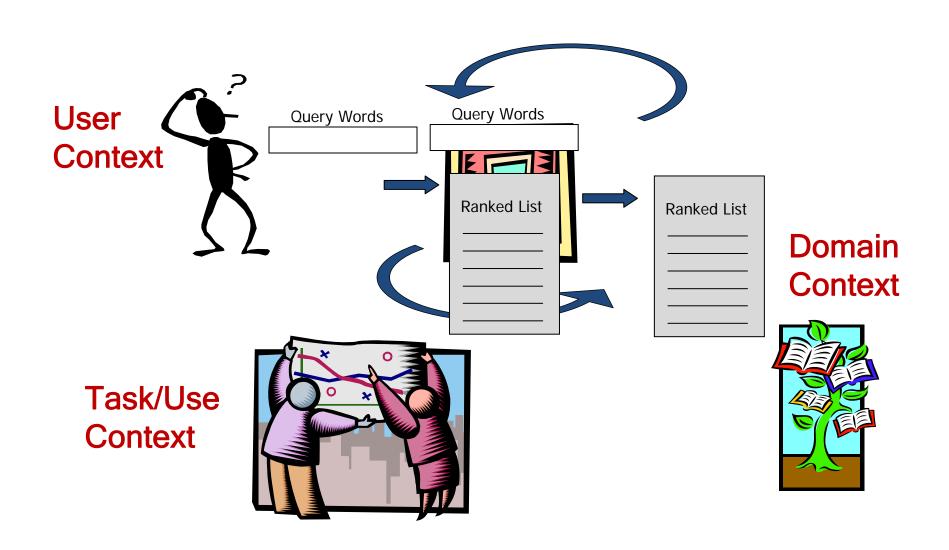
Personalization and Search

Jaime Teevan Microsoft Research

Information Retrievalery

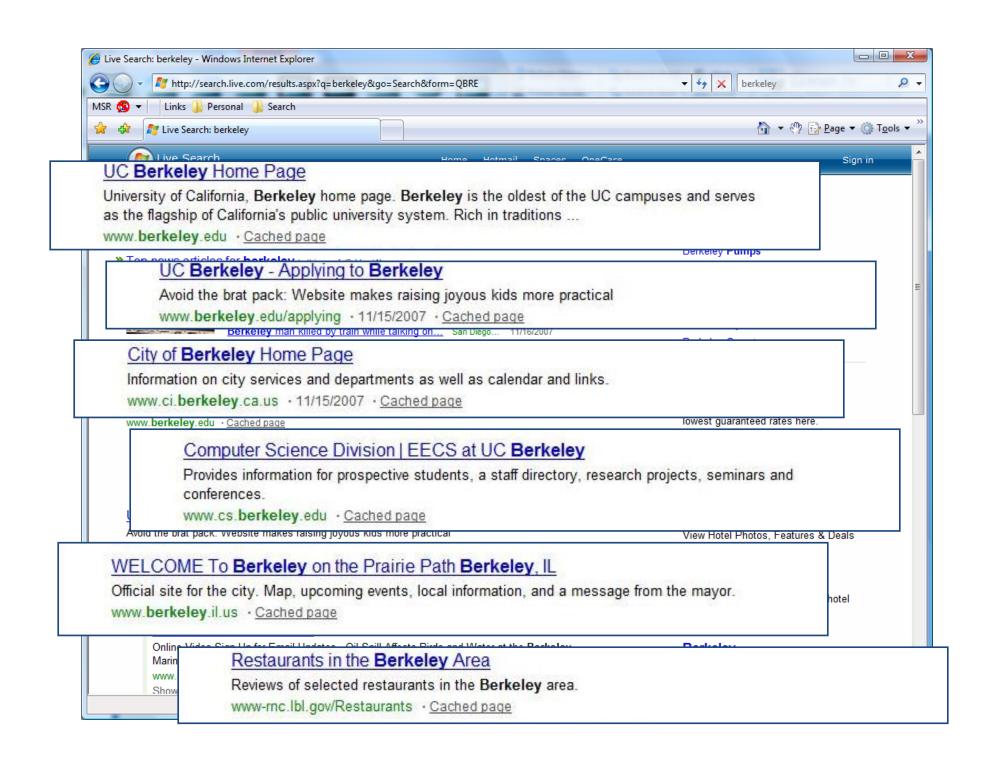


Personalization and Search

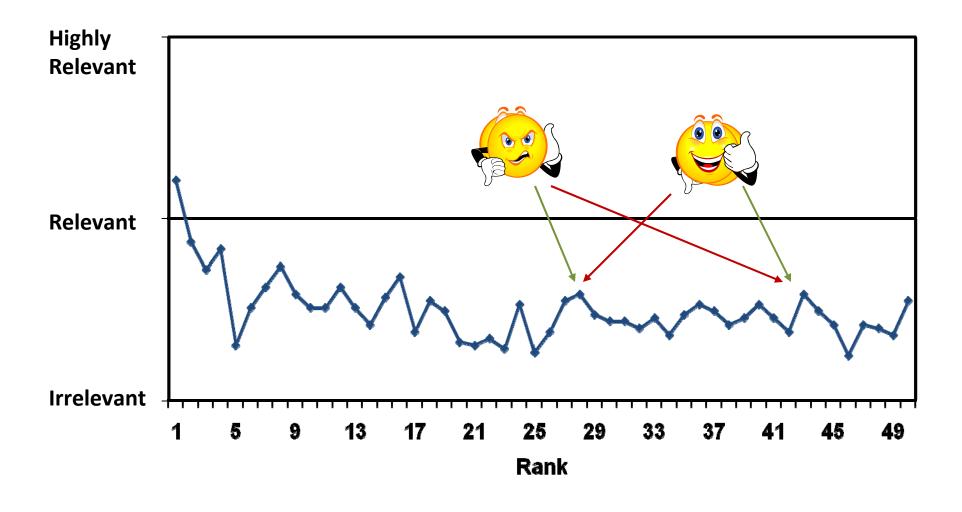
- Measuring the value of personalization
 - Do people's notions of relevance vary?
- Understanding the individual
 - How can we model a person's interests?
- Calculating personal relevance
