Beyond the Commons

Investigating the Value of Personalizing Web Search

Jaime Teevan[º], Susan T. Dumais[º] and Eric Horvitz[º]

[º]MIT, [º]Microsoft Research

Google Search: pia workshop - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://www.google.com

Google

Web

Privacy Training: PRIVA-C

PIA Workshop Program - 1 D

Privacy Impact Assessments.

www.priva-c.com/corporate/trai

Ideas for the **PIA Worksh**

Ideas for the **PIA Workshop**.

Workshop (20.0Kb 07sec @ 2

www.nice.org.uk/page.aspx?o

Daily Times - Site Edition

Two-day **PIA workshop** ends.

National Assembly Question H

www.dailytimes.com.pk/default

1 - Number One News Re

PIA workshop on customer s

To be inaugurated by Director

www.jang.com.pk/thenews/ma

[PDF] Direzione KAS-**PIA**.

File Format: PDF/Adobe Acrob

Regional Stakeholder **Worksh**

Initiative by the incoming EU C

www.minambiente.it/Sito/setto

[PDF] Annex ME 2 **PIA Wor**

File Format: PDF/Adobe Acrob

Participatory impact assessme

aim was to develop impact ass

www.naro.go.ug/Information/C

Similar pages

MSN Search: pia workshop - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://search.msn.com/results.aspx?q=pia+workshop&FORM=QBRE Go

pia workshop

MSN Search: pia work...

Web Desktop News Images

pia workshop

+Search Builder Help

Web Results

Page 1 of 88,316 results containing p

GAIN -- Graphic Arts Inform

... from the Sheetfed Offset Press

by WhatTheyThink.com and GAIN

www.gain.net Cached page

PIA/GATF - Conference

... you're accepting a **PIA**/GAT

workshop to assist with the u

www.gain.net/conference/brochu

Show more results from "ww

Privacy Training: PRIVA-C F

... Treasury Board Secretariat **PIA**

- 1 Day (Tentative Agenda): Sessi

www.priva-c.com/corporate/training_p

Privacy Training: PRIVA

... TM | PrivacyHORIZON |

Workshop Training for Train

www.priva-c.com/corporate/traini

Show more results from "ww

PiA - Primary Immunodeficie

... The Annual **PiA** Mark Steanson'

great support in 2004 To learn mo

www.pia.org.uk/medical_students_wo

PiA - Primary Immunode

... AGM News Applications for

Contribute to the **PiA** Website

www.pia.org.uk Cached page

Show more results from "ww

MSN Search: pia workshop - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://search.msn.com/results.aspx?q=pia+workshop&FORM=QBRE Go

pia workshop Search Web Form Fill Blocked (1) Spaces

Opening...

Web Desktop News Images Local (BETA) Encarta

pia workshop Search Near Me

msn Search

+Search Builder Settings Help Español

Personalized Web Results

1-100 containing **pia workshop**

My Search: Personalized WebCache Desktop Web

PIA 2005 - Workshop on New Technologies for Personalized Information ...

PIA 2005 – Workshop on New Technologies for Personalized Information Access July 24th and 25th, 2005, Edinburgh Scotland, UK Information access is one of the hottest topics of information society ...

http://irgroup.cs.uni-magdeburg.de/pia2005/

Privacy Training: PRIVA-C Workshop Registration

... TM | PrivacyHORIZON | PrivacySOURCE Quick Search: PRIVA- C TM Training CPO Workshop PIA Workshop Training for Trainers Locations Registration Consulting Technology Site Map | Contact Us ...

http://www.priva-c.com/corporate/training_register.asp

Intoinfo Consulting Group - A Step-by-Step Guide to Privacy Impact ...

... and outcomes of the PIA Understanding the costs and risks related to not doing a PIA After attending this workshop, participants have a better understanding of what type of activities could have an ...

http://www.intoinfo.com/content/misc/privacy.asp

Potential Industrial Applications of Mixed and Augmented Reality

The International Workshop on Potential Industrial Applications of Mixed and Augmented Reality October 7th, 2003. The National Center of Sciences, Tokyo, Japan • Home • Call ...

http://www.mr-system.com/pia/

Intoinfo Consulting Group - A Step-by-Step Guide to Privacy Impact ...

