

## Personalized Information Access

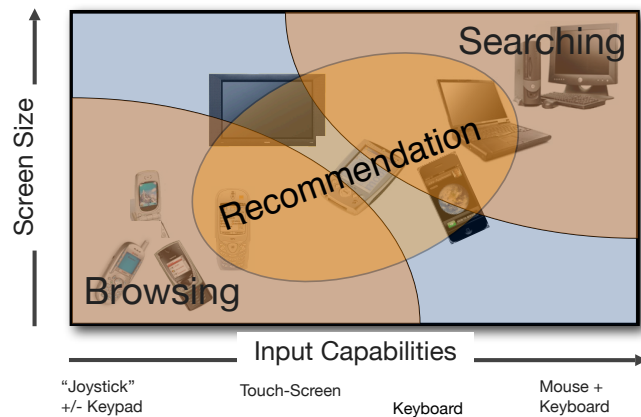
Barry Smyth  
University College Dublin, Ireland

## Key Ideas...

- Information Access Modalities
  - Navigation, Search, Recommendation
- Interactions & Feedback
  - Mining users interactions rather than item content.
- Profiling & Privacy
  - The Economics of Personalization

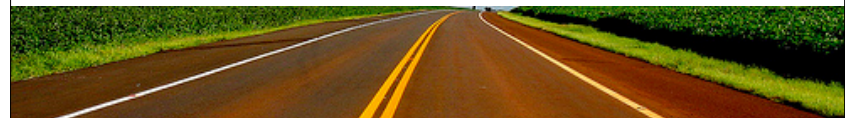


## Information Access Modalities



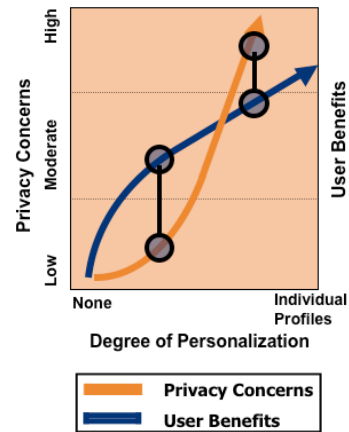
## Lessons learned along the way ...

- Understand user problems to identify application sweetspots.
- Exploit domain constraints. Keep it simple. Build for scale.
- Learn from live-user trials to understand real benefits.



## The Case for Personalization

- Information Overload
- Device Variability
  - PC, Mobile, TV, ...
- Interactions ⇒ Adaptation
- Profiling & Privacy
  - Individuals / sessions / communities
- The Economics of Personalization
  - Benefits & Costs



## The Road to Personalization

- Personalization | Browsing
  - Menu adaptation in mobile portals; Large-scale navigation prediction
- Personalization | Recommendation
  - Targeted feedback in conversational recommendation; Session adaptation & mining compound critiques
- Personalization | Search
  - Exploiting query repetition and selection regularity; Community-based profiling for anonymous personalization

user profiles

session profiles

community profiles



## 1. Personalization | Browsing

Adapting the Structure of Mobile Portals to the Navigation Patterns of Users

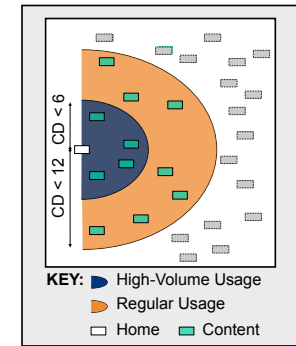
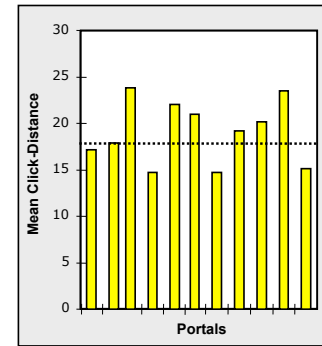
## Browsing on the Mobile Internet



## Browsing on the Mobile Internet

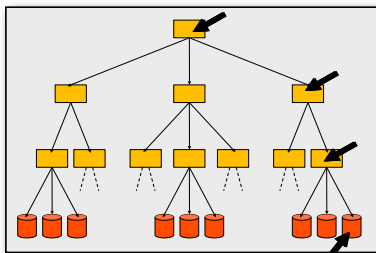


## The Click-Distance Problem



Mobile portals are compromised by large click-distances that render up to **75% of content all but invisible** to the most determined of users.

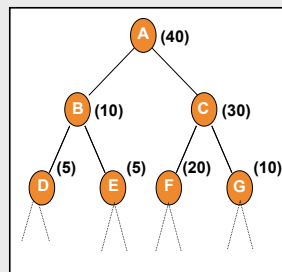
## Mining Navigation Trails



Recording user navigation trails provides a basis for profiling ...

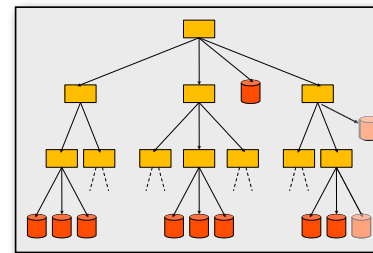
... which in turn can be used as the basis for a probabilistic navigation model.

### Navigation Profile



$$P(F|A) = P(C|A) \cdot P(F|C) \\ = (30/40) \cdot (20/30) \\ = 0.5$$

## Link Reordering & Promotion



### Menu Construction

Each personalized version of a requested menu,  $m$ , is constructed by adding the  $k$  most probable options, which are descendants of  $m$ .

By default all of  $m$ 's options are then ordered according to their access probabilities.



