A Framework for Personalization:

When do you want to go *Where Everybody Knows Your Name* (and mailing address, and preferences, and last 50 web pages visited)?

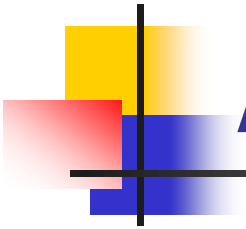
Susan Dumais

Microsoft Research

sdumais@microsoft.com

http://research.microsoft.com/~sdumais

Delos-NSF Workshop: June 18-20, 2001



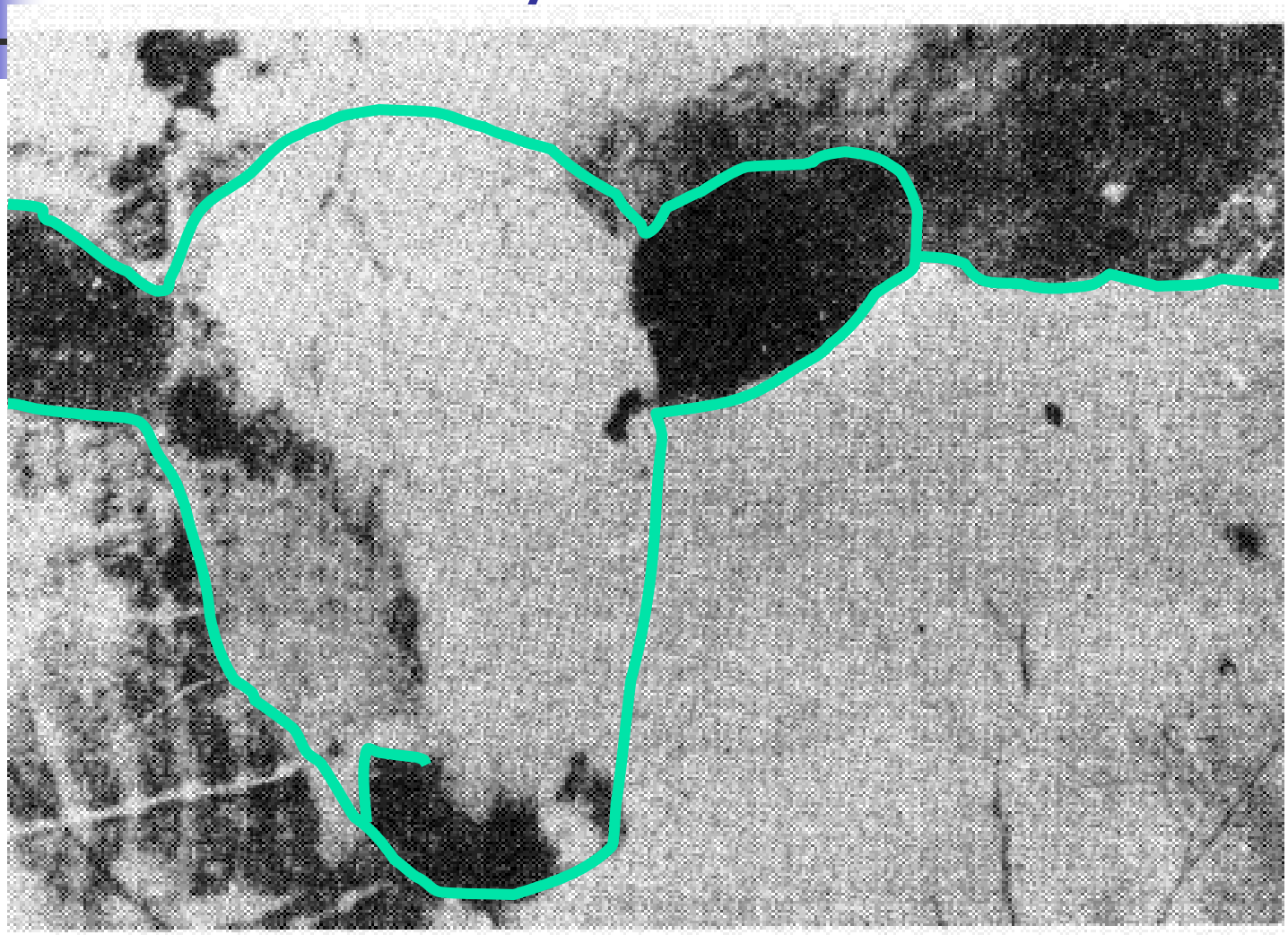
A Working Definition

- Outcome(t) = f(Action(t), PersonalHistory(t-n))
- Examples,
  - Relevance feedback
  - Content-based filtering
  - Collaborative filtering
  - Caching, history lists, auto completion, MRU
  - Implicit queries, Rememberance Agent, Watson, Kenjin
  - MyYahoo!, MyAOL, MyMSN, MyLibrary, etc.
  - AltaVista's MySearch, iLOR

  …

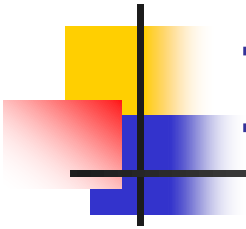
A Demonstration:

What do you see?