 - How can we use the model to measure relevance?
- Other ways to personalize search
 - What other aspects can we personalize?

Personalization and Search

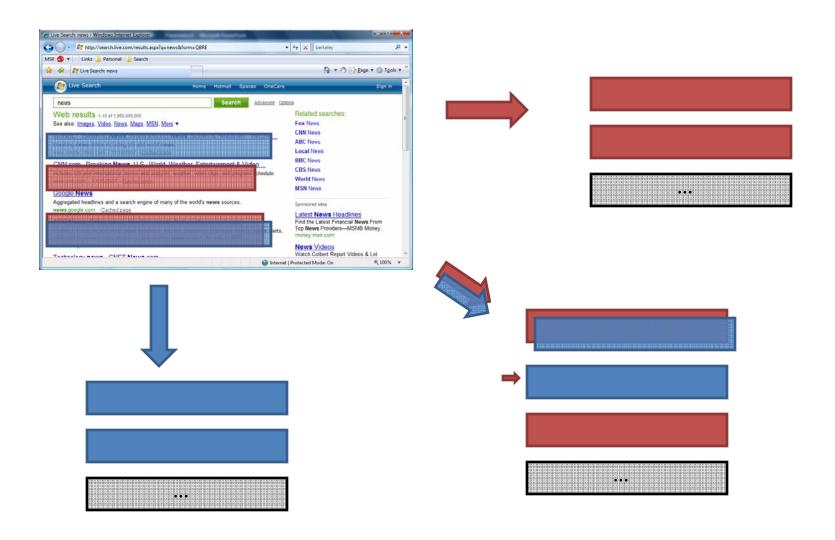
- Measuring the value of personalization
 - An example
 - Lots of relevant results ranked low
 - Best group ranking v. individual ranking
- Understanding the individual
- Calculating personal relevance
- Other ways to personalize search