... and outcomes of the PIA Understanding the costs and risks related to not doing a PIA After attending this workshop, participants have a better understanding of what type of activities could have an ...

http://www.intoinfo.com/content/misc/privacy.asp?print=true

Architecture for Education

... Environments Thinking the Way They Think Academy of Neuroscience for Architecture K-6 Classroom Workshop "Design Down" Process Practice Theory, Pedagogy, and the Design of Learning Environments ...

http://www.aia.org/PIA/cae/

Done Internet

Outline

- Value of personalizing web search
- Study of personal relevancy
  - Rank and rating
  - Same query, different ratings
- Search engines are for the masses
  - Much room for improvement
- Implications for personalized search

What is the Value of Personalization?

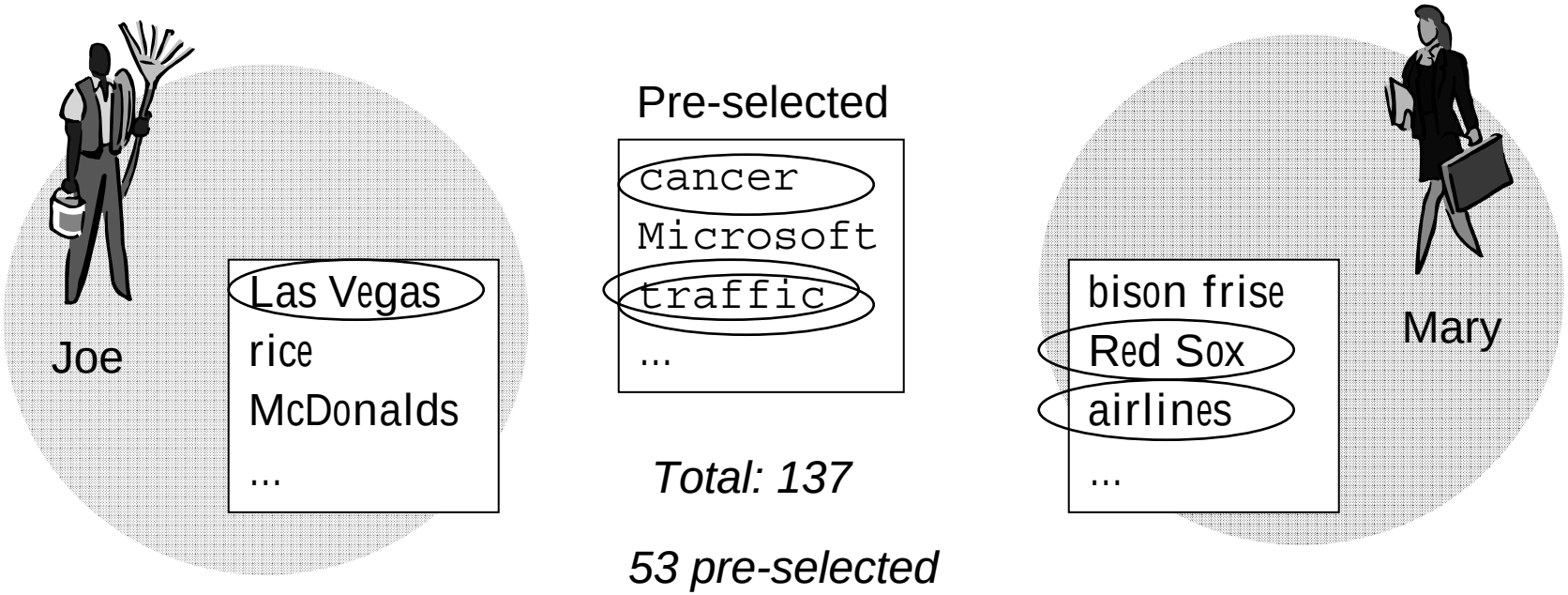
- Do people want different results for the same query?
- How much improvement can be gained by making a generic search engine better?
- How much will personalization help?