## Link Reordering & Promotion

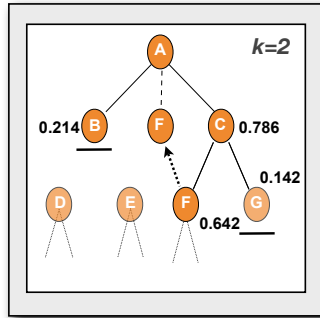
### Scalability Issues [Smyth & Cotter, AH 2002]

- Depth-limited, breadth-first search for real-time promotion and reordering.

$$P(o|m) \geq P(o'|m')$$

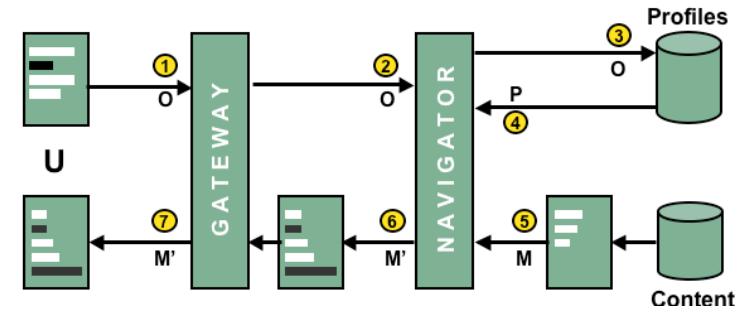
if  $o' \in m'$  and  $\text{Descendant}(m', m)$

Expand  $o'$  iff  $P(o'|m) > k^{\text{th}}$  best probability so far.

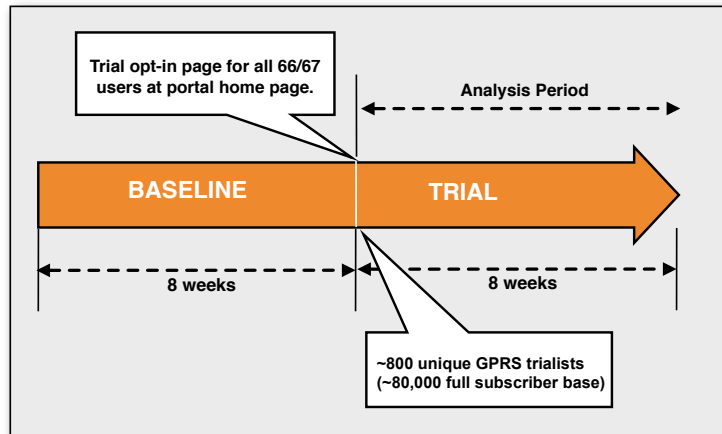


- Mobile OpCo requirements: approx. 1000 personalized pages per second with a latency of <500msec.

