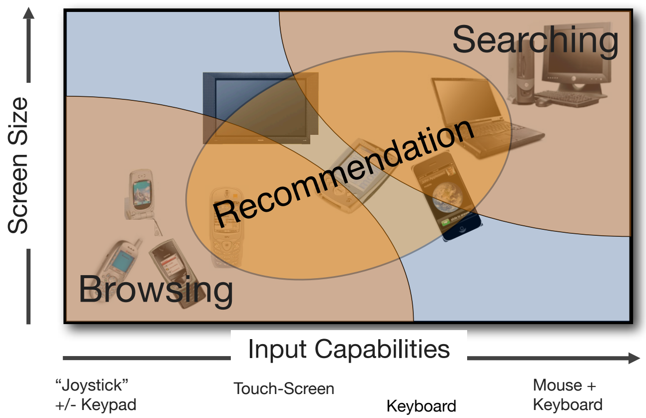
Personalized Information Access

Barry Smyth
University College Dublin, Ireland

Key Ideas...

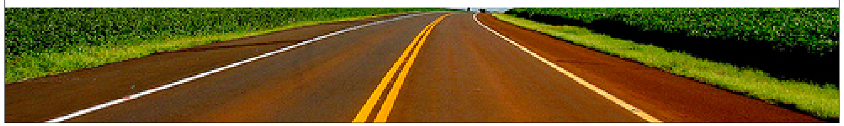
- Information Access Modalities
  - Navigation, Search, Recommendation
- Interactions & Feedback
  - Mining users interactions rather than item content.
- Profiling & Privacy
  - The Economics of Personalization

Information Access Modalities



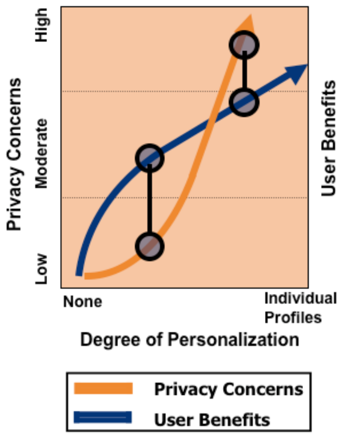
Lessons learned along the way ...

- 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



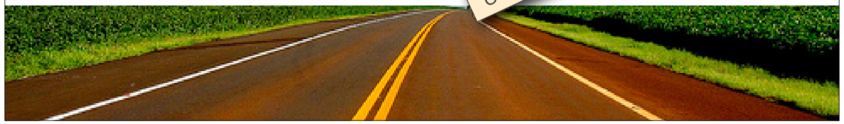
The Road to Personalization

- Personalization | Browsing
  - Menu adaptation in mobile portals; Large-scale, p... navigation prediction
- Personalization | Recommendation
  - Targeted feedback in conversational recomm... -session adaptation & mining compound critiques
- Personalization | Search
  - Exploiting query repetition and selection regularity i... arch; Community-based profiling for anonymous personalizati...

user profiles

session profiles

community profiles



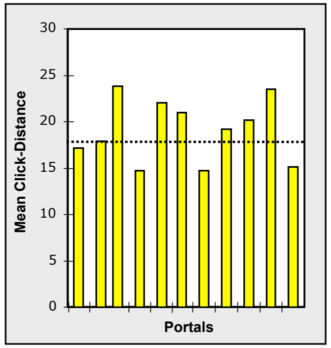
1. Personalization | Browsing

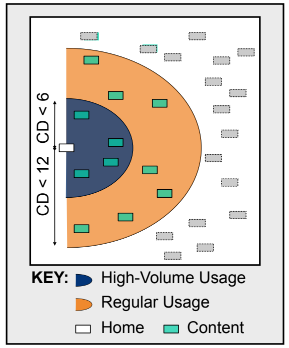
Adapting the Structure of Mobile Portals to the Navigation Patterns of Users

Browsing on the Mobile Internet

Browsing on the Mobile Internet

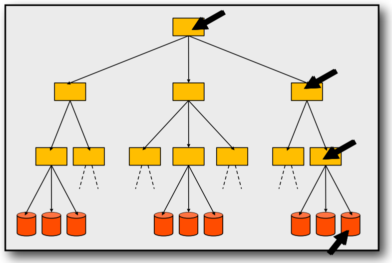
The Click-Distance Problem





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

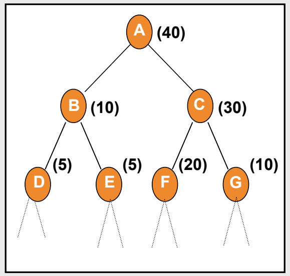
Mining Navigation Trails



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

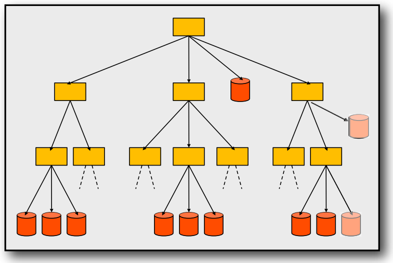
... which in turn can be used as the basis for a probabilistic navigation model.