Study of Personal Relevancy

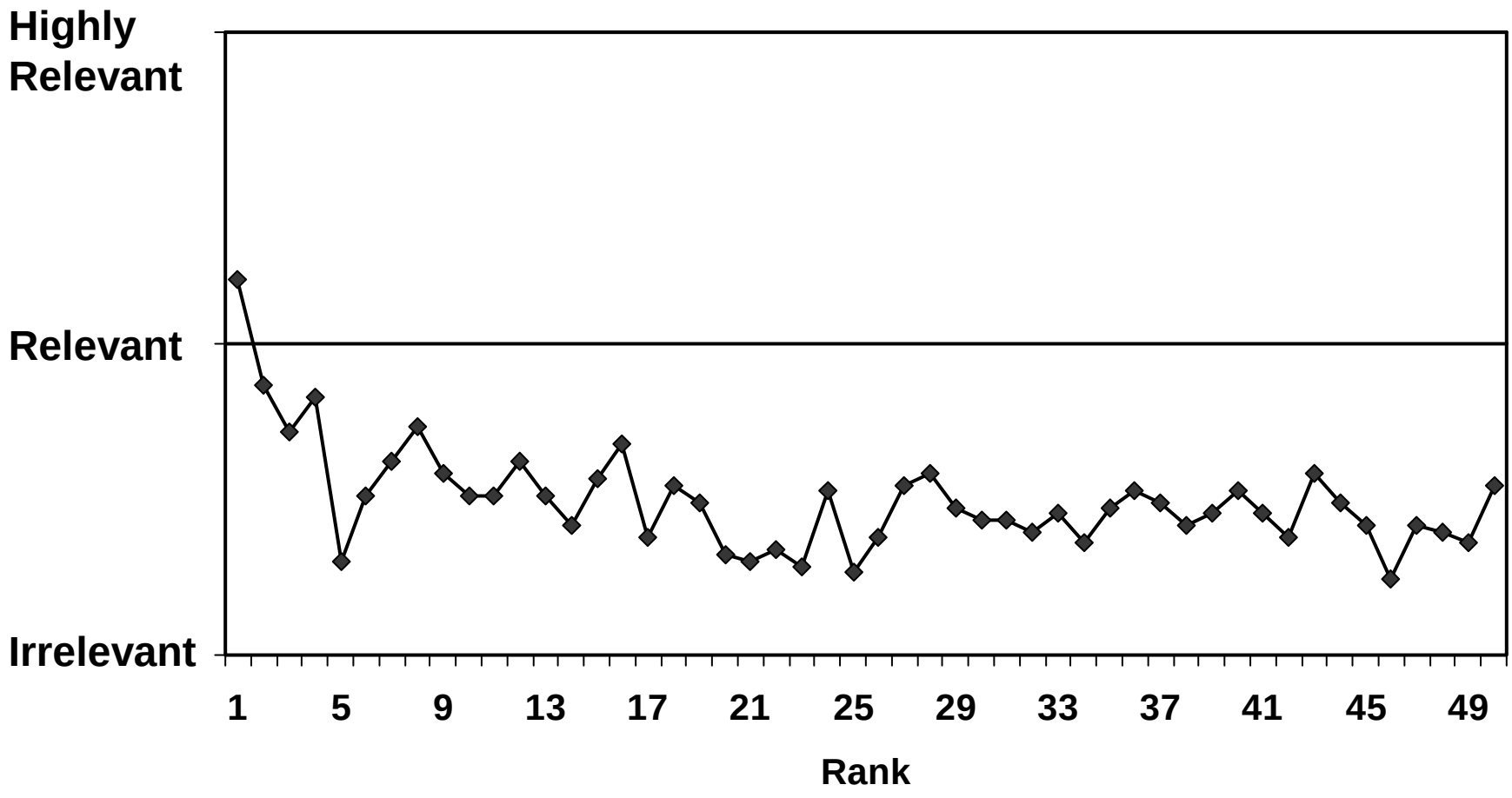
- 15 participants
  - Microsoft employees
  - Managers, support staff, programmers, etc.
- Evaluate 50 results for a query
  - “Highly relevant”
  - “Relevant”
  - “Irrelevant”
- Based on personal preference
- Longer description of information goal
- ~10 queries per person