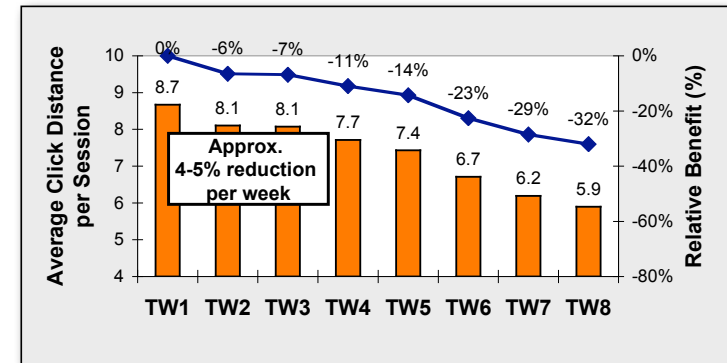
## Deployment Architecture



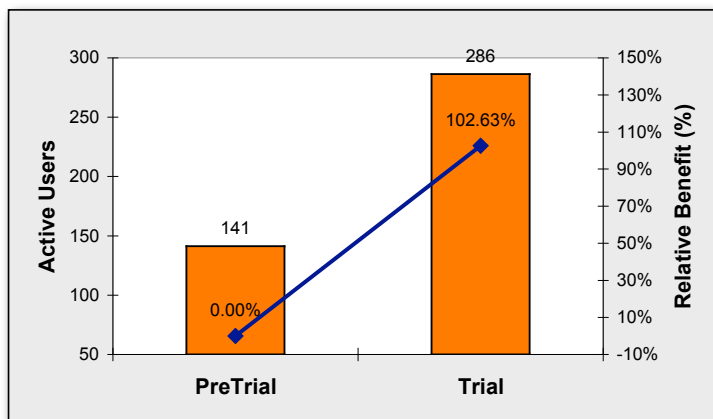
## Evaluation



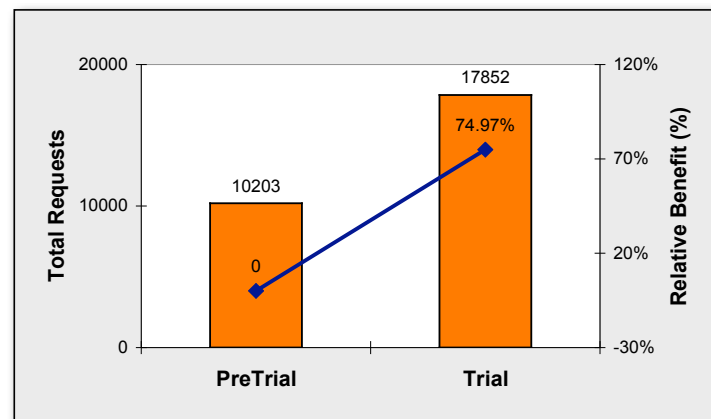
## Click-Distance



## Active Users



## Total Requests



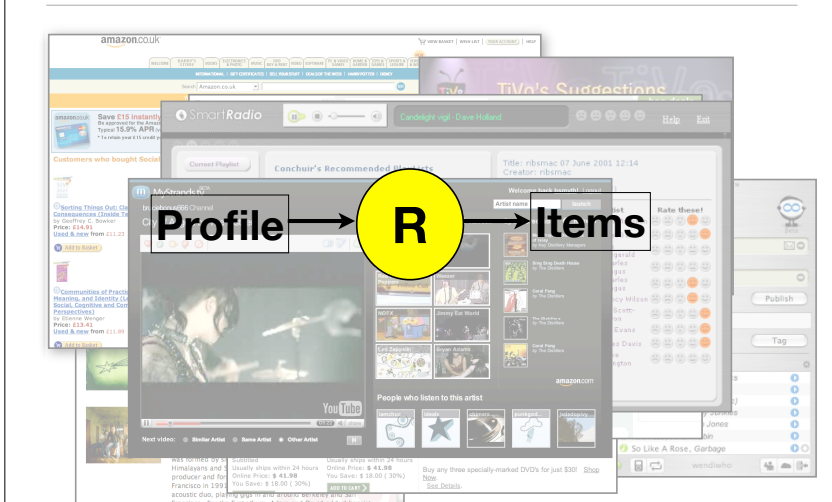
## Lessons Learned

- Selling AI ... *Return on Investment* is key!
  - Usability → Click-Distance → Usage
- The "AI Blackbox" vs the need for operator controls.
  - Integrated portal management and personalization administration
  - Personalization changes the way that operators manage their portals!
- User perceptions & the personalization-privacy trade-off
  - Satisfaction, usage, bandwidth, privacy, 90%+ opt-in rates.
- See also: [Smyth et al. IAAI 2007, AI Magazine (forthcoming)]

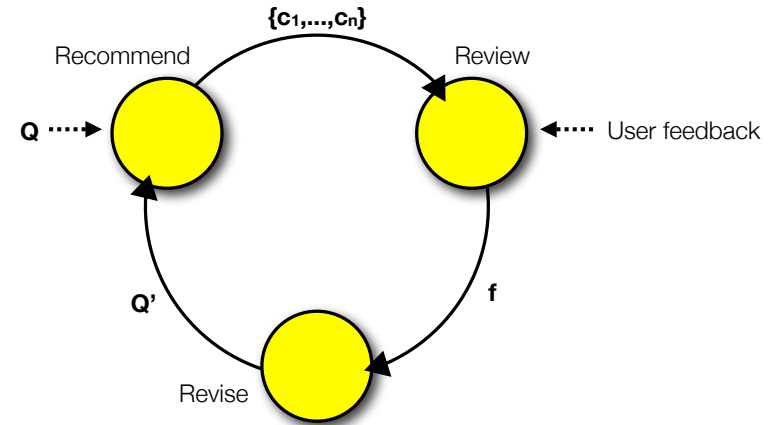
## 2. Personalization | Recommendation

Mining Feedback in Conversational Recommender Systems

## Recommender Systems (Single-Shot)



## Conversational Recommender Systems



## Critiquing-Based Recommenders

“Show me more like this but cheaper”  
(Unit critique over the *price* feature)

Adjust your preferences in real time and let us find the right product for you

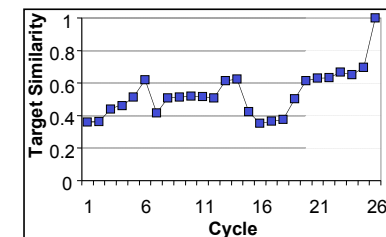
**Recommend(c,f):-**

Brand	Canon
Model	EOS-300D
Price (\$)	871.0
Resolution (M Pixels)	6.29
Optical Zoom (X)	10.0
Weight (grams)	645.0
Storage Type	Compact Flash

1. Filter remaining cases by critique, **f**.
2. Rank filtered cases by similarity to, **c**.
3. Return top **k** ( $k=1$ ) most similar cases.

## The Problem with (Unit) Critiques ...

- Session Analysis
  - Protracted sessions in simple product-spaces.
  - False-leads.
  - Changes of mind.
- Dynamic Critiquing
  - Unit  $\Rightarrow$  Compound Critiques
- Incremental Critiquing
  - How to adapt recommendations in response to new and changing preferences?



[Reilly, Zhang et al. EC 2007]

**Standard Unit Critiques**

**Compound Critiques**

**Recommendation**

## Dynamic Critiquing: Creating Compound Critiques

Not Satisfied with the result? you may select other recommendations listed below

- 1. Larger Hard-Disk, Lighter, Longer Battery Life and Cheaper.**  
But with Different Brand, Slower CPU, Smaller Screen and Different OS. [I like this](#)  
[>>see product detail<<](#)
- 2. Faster CPU and Larger Hard-Disk.**  
But with Different Brand, Different Type of CPU, Heavier, Different OS, Shorter Battery Life and More Expensive. [I like this](#)  
[>>see product detail<<](#)
- 3. Larger Hard-Disk and Cheaper.**  
But with Different Brand, Slower CPU, Smaller Screen, Heavier, Different OS and Shorter Battery Life. [I like this](#)  
[>>see product detail<<](#)