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***.

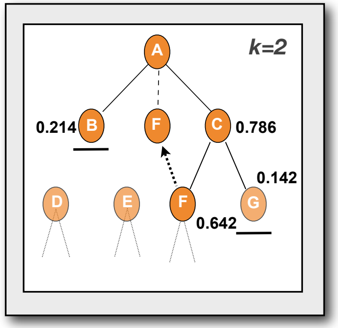
By default all of ***m*** 's options are then ordered according to their access probabilities.

Link Reordering & Promotion

- Scalability Issues [Smyth & Cotter, AH 2002]
  - Depth-limited, breadth-first search for real-time promotion and reordering.

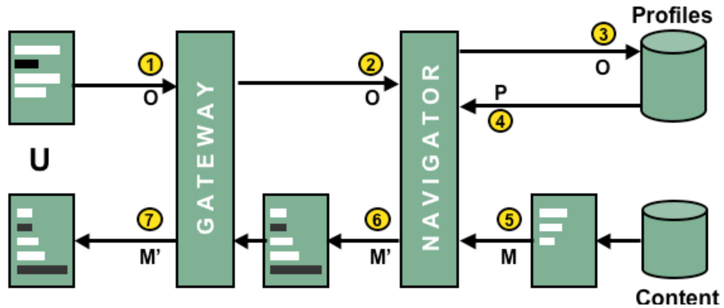
P(o|m) ≥ P(o'|m')
if o'є m' and Descendant(m',m)

Expand o' iff P(o'|m) > kth best probability so far.

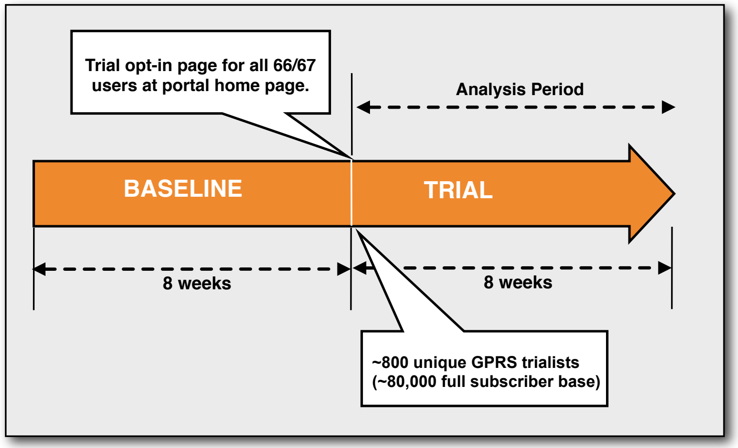


- Mobile OpCo requirements: approx. 1000 personalized pages per second with a latency of <500msec.