Study of Personal Relevancy

- Query selection
  - Previously issued query (based on diary)
  - Chose from 10 pre-selected queries

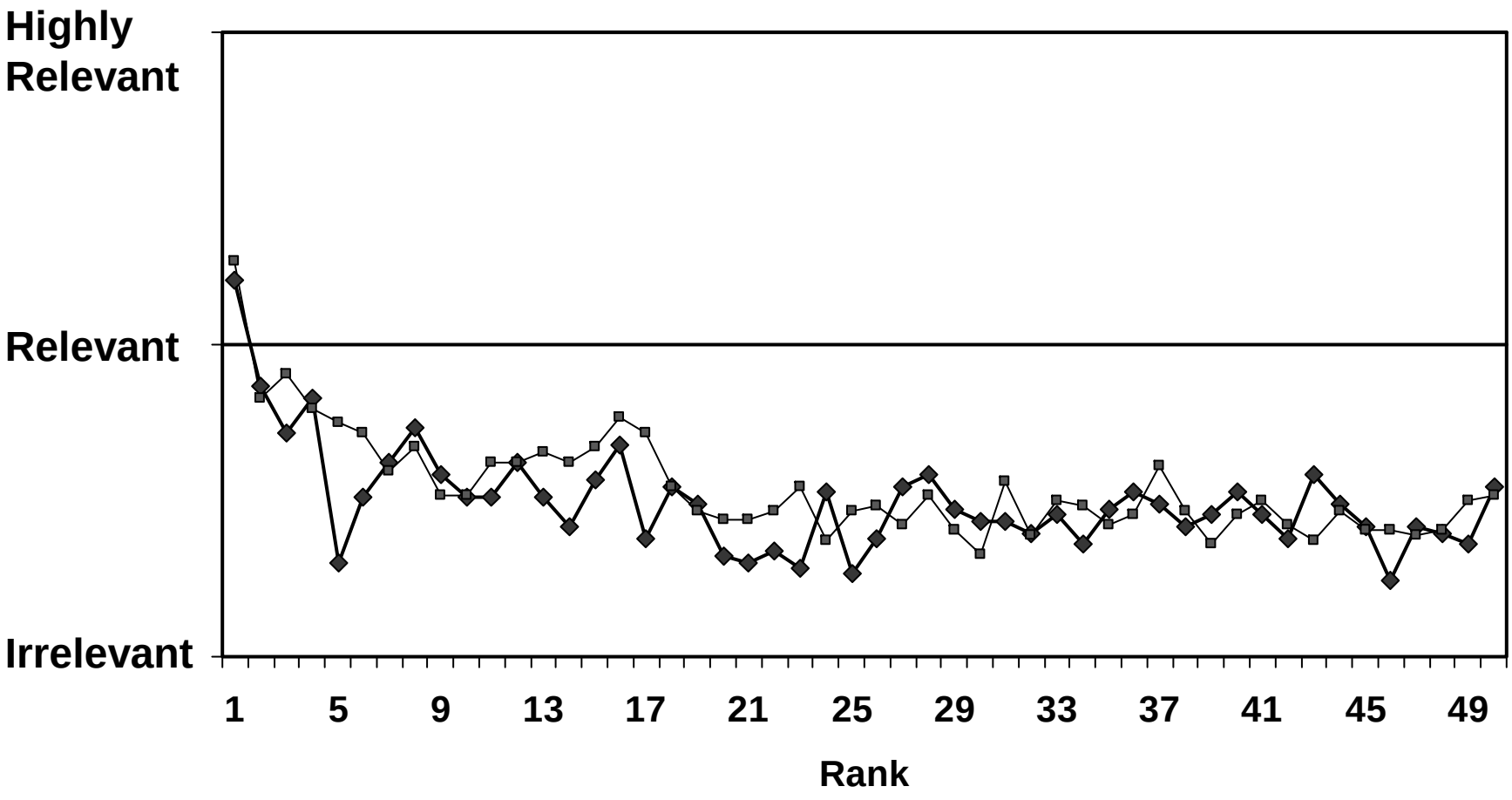


Relevant Results Have Low Rank
(Pre-selected queries)