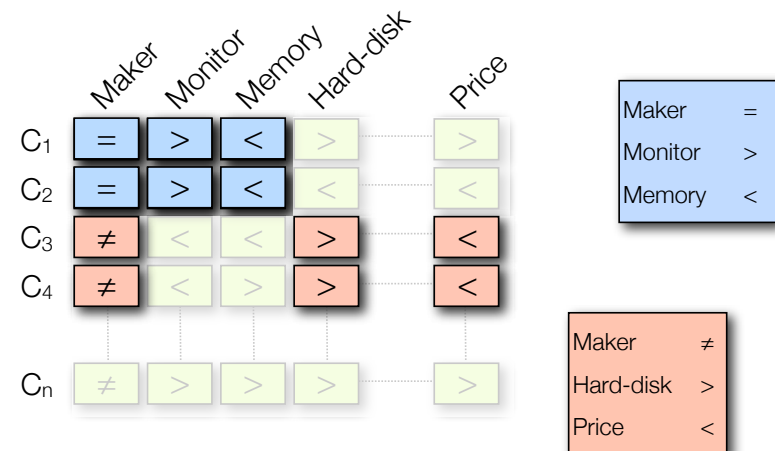
Generate compound critiques on-the-fly by using data mining techniques to extract common patterns of critiques from the remaining cases (relative to the current recommendation)

## Critique Patterns

	Recommendation	Alternative	Critique Pattern
Manufacturer	Compaq	Sony	!=
Monitor	14	12	<
Memory (MB)	512	256	<
Hard-Disk (GB)	40	30	<
Processor	Pentium 3	Pentium 3	=
Speed (Mhz)	1400	900	<
Type	Desktop	Laptop	!=
Price (€)	1500	3000	>

Each remaining case is converted in to a *critique pattern* to reflect the differences between it and the current recommendation....

## From Critique Patterns to Compound Critiques



## Ranking & Selecting Compound Critiques

- Use *A priori* to produce compound critiques (for the current cycle) containing 2-3 unit critiques
  - Large lists of candidates (50+/cycle in typical scenarios)
- Ranking based on support and confidence thresholds
  - Low-support critiques shown to provide best balance of filtering power and likely applicability ⇒ Best-k (k=3) compound critiques to present
- See also:
  1. Comparison of ranking strategies (Reilly et al, ECCBR 2004)
  2. MAUT critique mining (Zhang & Pu, AH 2006; Reilly, Zhang et al, IUI'07)

## Incremental Critiquing

The screenshot shows the QuirkShop.com website interface for finding a digital camera. The search criteria are set to Canon, 7x optical zoom, 512 MB memory, 780g weight, 6.2 M Pixels resolution, Large size, Magnesium case, and 995 price. The interface shows a list of matching cameras with 'EXPLAIN' and 'PICK' buttons. A red circle highlights the 'PICK' button for the first matching camera: '1. Less Memory and Lower Resolution and Cheaper'.

## Incremental Critiquing

- Incremental critiquing constructs an *in-session* user model, **U**, from the critiques provided by the user during each cycle.
- Sometimes critiques are inconsistent with previous critiques stored in **U**
  - *Contradictory critiques* - [Price > €750] vs. [Price < €500]
  - *Complementary critiques* - [Resolution > 5M] vs. [Resolution > 6.2M]
- The user model is maintained by a “*newer is better*” strategy in which more recent critiques over-ride past critiques.

## Incremental Critiquing

- Recommending a new case ...
  - Incremental critiquing considers the user’s critique history (**U**) during recommendation.
  - Each candidate case **c'** is ranked according to its **similarity to c** (the current case) *and* its **compatibility with U**.
  - Candidates are preferred if they are similar to **c** *and* if they satisfy many of the users past critiques.

$$\text{Compatibility}(c', U) = \frac{\sum v_i \text{ satisfies}(U_i, c')}{|U|}$$

$$\text{Quality}(c, c', U) = \text{Compatibility}(c', U) * \text{Similarity}(c, c')$$



# Evaluation

## E-Commerce Scenario

- Online digital camera store
- 200+ cameras

## Trial Setup

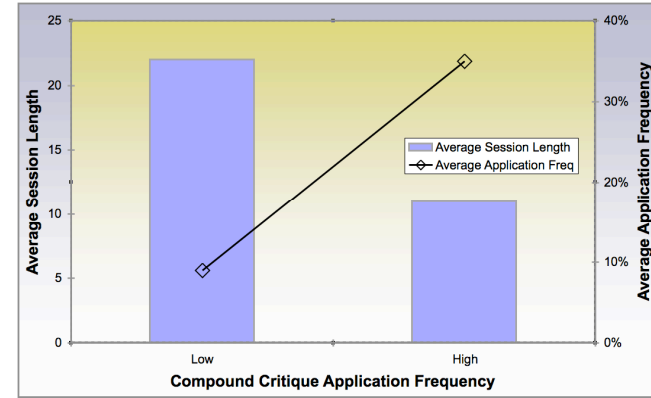
- 76 participants were asked to use the online store to 'shop' for a digital camera of their choice.
- Low vs high freq. use of compound critiques (0→100%, median ≈ 20%)

## Systems Compared

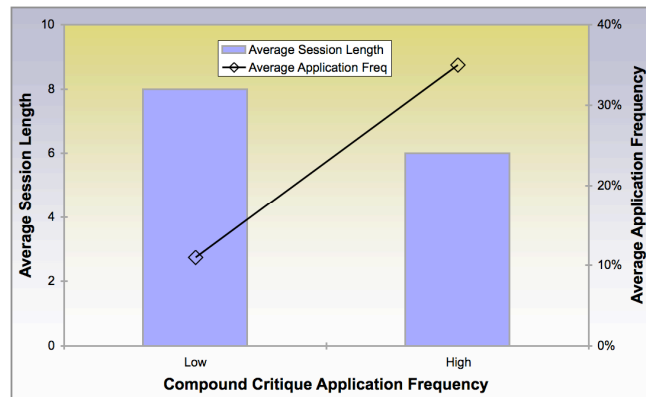
- Standard vs Incremental Critiquing x Unit vs Compound Critiques



