions that if you are ipc2001 (15) ★, you can respond to comments. The bottom section lists 16 recent comments from various users, each with a date, item number, and a short excerpt of the praise.

The screenshot shows a Zagat.com review for the restaurant LA BELLE VIE. The top part of the page displays the restaurant's name and address: Out of Town, 312 S. Main St. (Nelson St.), Stillwater, MN, 55082-5169 (651) 430-3545. Below this is a brief description: "Supercreative" French-Mediterranean cuisine courtesy of two former D'Amico Cucina chefs earns this "charming" Stillwater newcomer enthusiastic foodie support, and the "historic" storefront setting is an eye-pleaser; while some speak only of "great promise" and opine it "could be outstanding," converts are convinced it's already a "very special" "escape" that's well "worth the drive." There are links to "VOTE ON THIS RESTAURANT" and "find similar restaurants". The "features list" on the left includes: Credit Cards Accepted, Game & Season, Neighborhood Newcomers, Open Sunday, Party Spaces, Quiet Conversation, Romantic Spots, Visitors on Expense Accounts, Wine/Bier Only, Winning Wine Lists, Worth a Trip. The "find similar restaurants" section allows users to refine their search by rating (26 or above, 25 or above, 23 or above, 21 or below), cost (\$43 or below), neighborhood (Out of Town), and cuisine (French). A button labeled "show me more like this" is also present.

Community Creates Content

The screenshot shows a Microsoft Internet Explorer window displaying the Epinions.com website. The URL in the address bar is http://www.epinions.com/electronics/photo/cameras/all-olympus-digital-olympus_digital_cameras. The page title is "Reviews of Olympus Digital Cameras - Microsoft Internet Explorer".

The main content area displays several reviews for Olympus cameras:

- Olympus D-500L** by [mikarsh](#) (Nov 6 '99)
Product rating: The camera I have is the now discontinued 500L. It's a 1.3 megapixel camera, but takes photos at a lower resolution, maxing about at 1024x768....
- This makes owning a digital camera as much fun as I thought it would be** by [jmcnealy](#) (Oct 5 '99)
Product rating: I won't kid you: I'm no photographic expert-- digital, analog or otherwise. But I love monkeying with my web site, enjoy chronicling travel and other....
- It's no Point and Shoot...** by [Dale](#) (Sep 10 '00)
Product rating: I've had this camera since early 1999. It was the perfect replacement for my Non-Zoom Kodak Digital camera. I instantly saw superior results in...
- Olympus D-600L** by [PCPhotoReview](#)
Olympus has made a quantum leap with a new level of consumer Digital Still camera performance with the D-600L...
- Olympus C-2000** by [Imagine Resource](#)
We've now had the opportunity to test a full-production model of the Olympus C-2000, and have found its image quality to be absolutely first rate....
- Olympus C-2000Z** by [Steve's Danglers](#)
The Olympus C-2000Z is a 2.1-megapixel digital camera capable of capturing SHQ (Super High Quality) 1600 x 1200 pixel images in standard JPEG compressed or TIFF uncompressed format....

On the right side of the page, there are two large advertisements:

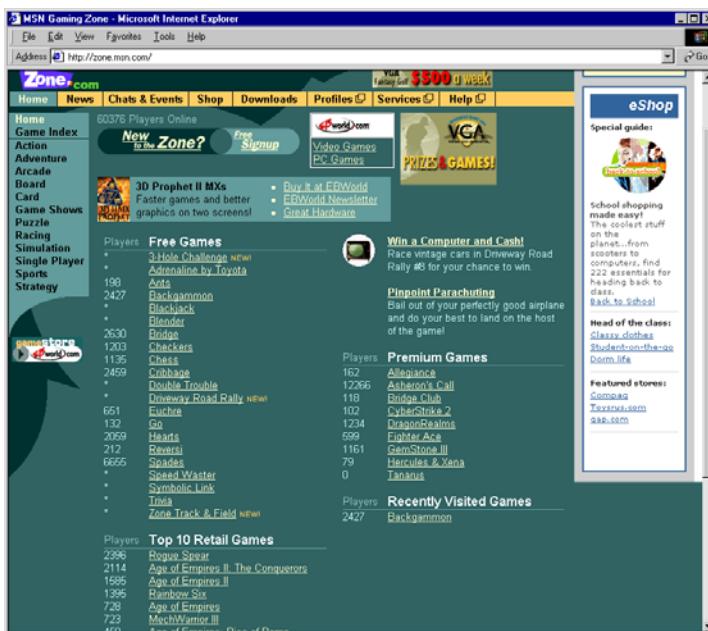
- buy.com** Panasonic \$779.99 CLICK HERE
- BUILD YOUR OWN SERVER MONSTER** CI HOST Click To Start

At the bottom of the page, there are links for "Read all 54 Member Opinions" and "Suggest a Link".