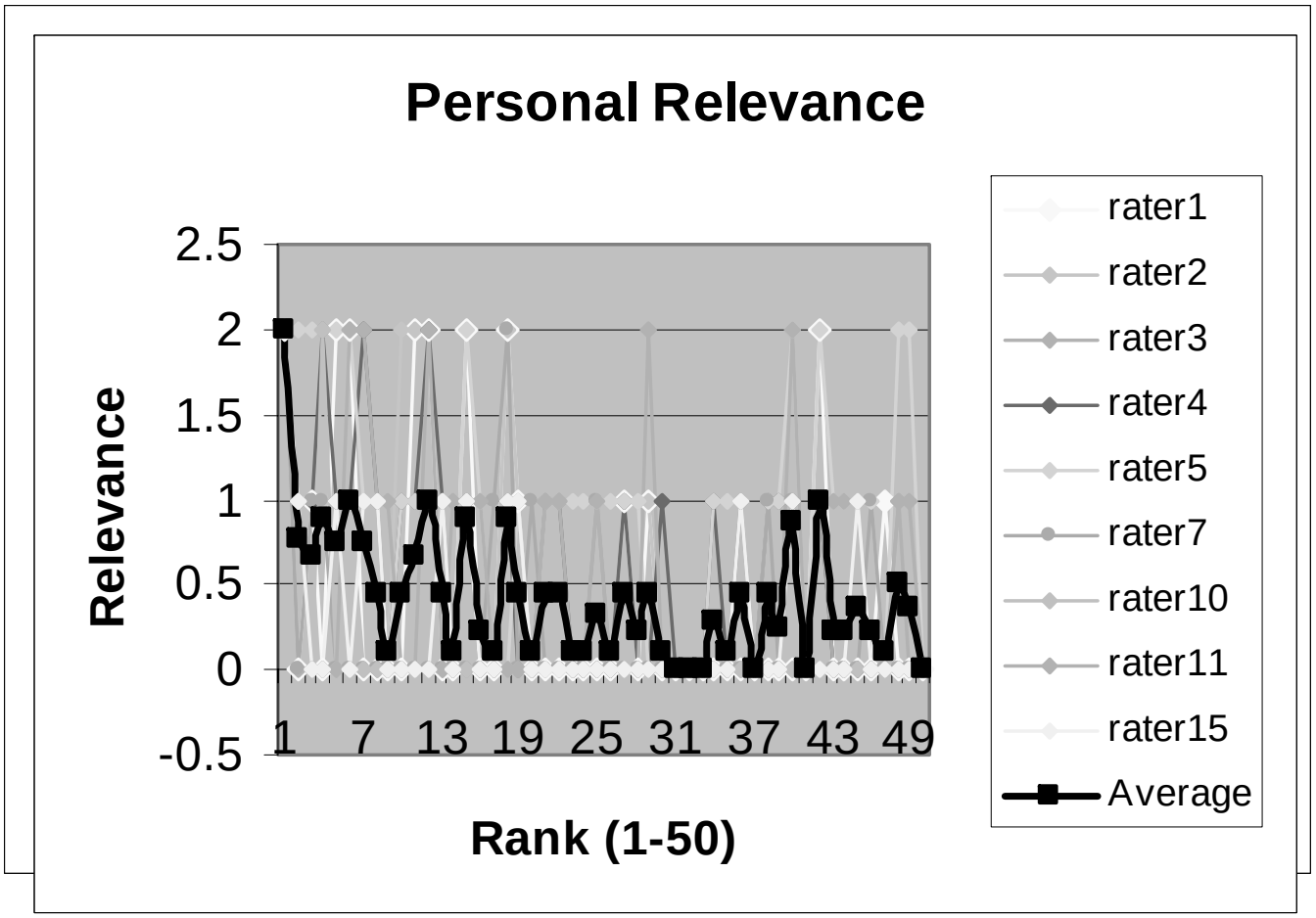
Relevant Results Have Low Rank

(+ Self-selected queries)