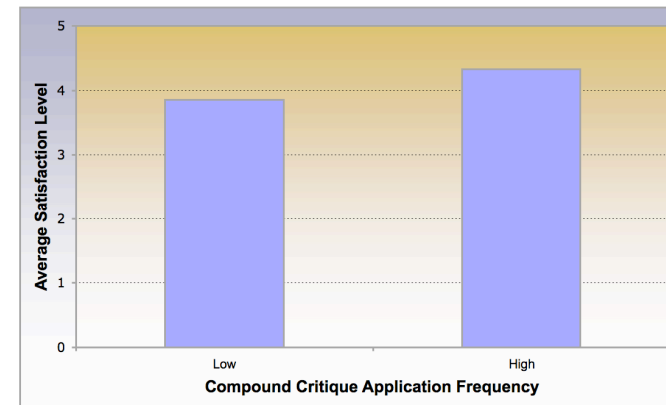
# Standard Critiquing



# Incremental Critiquing



# User Satisfaction



## Lessons Learned

---

- Critiquing is an effective form of feedback, but unit critiques can lead to protracted sessions (false-leads, dead-ends etc.)
- Dynamic, compound critiques offer a richer form of feedback.
- Incremental critiquing facilitates a more natural form of product exploration that adapts to the in-session behaviour of the current user.
- Significant benefits in session length without compromising user satisfaction.
- No persistent profiles ⇒ Limited privacy implications.

## 3. Personalization | Search

---

Harnessing Community Expertise in Collaborative Web Search

## Collaborative Web Search

---

- Observations on Web Search ...
  - Vague queries and the vocabulary gap.
  - Repetition & regularity in communities of like-minded searchers.
  - Query profiling is a privacy nightmare!
- Community-based personalization
  - Leverage existing search engine resources.
  - A Case-Based Reasoning Perspective.
    - Reuse of search cases ⇒ search expertise ⇒ community promotions
  - No need for individual search profiles.

## The Challenges of Web Search

---

- Vague Queries
  - Jaguar, Jaguar, Jaguar, ....
- Query Vocabulary ≠ Indexing Vocabulary (The Vocabulary Gap)
- Generic Search vs Community-Based Search
  - Communities are commonplace on the modern Web (ad hoc, formal, ...)
- Communities often search for similar things in similar ways
  - Query repetition and selection regularity is commonplace within communities or like-minded searchers.

## Vague Queries

Google search results for "michael jordan". The search bar shows "michael jordan" and the search button is labeled "Search". Below the search bar, there are tabs for "Web", "Video", "Images", and "News". The results show "Results 1 - 20".

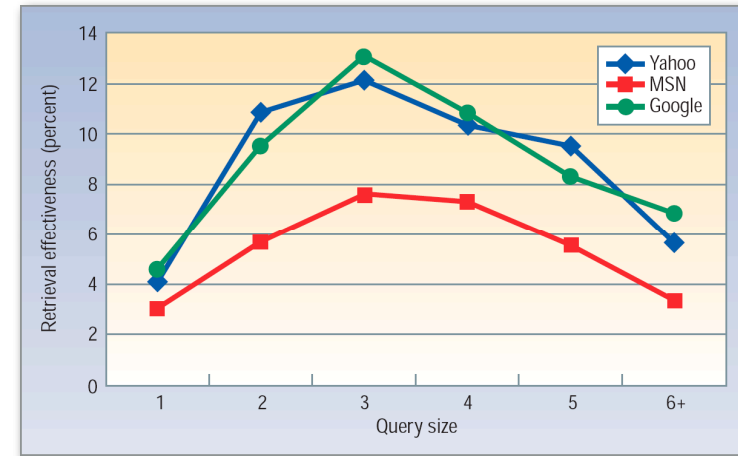
**Michael Jordan - Wikipedia, the free encyclopedia**  
 Michael Jeffrey Jordan (born February 17, 1963) is a retired American professional basketball player. Widely considered one of the greatest basketball ...  
[en.wikipedia.org/wiki/Michael\\_Jordan](http://en.wikipedia.org/wiki/Michael_Jordan) - 202k - [Cached](#) - [Similar pages](#) - [Note this](#)

**NBA.com: Michael Jordan Bio**  
 Michael Jordan | 23. Season statistics & Notes · Season splits · Game-by-game stats · Bio · Printable player file, 2002-03 Statistics. PPG, 20.0. RPG, 6.10 ...  
[www.nba.com/playerfile/michael\\_jordan.html](http://www.nba.com/playerfile/michael_jordan.html) - 139k - [Cached](#) - [Similar pages](#) - [Note this](#)

**NBA.com: Michael Jordan Summary**  
 Michael Jordan By acclamation, Michael Jordan is the greatest basketball player of all time. Although, a summary of his basketball career and influence on ...  
[www.nba.com/history/players/jordan\\_summary.html](http://www.nba.com/history/players/jordan_summary.html) - 47k - [Cached](#) - [Similar pages](#) - [Note this](#)  
 [ More results from www.nba.com ]

**23Jordan - A Michael Jordan Tribute**  
 A Michael Jordan tribute. Find the latest information, stats, memorable pictures, and a biography of Michael Jordan, the six-time NBA champion and league ...  
[www.23jordan.com/](http://www.23jordan.com/) - 24k - [Cached](#) - [Similar pages](#) - [Note this](#)