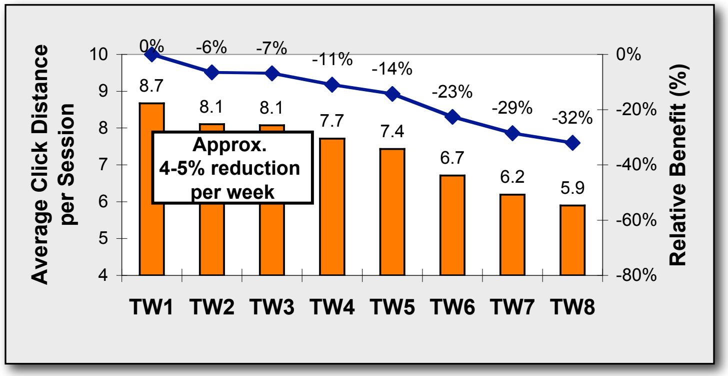
Deployment Architecture



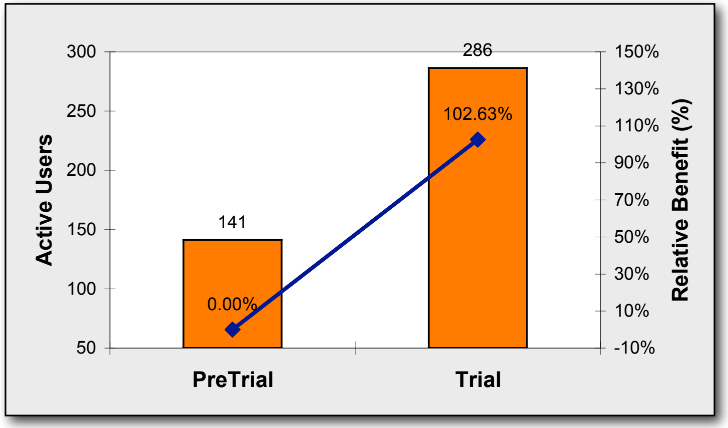
Evaluation



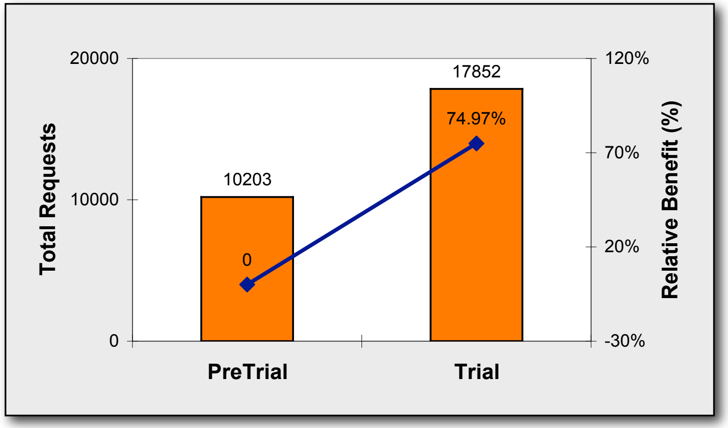
Click-Distance



Active Users



Total Requests



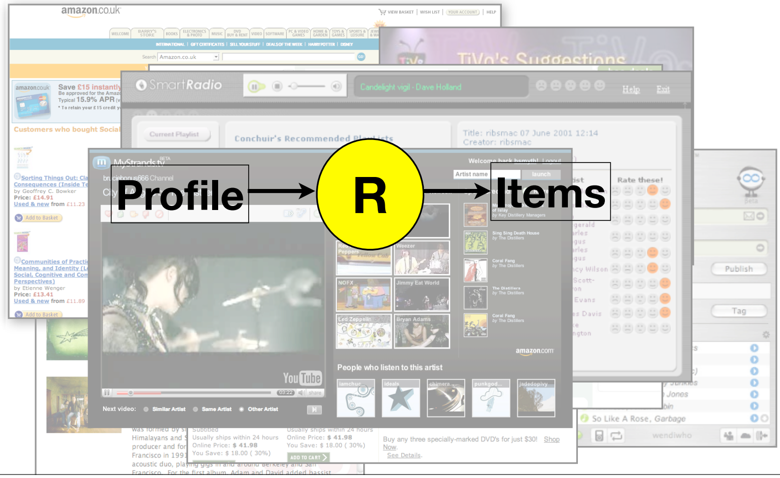
Lessons Learned

- 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)]

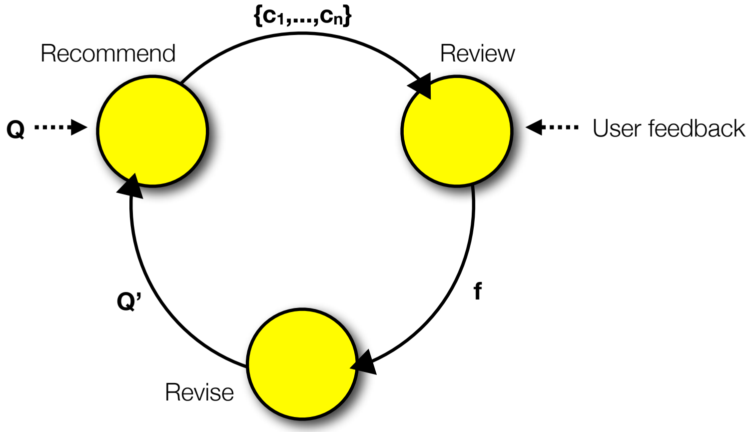
2. Personalization | Recommendation

Mining Feedback in Conversational Recommender Systems

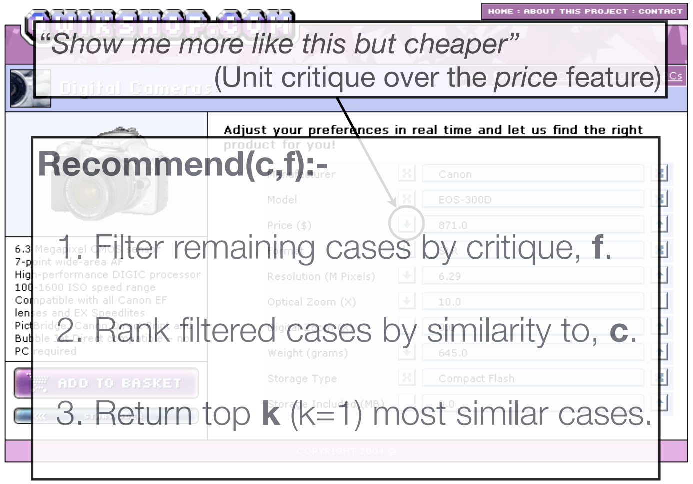
Recommender Systems (Single-Shot)