Relevant Results Have Low Rank

(Some low ranks highly rated; Raters disagree)



Same Results Rated Differently

- Average inter-rater reliability: 56%
- Different from previous research
  - Belkin: 94% IRR in TREC
  - Eastman: 85% IRR on the Web
- Asked for *personal* relevance judgments, rather than general topical relevance

- Some queries more correlated than others

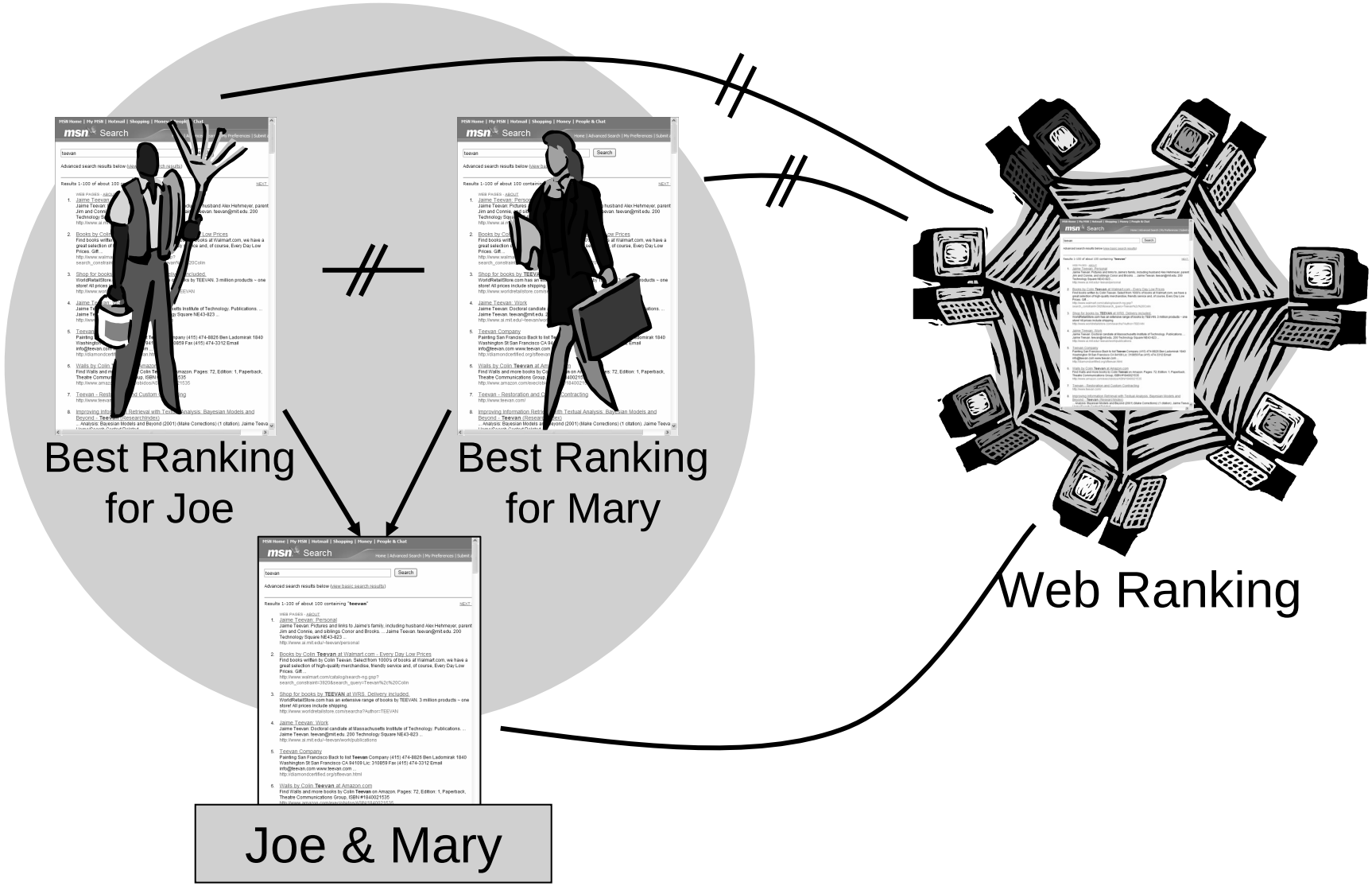
Same Query, Different Meaning/Intent

- Query: *cancer*
- Different meanings
  - “Information about the astronomical/astrological sign of cancer”
  - “information about cancer treatments”
- Different intents
  - “is there any new tests for cancer?”
  - “information about cancer treatments”