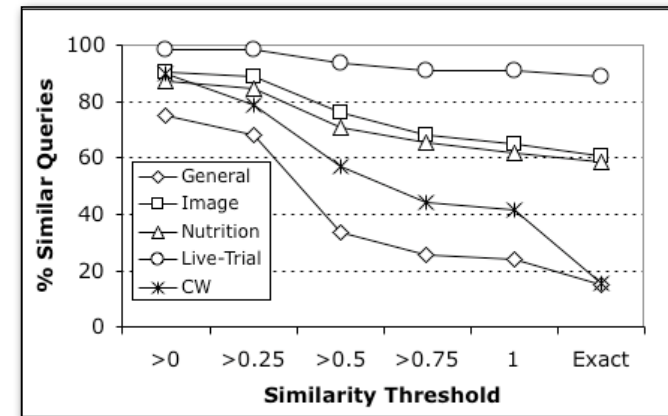
## The Vocabulary Gap



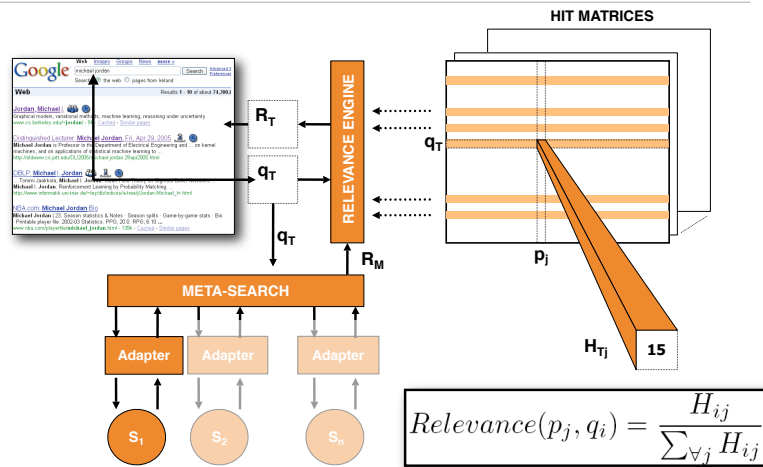
## Generic vs Community-Based Search

- Generic search engines (Google, Yahoo, Ask, ...) appear to dominate the world of Web search
  - ... but many search sessions reflect the niche interests of communities of like-minded individuals
  - ... and there are increasingly many examples of searchers searching in a community context
  - Some searchers are looking for information about a particular term-paper;
  - Researchers are searching for work related materials;
  - Visitors to a themed web site are using syndicated search services.
- 
- The screenshot shows the Autonet website interface. It features a search bar at the top with the text "autonet.ie - Ireland's foremost motorist website - new cars, used cars, dealers, motov, reviews and more". Below the search bar, there are navigation tabs: "AUTONET HOME", "USED CARS", "AUTONET DEALERS", "FINANCE", and "DEALERS ONLY". The main content area displays various car listings, including a Toyota Avensis. There are also sections for "USED CARS", "CAR IMPORTS", "MOTORBIKES", and "CAN'T FIND YOUR CAR". At the bottom, there is a "YOUR COMMENTS" section and a "USEFUL CAR STUFF" section.

## Repetition & Regularity in Web Search



# Collaborative Web Search Architecture



# A Case-Based Reasoning Perspective

- Hit-Matrix  $\approx$  Case-Base
  - Cases capture community search experiences as a query and a set of selections.
- Query Reuse  $\approx$  Case Retrieval
  - Select search cases for similar queries.

$$Sim(q_T, c_i) = \frac{|q_T \cap Spec(c_i)|}{|q_T \cup Spec(c_i)|}$$

	Spec(C <sub>i</sub> )		Sol(C <sub>i</sub> )			
jaguar pics	10	33				
jaguar			21	12		56
used jaguar cars			5		25	14
	car-photos.com	jaguarcars.com	Motorworld.com	usedcars.com	JustJags.com	

[Smyth et al., JIR 2006, UMUI 2004]

# A Case-Based Reasoning Perspective

- Promotion  $\approx$  Case Reuse
  - Most relevant selections are promoted within the result-list.
  - Result relevance based on query relevance and query similarity.

$$\frac{\sum_{i=1..n} Relevance(p_j, q_i) \cdot Sim(q_T, q_i)}{\sum_{i=1..n} Exists(p_j, q_i) \cdot Sim(q_T, q_i)}$$

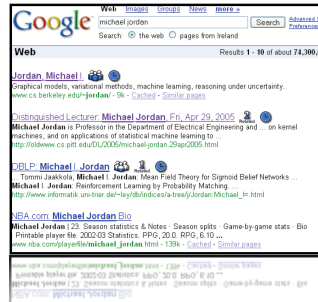
$$\begin{aligned}
 &Relevance("jaguar", jaguarcars.com) \\
 &= \frac{"jaguar" \quad "jaguar pics" \quad "used jaguar cars"}{(21/89) \cdot 1 + (33/43) \cdot 1/2 + (5/44) \cdot 1/3}{(1 + 1/2 + 1/3)} \\
 &= 0.36
 \end{aligned}$$

jaguar pics	10	33				
jaguar			21	12		56
used jaguar cars			5		25	14
	car-photos.com	jaguarcars.com	Motorworld.com	usedcars.com	JustJags.com	

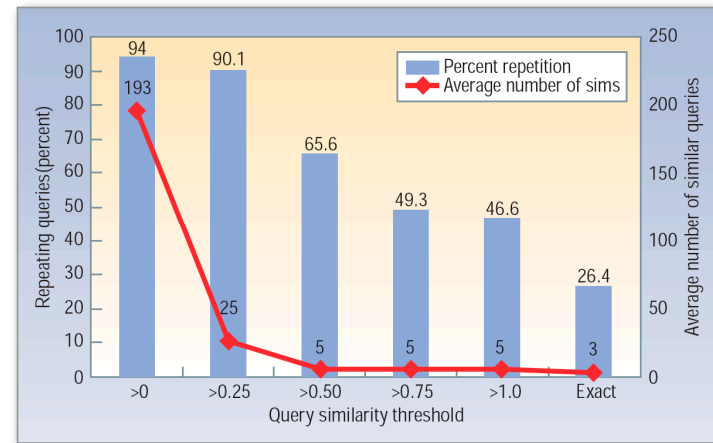
The screenshot shows a Google search for "michael jordan". The results are filtered to "Web" and show "Results 1 - 10 of about 74,300". The top result is for "Jordan, Michael I.", a Distinguished Lecturer at Michigan State University. Other results include a DBLP entry for Michael I. Jordan and an NBA.com profile for Michael Jordan.