Conversational Recommender Systems

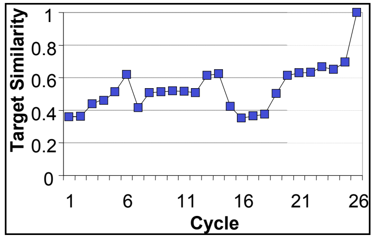


Critiquing-Based Recommenders

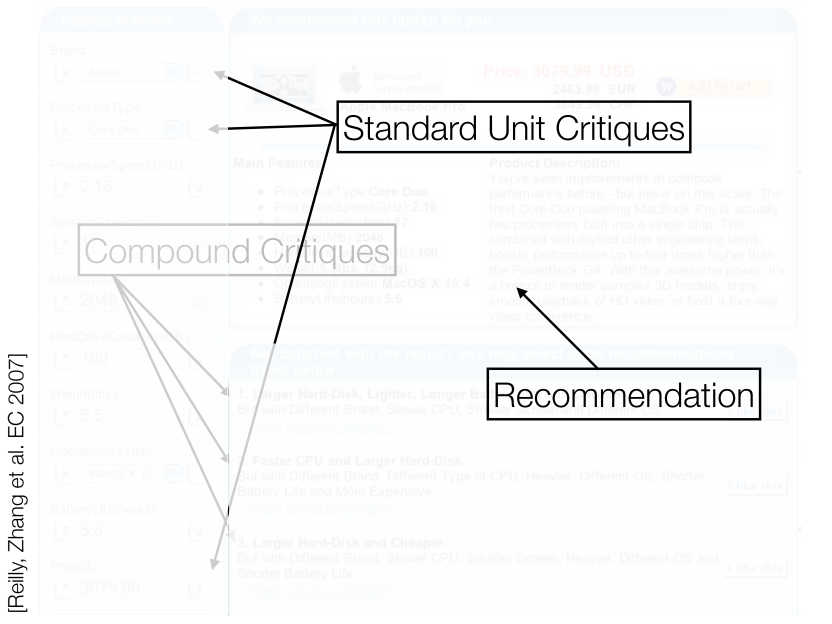


The Problem with (Unit) Critiques ...

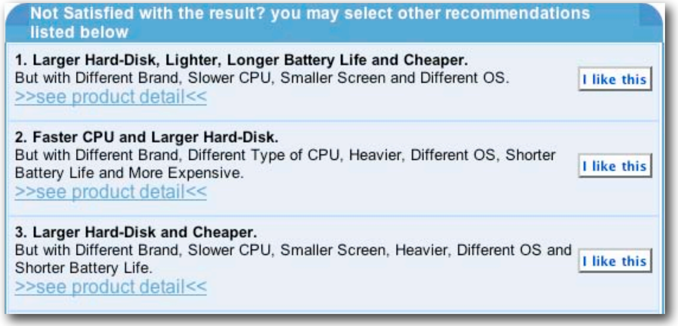
- Session Analysis
  - Protracted sessions in simple product-spaces.
  - False-leads.
  - Changes of mind.



- Dynamic Critiquing
  - Unit ⇒ Compound Critiques
- Incremental Critiquing
  - How to adapt recommendations in response to new and changing preferences?



Dynamic Critiquing: Creating Compound Critiques



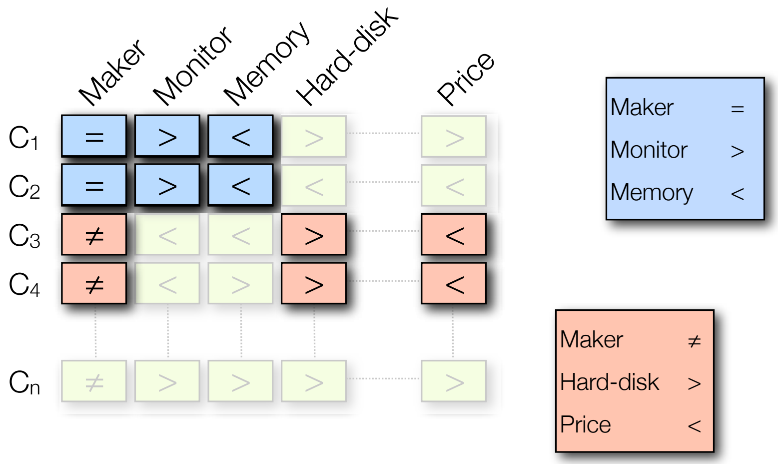
Generate compound critiques on-the-fly by using data mining techniques to extract common patterns of critiques from the remaining cases (*relative to the current recommendation*)

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 *Apriori* 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

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.

$$\text{Compatibility}(c',U) = \frac{\sum_{\forall i} \text{satisfies}(U_i,c')}{|U|}$$

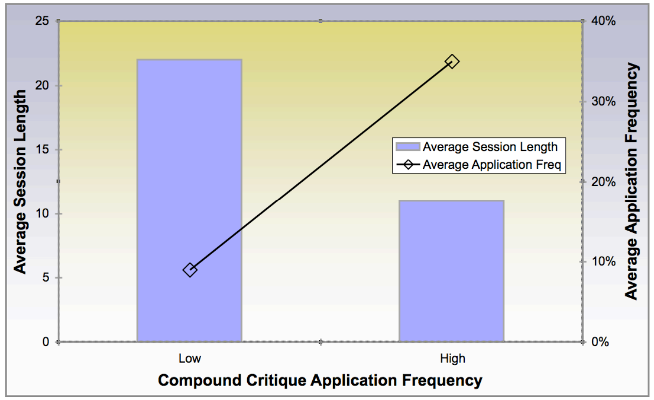
$$\text{Quality}(c,c',U) = \text{Compatibility}(c',U) * \text{Similarity}(c,c')$$