The screenshot shows a web browser window with the URL http://en.wikipedia.org/wiki/Information_retrieval. The page title is "Information retrieval". The content discusses the science of searching for information in documents, mentioning automated systems, public libraries, and search engines like Google, Yahoo, and Live.com. A sidebar on the left contains links for navigation, interaction, and search, along with a toolbox and a contents table.

The screenshot shows a web browser window with the URL <http://www.youtube.com/>. The page features a "Videos being watched right now..." section with thumbnail images. Below it is a "Promoted Videos" section and a "Featured Videos" section. A large video player on the right is playing a video titled "iPhone Magic". A sidebar on the right includes a "Member Login" form and a "UpStage by Samsung Top 20 Lip-Sync Finalists" section.

Community is Content



The screenshot shows the Facebook homepage as it appeared in 2008. The interface includes a login form on the left with fields for Email and Password, and links for Register, Site Tour, and Find Friends. The main area features a News Feed with several posts: one from Raquel Sampson, a photo from Avraham Callahan, and a photo from Ashley Sculley. Below the news feed are sections for Requests (1 friend request) and Notifications (3 event invitations, 1 group invitation). A sidebar on the right titled "Site Tour" provides links to Getting Started, Your Profile, Photos and Notes, Your Home Page, and Your Privacy Settings. A large green "Register" button is also present in the sidebar.

The screenshot shows the Meetup.com homepage. The top features the Meetup logo and the tagline "Whatever your interest. Wherever you are." Below this, there's a section titled "Some of May's 54,179 Meetups:" with thumbnail images and descriptions for "Bergen County Scrapbooking Meetup" (with a quote about moving to the suburbs), "Slovenia New Technology Meetup" (with a quote about interesting things coming again), "Boston Social Justice Meetup" (with a quote about being the best volunteer-run networking group), and "New York Meditation Meetup" (with a quote about nice groups of people). To the right, there's a search form for "Enter an interest:" with fields for "e.g. Knitting, Moms, Dogs, Baseball" and dropdowns for "Country" (set to USA) and "ZIP Code". Below the search form, there are links to "all interests" and "all cities". On the right side, there are logos for TIME, Newsweek, and the Toronto Sun, each with a quote about the benefits of Meetup.com. The TIME quote says it's a convenient, non-threatening way to connect to other people who share similar interests and live nearby. The Newsweek quote says it helps groups of strangers organize monthly powwows at local watering holes. The Sun quote says it's important to stay true to your passions or interests, whatever they may be, and offers an opportunity to reconnect with yourself and others in a stress-free environment.

Current Research & Advanced Topics

Recommending Research
Papers

Recommending Research Papers

- Using Citation Webs
- For a full paper, we can recommend citations
 - A paper “rates” the papers it cites
 - Every paper has ratings in the system
- Other citation web mappings are possible, but many have problems

Demonstration #1

- Steps
 - Select user
 - Select paper
 - Select algorithm
 - See recommendations

What We Found

- Results published in McNee et al. (CSCW 2002):
 - Yes, we can make recommendations this way!
 - offline analysis showed that best algorithms could find half of recommendable withheld references in top 10, $\frac{3}{4}$ in top 40 recs
 - online experiments showed best algorithms gave recommendations more than half of which were relevant, and more than half of which were novel
 - Users like it!
 - more than half of users felt useful (1/4 to 1/3 said not)
 - 1-2 good recs out of 5 seemed sufficient for use
 - Different algorithms have different uses
- Further exploration in Torres et al. (JCDL 2004)

Phase II

- Shifted our focus to ACM Digital Library
- Greater exploration of user tasks:
 - awareness services
 - keeping track of a community
- More automation
 - find own bibliography from citations
 - find collaborators
- Thinking about “researcher’s desktop”

Moving Forward

- Collaboration
 - Computer Scientists (HCI, recommenders)
 - Librarians (field work, domain expertise, “real-life” service deployment)
- Research methods
 - Offline data gathering and feasibility studies
 - Online pilots and controlled experiments
 - Online field studies (including random-assignment studies)

What's Next?

- Short-Term Efforts
 - Task-specific recommendation
 - Understanding personal bibliographies
 - Privacy issues
- Longer-Term Efforts
 - Toolkits to support librarians and other power users
 - Exploring the shape of disciplines
 - Rights issues