## Evaluation

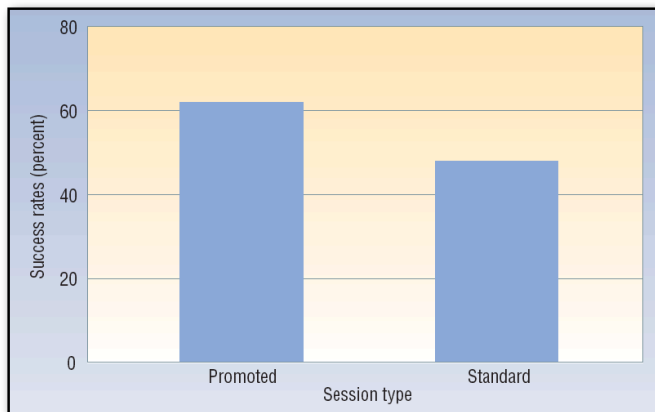
- Enterprise Search Scenario [Coyle & Smyth, ACM TOIT, IEEE Computer 2007]
  - Community composed of the employees of a local software company; search redirected through CWS-enhanced Google.
  - Approx. 70 employees over a 10 week period > 12,600 individual search sessions
- Methodology
  - Promoted vs. Standard Sessions
  - Successful vs. Failed Sessions
  - Promotion Sources (Self vs. Peer)



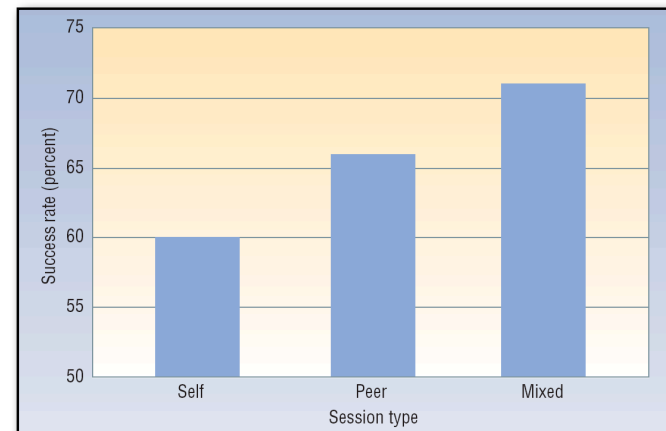
## Repetition & Regularity in Web Search

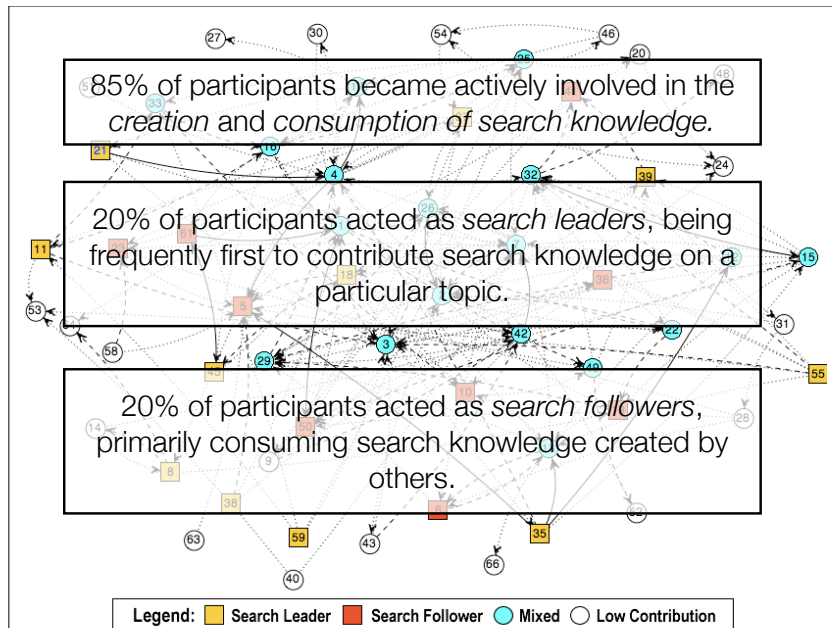


## Average Success Rates



## Sources of Promotions





## Lessons Learned

- Many Web search scenarios have a community context.
- Communities of like-minded searchers search for similarity information in similar ways.
- Community search histories can be harnessed as a source of search expertise and used to enhance standard search results.
- Aggregating community search experiences facilitates a form of anonymous, personalized Web search.

## Further Work...

- Beyond click-thru mining [Boydell & Smyth, IUI 2007, SIGIR 2006]
  - Snippet-based Collaborative Web Search
  - Generating community-focused result summaries
- Click-spam and trust-based filtering [Briggs & Smyth, AIR, IUI, ECIR 2007]
  - Malicious users ⇒ forced promotion
- Towards Inter-Community Web Search [Freyne & Smyth, AH, ECCBR 2006]
  - Sharing result promotions between related search communities.
- Community expertise → folksonomy learning

Communities About I-Spy

the smarter way to search

6 nations Search PRIVATE SEARCH

rugby union: Your Search for 6 nations returned 34 Results | Displaying 1 - 10 | Result Page: 1 2 3 4 Next

Related Information I-Spy Recommends

Recent Queries

- 6 nations championship
- geordan murphy pictures
- geordan murphy
- haka
- lions tour

Recent Web Pages

- Six Nations Rugby 2005
- Sporting Heroes - Photo...
- Leicester Tigers - Ceor...
- New Zealand Haka Haka...
- British and Irish Lions...