Same Intent, Different Evaluation

- Query: *Microsoft*
- Same intents
  - "information about microsoft, the company"
  - "Things related to the Microsoft corporation"
  - "Information on Microsoft Corp"
- 31/50 rated as *relevant* or *highly relevant*
  - All three agree only for www.microsoft.com
  - Only 6/31 do more than one agree
  - Inter-rater reliability: 62%

Search Engines are for the Masses

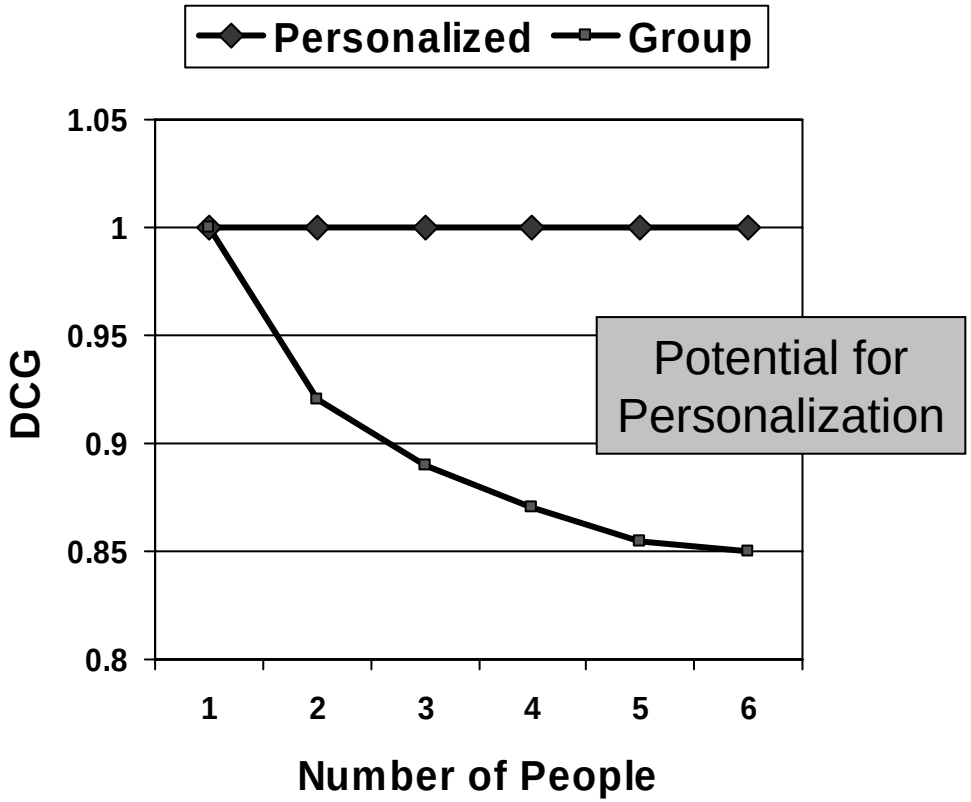


Search Engines are for the Masses

- Best common ranking for a query
  - $\mathrm{DCG}(i) = \begin{cases} \mathrm{Gain}(i), & \text{if } i = 1 \\ \mathrm{DCG}(i-1) + \mathrm{Gain}(i)/\log(i), & \text{otherwise} \end{cases}$
  - Sort results by number marked *highly relevant*, then by *relevant*
- Compare best possible ranking with Web ranking
  - Measure distance with *Kendall-Tau*
  - Number of pair-wise disagreements
  - 0 = same; 1 = reverse order
- Web ranking more similar to common
  - KT(Web, Individual) = 0.47
  - KT(Web, Common) = 0.44