Evaluation

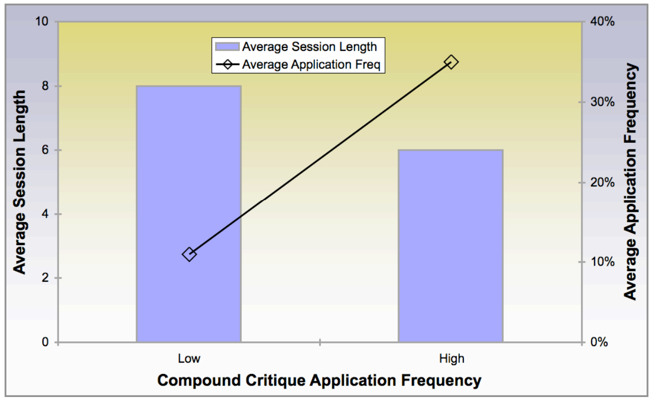
- E-Commerce Scenario
  - Online digital camera store
  - 200+ cameras

- 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

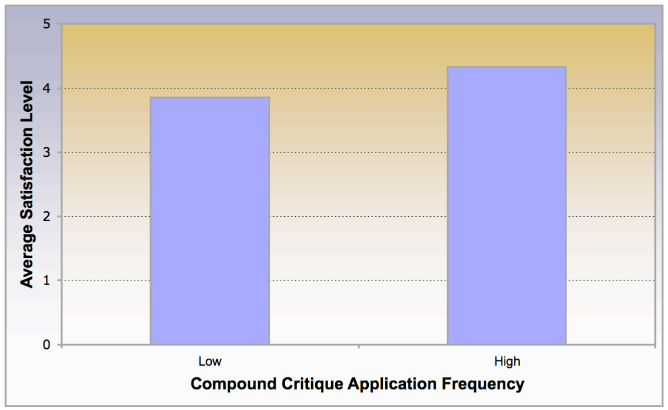
Standard Critiquing



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

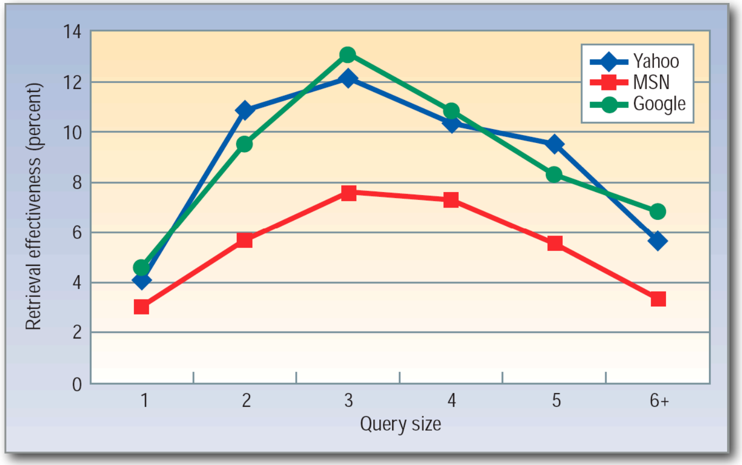
- Observations on Web Search ...
  - Vague queries and the vocabulary gap.
  - Repetition & regularity in communities of like-minded searchers.
  - Query profiling is a privacy nightmare!
- Community-based personalization
  - Leverage existing search engine resources.
  - A Case-Based Reasoning Perspective.
    - 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.