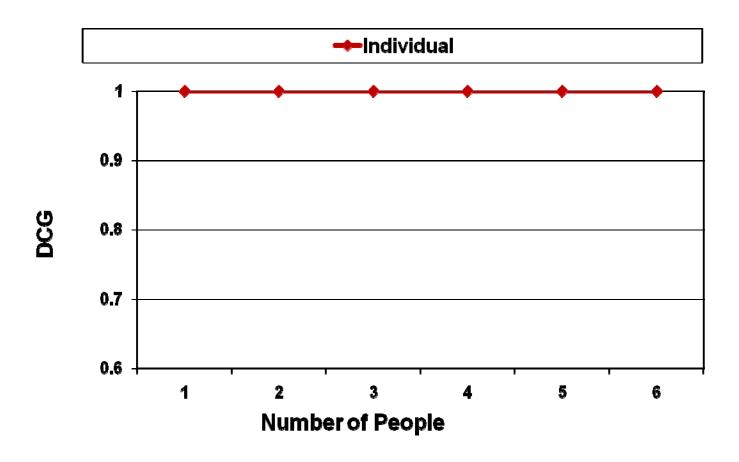
Relevant Content Ranked Low



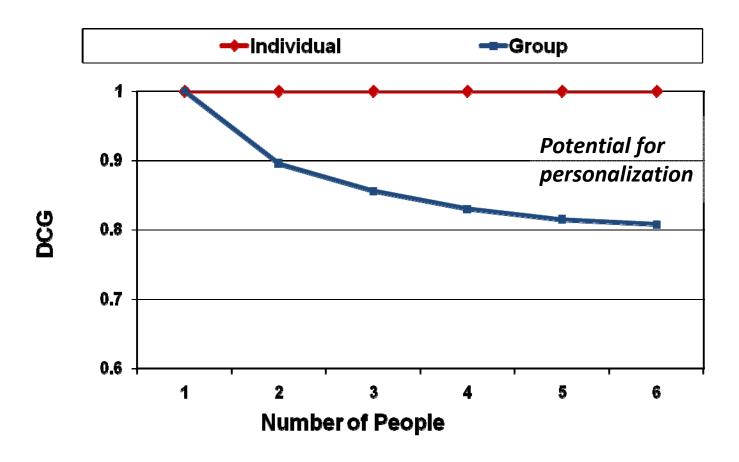
Best Rankings



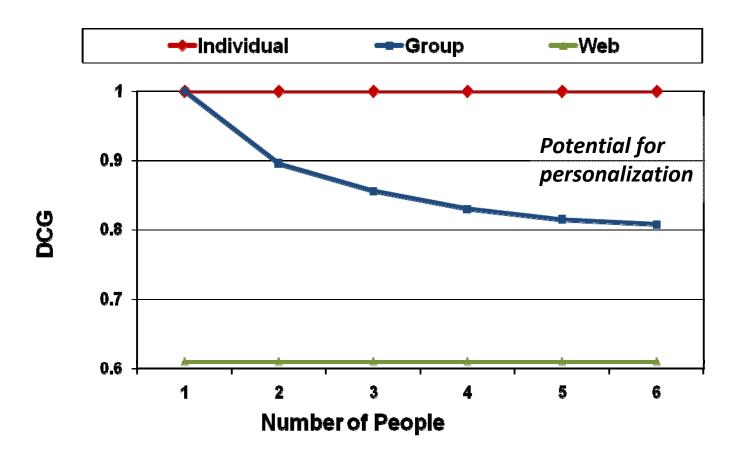
Potential for Personalization



Potential for Personalization



Potential for Personalization



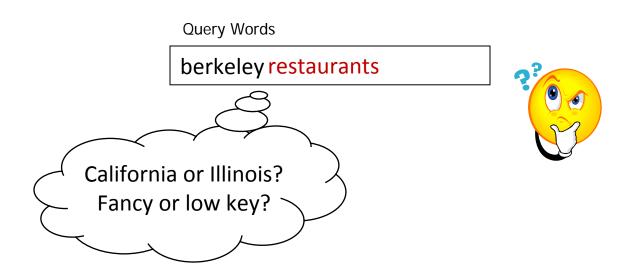
Overview

- Measuring the value of personalization
- Understanding the individual
 - Gather information beyond the query
 - Explicit v. implicit
 - Client-side v. server-side
- Calculating personal relevance
- Other ways to personalize search

Learning More Explicitly v. Implicitly

Explicit

- User shares more about query intent
- User shares more about interests
- Hard to express interests explicitly



Learning More Explicitly v. Implicitly

Explicit

- User shares more about query intent
- User shares more about interests

Intellectual property?