Popular Queries

- irish lions
- william weath ellis
- haka
- conrad smith
- 1981 springbok tour

Popular Web Pages

- Lions Tour Home

Rugby Union European Rugby (4) Irish Rugby (3) English Rugby (1)

The Rugby Union community recommends the following results:

**Six Nations Rugby 2005**

The Six Nations 2005 - Europe's Premier International Rugby Tournament ... superior to the World Champions in a close-fought struggle for the 6 Nations Championship in Paris ... The title "6 Nations Rugby" is a description of content and does ...

Related Queries 6 nations championship

**Six Nations Rugby Tournament**

6nations.co.uk brings you all the excitement of the six nations rugby tournament including online ticket sales ... Corporate Hospitality. 6 nations Tickets - Buy Now!! Venues. 6 nations Tickets. Basket. Conditions. Payment ...

Other Matching Results

**Official RBS 6 Nations 2005**

Official RBS 6 Nations 2005 ...

http://www.6nations.net

Communities About I-Spy

the smarter way to search

6 nations Search PRIVATE SEARCH

rugby union: Your Search for 6 nations returned 34 Results | Displaying 1 - 10 | Result Page: 1 2 3 4 Next

Related Information I.Spy Recommends

Rugby Union European Rugby (4) Irish Rugby (3) English Rugby (1)

**Recent Queries** (VIEW ALL)

- 6 nations championship
- geordan murphy pictures
- geordan murphy
- haka
- lions tour

**Recent Web Pages** (VIEW ALL)

- Six Nations Rugby 2005
- Sporting Heroes - Photo...
- Leicester Tigers - Geor...
- New Zealand Haka HakaL...
- British and Irish Lions...

**Popular Queries** (VIEW ALL)

- irish lions
- william webb ellis
- haka
- outward spiral
- 1981 springbok tour

**Popular Web Pages** (VIEW ALL)

- Lions Tour Home
- British and Irish Lions...
- Rugby Football Union ...

**The Irish Rugby community is 28% similar to the Rugby Union community and has the following to recommend**

**Rugby.ie: Home**

Ireland's premier Rugby news website providing up to the minute reports. ... returns to the Ireland squad for this year's 6 Nations following his injury-enforced absence from the autumn ... scrum-half for the RBS 6 Nations. Scotland hit by Macfadyen ...  
http://www.rugby.ie

**Official RBS 6 Nations 2005**

Official RBS 6 Nations 2005 ...  
http://www.6nations.net

**Ireland.com / Today / Sport / Rugby / Six Nations Tournament**

Full coverage of the 6 Nations Rugby Tournament from The Irish Times/ireland.com ... the worth of the Triple Crown need only think for ... pivotal playmakers in the Triple Crown and Championship campaigns of ... eased into a 15-6 lead by the break with ...  
http://www.ireland.com/sports/rugby/6nations/2004/news/0327/news3.htm

**Related Queries** rugby triple crown 6 nations triple crown

Web Images Groups News more »

Google cbr Search Advanced Search Preferences

Search:  the web  pages from Ireland

Web Results 1 - 10 of at

**Cbr** Sponsored Link

www.ebay.ie 1000s of new & used vehicles added every day. Find it here!

**Case-Based Reasoning [the AI-CBR Homepage]**

Featured papers, researchers and projects, links to CBR people, research projects, along with a comprehensive guide to software tools, ...  
www.ai-cbr.org/ - 24k - Cached - Similar pages

**Case-Based Reasoning [Commercial Organisations using CBR]**

This pages lists commercial organisations who are using Case-Based Reasoning provide further information. AI-CBR.  
http://www.ai-cbr.org/applied.html

**Wikipedia: Case-based reasoning**

Wikipedia Free Encyclopedia's article on 'Case-based reasoning'  
http://en.wikipedia.org/wiki/Case-based\_reasoning

**Comic Book Resources - Daily Comic Book News, Reviews, Previews ...**

Comic Book Resources - Daily Comic Book News, Reviews, Previews, Commentary and Message Boards.  
www.comicbookresources.com/ - 54k - 27 Aug 2006 - Cached - Similar pages

Result was selected by you and 9 others:	
<b>Recommender:</b>	<b>Trust:</b>
+ kkoyrn	94.4%
+ evanp	76.9%
- mabes25	72.6%
<b>Queries:</b>	
research cbr	
cbr publications	
<b>More recommenders...</b>	

## Conclusions

- Personalized Information Access
  - Navigation, Recommendation, Search
  - Profiling strategies (sessions, individuals, communities)
- Privacy and the Economics of Personalization
  - Trading privacy for features ...
  - Mobile subscriber take-up; Recent FaceBook/MySpace studies
- Next Steps ...
  - Social information access - integrating navigation, search and recommendation [Farzan, Coyle, et al, IUI, Hypertext 2007]
  - ...

## With special thanks to...

- University College Dublin
  - Evelyn Balfe, Oisín Boydell, Peter Briggs, Maurice Coyle, Jill Freyne, Michael O'Mahony, Karen Church, James Reilly, Kevin McCarthy, Lorraine McGinty
- Collaborators
  - ChangingWorlds Ltd.
  - Peter Brusilovsky, Rosta Farzan (University of Pittsburgh)
  - Pearl Pu, Jiyong Zhang, Ecole Polytechnique Fédérale de Lausanne (EPFL)
  - Maria Salamo, Universitat de Barcelona