Much Room for Improvement

- Group ranking
  - Best improves on Web by 23%
  - More people → Less improvement
- Personal ranking
  - Best improves on Web by 38%
  - Remains constant



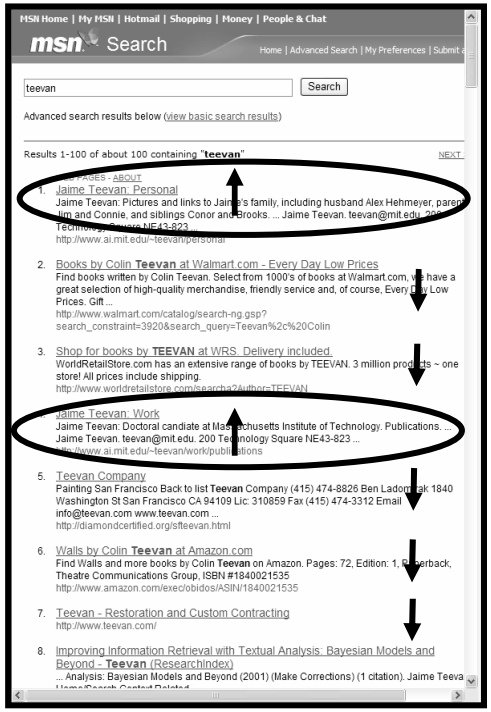
How to Close the Gap

- Aid user to better specify search goal
  - Longer queries; Explicit profile
- Interaction with user
  - Query suggestion; Relevance feedback; Meta-data
- Infer search goal automatically
  - Previous query
  - Richer model (content, usage)
  - And … Re-rank results -> PS Prototype

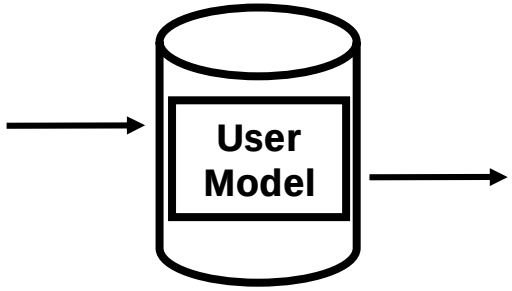
- *Minimize* upfront *work* by user (e.g., no explicit profile); *Maximize* user *control*

Personalized Search (PS): Overview

Step 1:
Retrieve web search results, n>>10



Step 2:
Compute similarity (result, user)



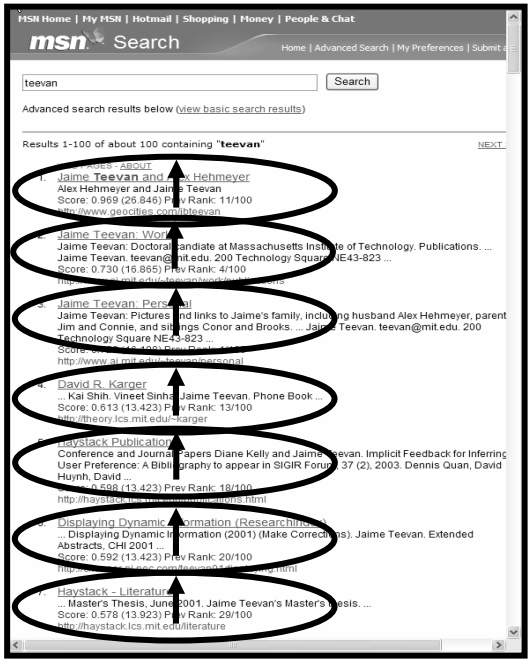
User Model:

* Content + Activity
* Rich and unstructured

Client-Side:

* All storage and processing

Step 3:
Re-rank search results



Summary

- Personal relevance study
- Different ratings even with similar goals
- Making everyone happy means making the individual less happy
- Implications for improving search
  - PS paper to appear at SIGIR 2005

Thank you!
Questions / Comments ???

Contact: Susan Dumais
sdumais@microsoft.com
http://research.microsoft.com