Hard to express interests explicitly

Tobacco and guns

Rock climbing?

Arts	Business	Computers
Games	Health	Home
Kids and Teens	News	Recreation
Reference	Regional	Science
Shopping	Society	Sports

Learning More Explicitly v. Implicitly

Explicit

- User shares more about query intent
- User shares more about interests
- Hard to express interests explicitly

Implicit

- Query context inferred
- Profile inferred about the user
- Less accurate, needs lots of data

Profile Information

- Behavior-based
 - Click-through
 - Personal PageRank
- Content-based
 - Categories
 - Term vector



[topic: computers]



→ [computers: 2, microsoft: 1, click: 4, what: 3, tablet: 1]

Profile Information

- Behavior-based
 - Click-through
 - Personal PageRank
- Content-based
 - Categories
 - Term vector

Server information

- Web page index
- Link graph
- Group behavior







Server-Side v. Client-Side Profile

Server-side

- Pros: Access to rich Web/group information
- Cons: Personal data stored by someone else

Client-side

- Pros: Privacy
- Cons: Need to approximate Web statistics

Hybrid solutions

- Server sends necessary Web statistics
- Client sends some profile information to server

Match Individual to Group

- Can use groups of people to get more data
- Back off from individual → group → all
- Collaborative filtering

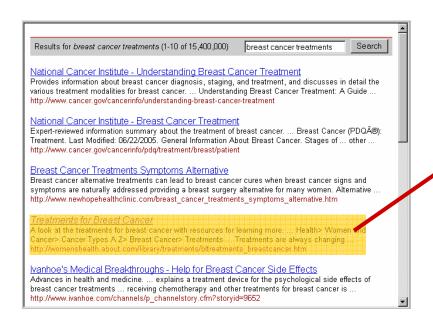


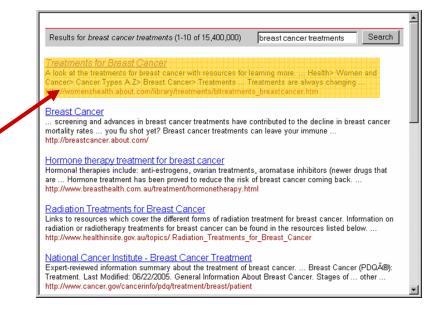
Overview

- Measuring the value of personalization
- Understanding the individual
- Calculating personal relevance
 - Behavior-based example
 - Content-based example
- Other ways to personalize search

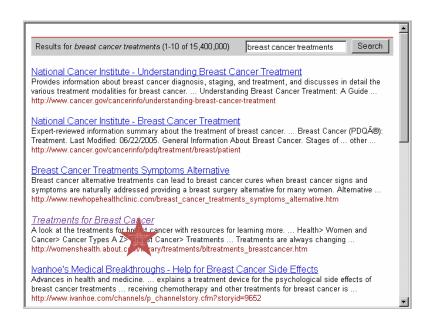
Behavior-Based Relevance

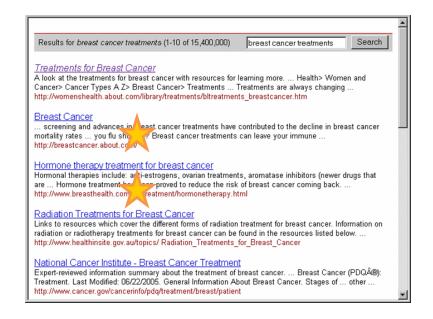
- People often want to re-find
- People have trusted sites
- Boost previously viewed URLs