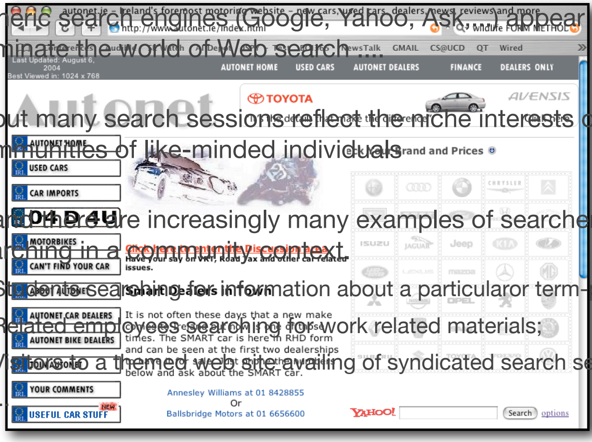
Vague Queries

The Vocabulary Gap

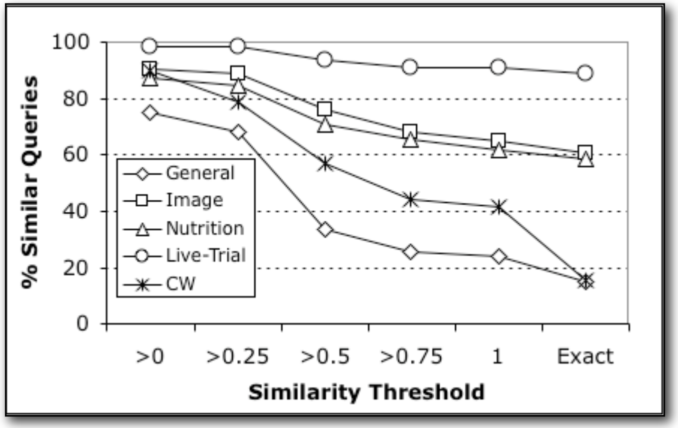


Generic vs Community-Based Search

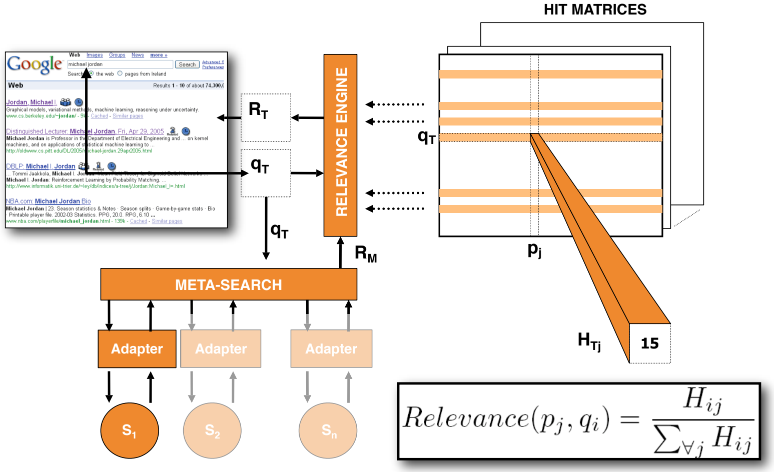
- Generic search engines (Google, Yahoo, Ask, ...) appear to dominate the world of Web search ....
- ... but many search sessions reflect the niche interests of communities of like-minded individuals
- ... and there are increasingly many examples of searchers searching in a community context
  - Students searching for information about a particularor term-paper;
  - Related employees searching for work related materials;
  - Visitors to a themed web site availing of syndicated search services.
  - ...



Repetition & Regularity in Web Search



Collaborative Web Search Architecture

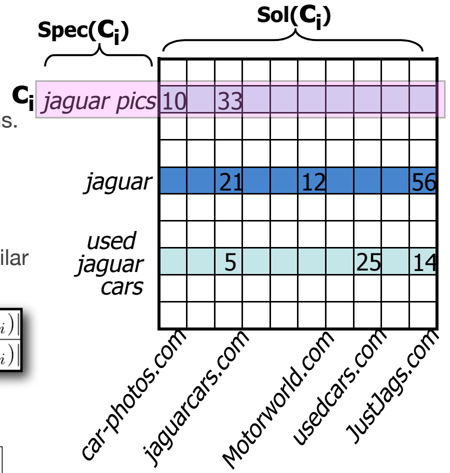


$$Relevance(p_j, q_i) = \frac{H_{ij}}{\sum_{\forall j} H_{ij}}$$

A Case-Based Reasoning Perspective

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

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



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

A Case-Based Reasoning Perspective

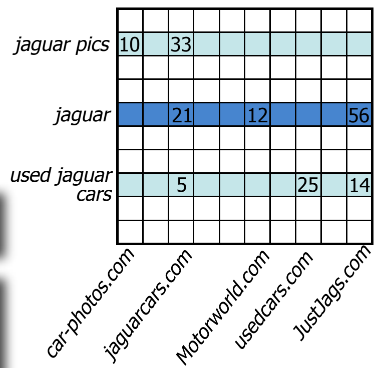
- Promotion ≈ Case Reuse
  - Most relevant selections are promoted within the result-list.
  - Result relevance based on query relevance and query similarity.

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

Relevance("jaguar",jaguarcars.com)

$$= \frac{\underbrace{(21/89)*1}_{\text{"jaguar"}}+\underbrace{(33/43)*1/2}_{\text{"jaguar pics"}}+\underbrace{(5/44)*1/3}_{\text{"used jaguar cars"}}}{(1 + 1/2 + 1/3)}$$