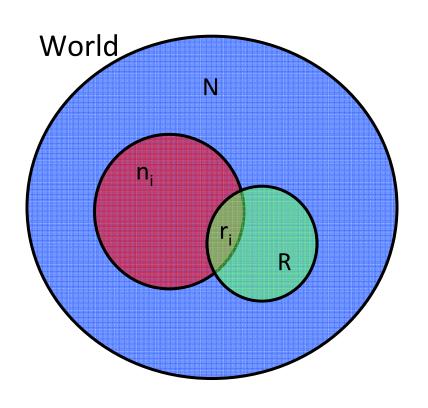




- Explicit relevance feedback
 - Mark documents relevant
 - Used to re-weight term frequencies







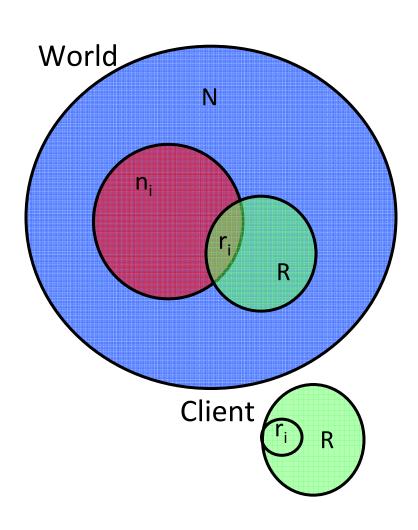
Score =
$$\sum tf_i * w_i$$

$$w_i = log \frac{(N)}{(n_i)}$$

$$w_i = log \frac{(r_i+0.5)(N-n_i-R+r_i+0.5)}{(n_i-r_i+0.5)(R-r_i+0.5)}$$

- Explicit relevance feedback
 - Mark documents relevant
 - Used to re-weight term frequencies
- Lots of information about the user
 - Consider read documents relevant
 - Use to re-weight term frequencies





Score =
$$\sum tf_i * w_i$$

$$w_i = log \frac{(N)}{(n_i)}$$

$$w_i = log \frac{(r_i+0.5)(N-n_i-R+r_i+0.5)}{(n_i-r_i+0.5)(R-r_i+0.5)}$$

$$w_{i} = log \frac{(r_{i}+0.5)(N'-n'_{i}-R+r_{i}+0.5)}{(n'_{i}-r_{i}+0.5)(R-r_{i}+0.5)}$$

Where: N' = N+R,
$$n_i' = n_i + r_i$$

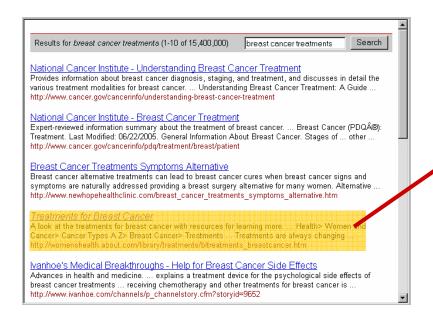
Personalization Performance

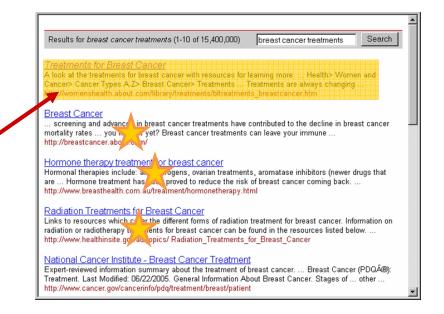
- Personalized search hard to evaluate
- Mostly small improvements despite big gap
- Identify ambiguous queries
 - Personalize: "berkeley"
 - Don't personalize: "uc berkeley homepage"
- Identify easily personalized queries
 - Re-finding queries

Other Ways to Personalize

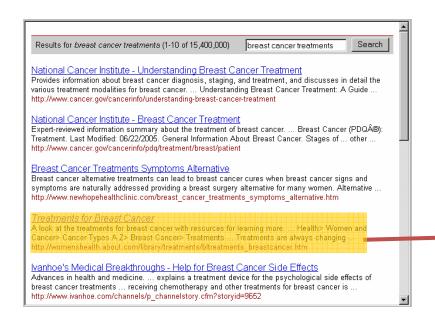
- Measuring the value of personalization
- Understanding the individual
- Calculating personal relevance
- Other ways to personalize search
 - Match expectation for re-finding queries
 - Personalized snippets