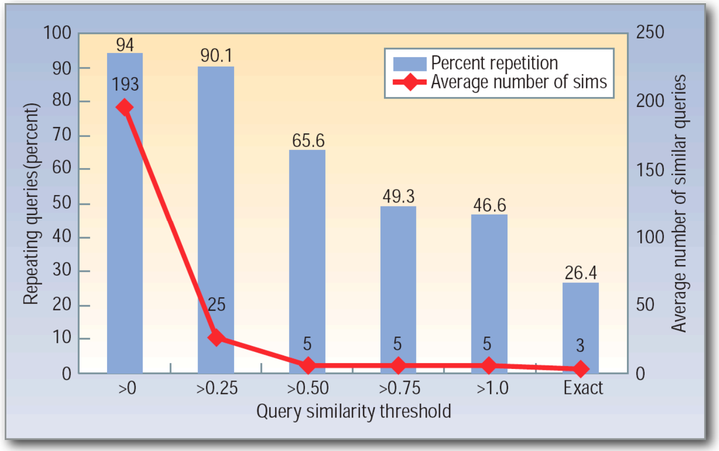
= 0.36



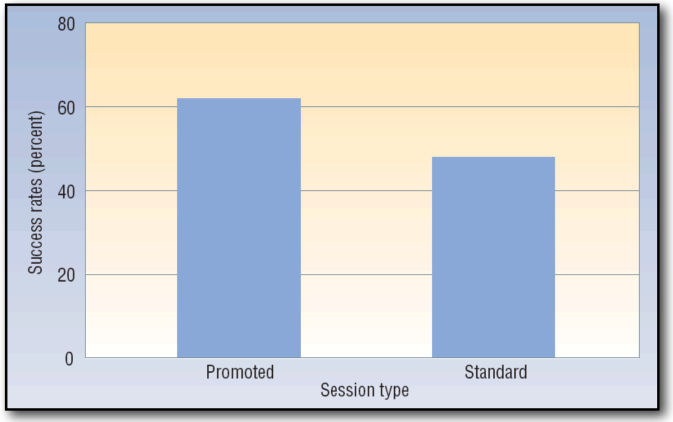
Evaluation

- 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)