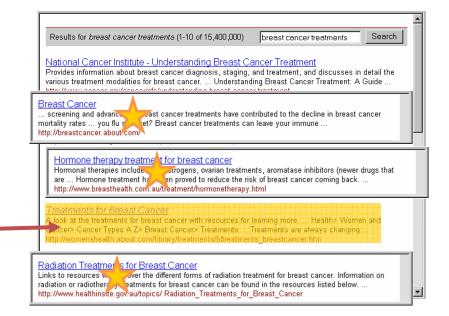
Ranking Results for Re-Finding





Ranking Results for Re-Finding





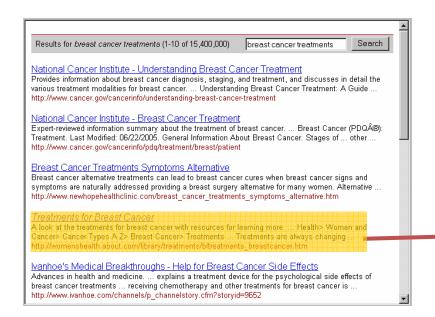
People Don't Notice Change

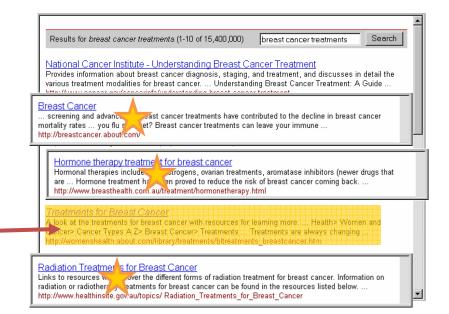


People Don't Notice Change



People Don't Notice Change





Snippets to Support Re-Finding

Query: "winery"

Winery - Wikipedia, the free encyclopedia

A **winery** is a building or property that produces wine, or a business involved in the production of wine, such as a wine company. Some wine companies own many wineries. Besides wine making equipment ...

en.wikipedia.org/wiki/Winery

If the person has visited the page before:

Winery - Wikipedia, the free encyclopedia Last visit: November 14, 2007

A **winery** is a building or property that produces wine, or a business involved in the production of wine, such as a wine company. Some wine companies own many wineries. Besides wine making equipment ...

Snippets to Support Re-Finding

Query: "winery"

Winery - Wikipedia, the free encyclopedia

A **winery** is a building or property that produces wine, or a business involved in the production of wine, such as a wine company. Some wine companies own many wineries. Besides wine making equipment ...

en.wikipedia.org/wiki/Winery

If the person has visited the page before:

Winery - Wikipedia, the free encyclopedia Last visit: November 14, 2007

New content: It has been suggested that Winery wastewater be merged into this article or section.

Interest-Based Snippets

Query: "winery"

Winery - Wikipedia, the free encyclopedia

A **winery** is a building or property that produces wine, or a business involved in the production of wine, such as a wine company. Some wine companies own many wineries. Besides wine making equipment ...

en.wikipedia.org/wiki/Winery

If the person is interested in Maui:

Winery - Wikipedia, the free encyclopedia

A **winery** is a building or property that produces wine, or a business involved in the production of wine, such as a wine company... For example, in Maui there is a pineapple **winery**. ...

Interest-Based Snippets

Query: "winery"

Winery - Wikipedia, the free encyclopedia

A **winery** is a building or property that produces wine, or a business involved in the production of wine, such as a wine company. Some wine companies own many wineries. Besides wine making equipment ...

en.wikipedia.org/wiki/Winery

If the person is interested in Maui:

Winery - Wikipedia, the free encyclopedia

A **winery** is a building or property that produces wine, or a business involved in the production of wine, such as a wine company... For example, in <u>Maui</u> there is a <u>pineapple</u> **winery**. ...

Summary

- Measuring the value of personalization
 - There's a big gap between group and individual
- Understanding the individual
 - Building a profile, explicit v. implicit
- Calculating personal relevance
 - Relevance feedback, boost click through
- Other ways to personalize search
 - Rank based on expectation, personalized snippets