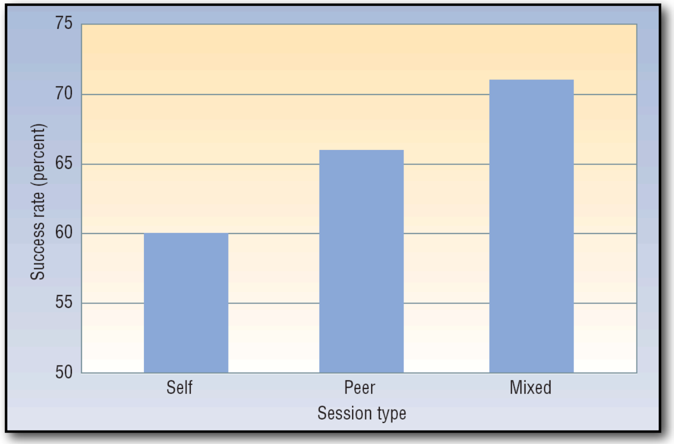
Repetition & Regularity in Web Search

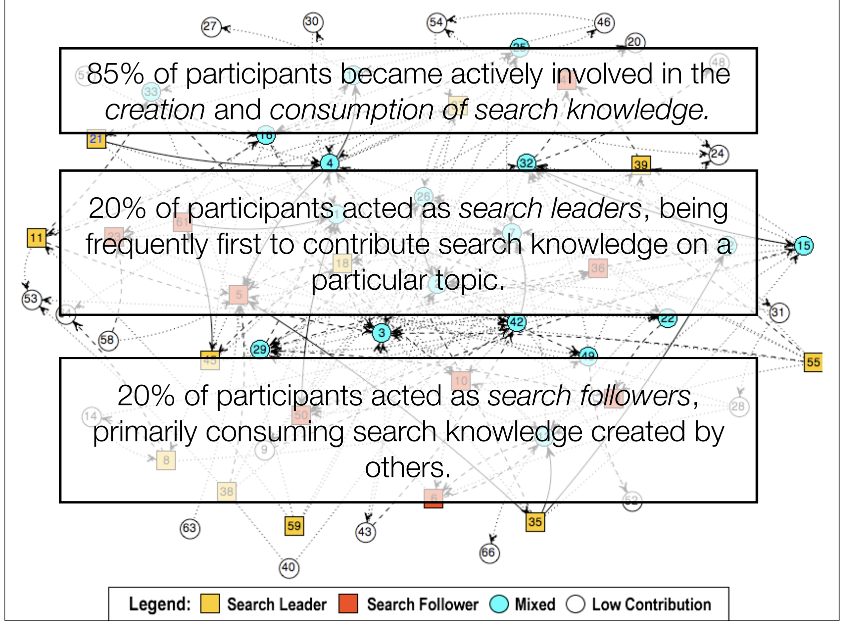


Average Success Rates



Sources of Promotions



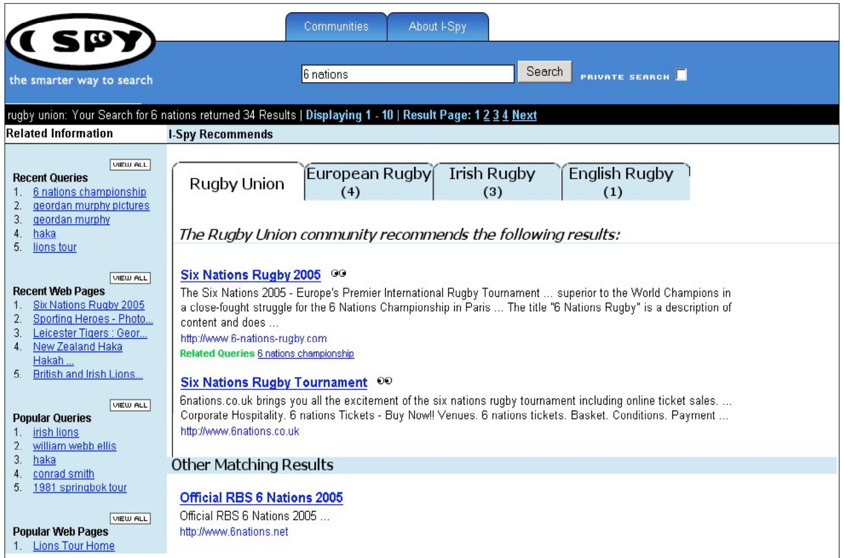


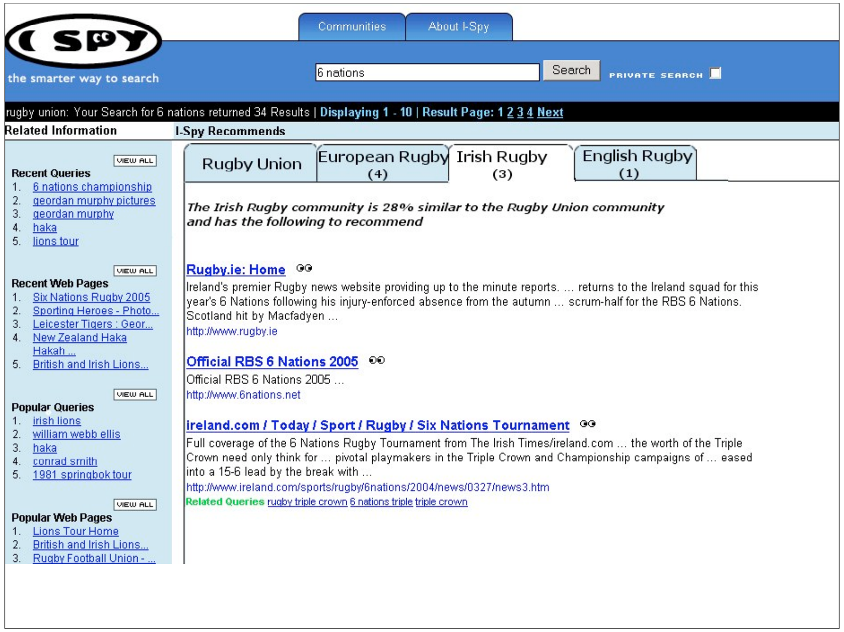
Lessons Learned

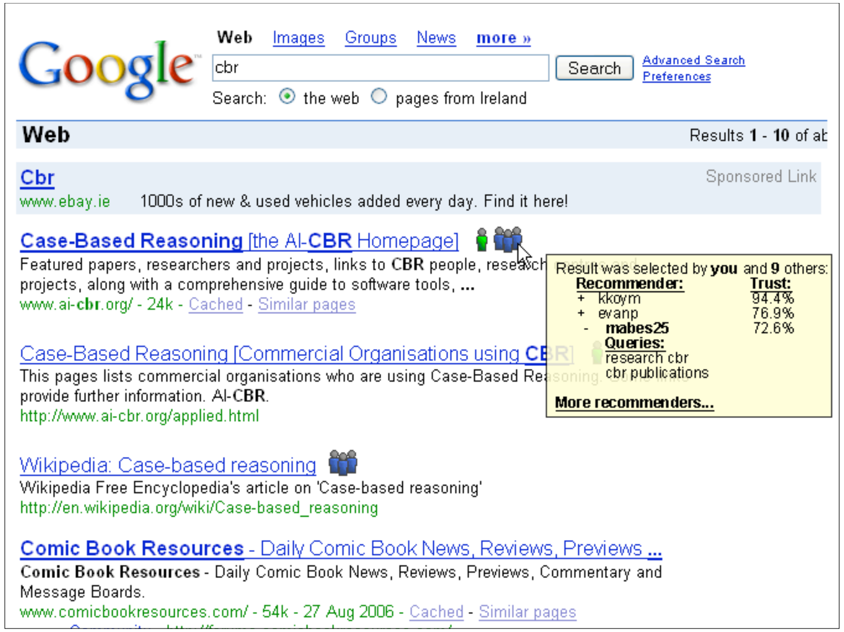
- Many Web search scenarios have a community context.
- Communities of like-minded searchers search for similarity information in similar ways.
- Community search histories can be harnessed as a source of *search expertise* and used to enhance standard search results.
- Aggregating community search experiences 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 experrise → folksonomy learning







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, Oisin 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