Many Kinds of Individual Differences

- Task – "info need"
  - Short-term, relevance feedback
  - Long-term, content-based filtering
- Preferences, e.g., CF
- Expertise, domain and application
- Cognitive aptitudes
  - Verbal, spatial, reasoning skills, etc.
- Demographics
  - Age, major, gender, location, etc.
- Cognitive styles, personality and affect



Individual Differences Are …

- Large
- Systematic
- Systems can often be modified to accommodate
  - E.g., robust systems
  - E.g., personalization

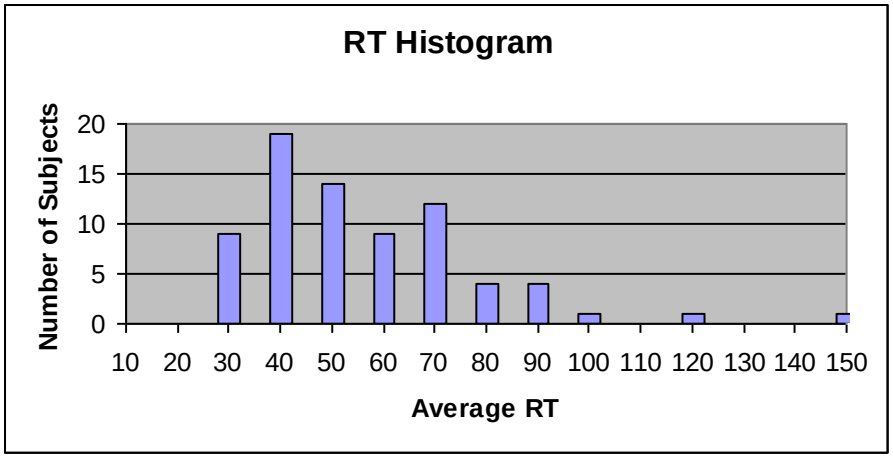
How Big Are Individual Diffs?

- E.g., Web searching (Chen & Dumais, CHI 2000)
  - 74 participants; Intermediate web/search experience
  - 30 search tasks (e.g., *Home page for "Seattle Weekly")*
  - Average RT (seconds) = 52.3 seconds

- Individual subjects' average RT:
  - 69, 30, 76, 48, 29, 68, 69, 49, 75, 62, 64, 69, 26, 89, 50, 44, 54, 35, 39, 30, 71, 56, 28, 59, 36, 67, 93, 37, 39, 49, 28, 89, 37, 36, 31, 47, 66, 62, 51, 30, 40, 38, 31, 70, 37, 36, 36, 88, 41, 50, 84, 68, 42, 58, 34, 25, 23, 22, 41, 62, 35, 41, 41, 60, 36, 56, 78, 144, 43, 58, 58, 45, 38, 115

Characterizing Indiv Diffs

- Histogram



- Max:Min 144, 22 = 6.5:1
- Q3:Q1 66, 36 = 1.8:1
- SD/X .42

Example Individual Diffs

Example Individual Differences			
Task	**Max/Min**	**Q3/Q1**	**SD/X**
Grocery Cashiers	2:1	1.5:1	
Text Editing	5:1	2:1	0.4
Information Retrieval	8:1	2.5:1	0.6
Computer Programming	20:1	3:1	0.8

Individual Diffs Correlated w/ Performance in HCI/IR Tasks

- Experience – both application and domain
- Reasoning (Egan et al.; Card et al.; Greene et al.)
- Spatial abilities (Egan & Gomez; Vicente et al.; Stanney & Salvendy; Allen)
- Academic major (Borgman)
- Verbal fluency (Dumais & Schmitt)
- Reading comprehension (Greene et al.)
- Vocabulary (Vicente et al.)
- Age (Egan et al.; Greene et al.; Konvalina et al.)
- Personality and affect
- Gender

Framework for Identifying and Accommodating Indiv Diffs

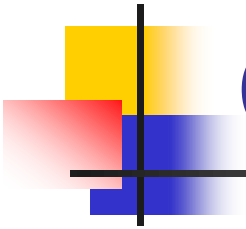
- **Assay** – which user characteristics predict differences in performance; many studies stop here
- **Isolate** – isolate the source of variation to a specific sub-task or design component
- **Accommodate** – do something about it
  - Often harder than you think …
    - E.g., Spatial ability and hierarchy navigation
    - E.g., Expertise
  - Evaluate!!!

Greene et al. *No IFs, ANDs, or ORs: A Study of Database Querying*

- Task: *Find all employees who either work in the toy department or are managed by Grant, and also come from the city London.*
- SQL – fixed syntax, logical operators, parentheses
  - E.g.,

```
SELECT      Name
FROM        Employee
WHERE       (Department = Toy
OR          Manager = Grant)
AND         City = London
```

- TEBI – just need attribute names and values; recognize alternatives from system-generated table
  - E.g., Name, Department = Toy, Manager = Grant, City = London



Greene et al. (Assay)

- Assessed individual characteristics:
  - Age, spatial memory, reasoning, integrative processing, reading comprehension & vocabulary
- Found large effects of:
  - Integrative processing (on accuracy, for SQL interface)
  - Age (on time, for SQL interface)

Greene et al.

<percent correct x integrative processing>

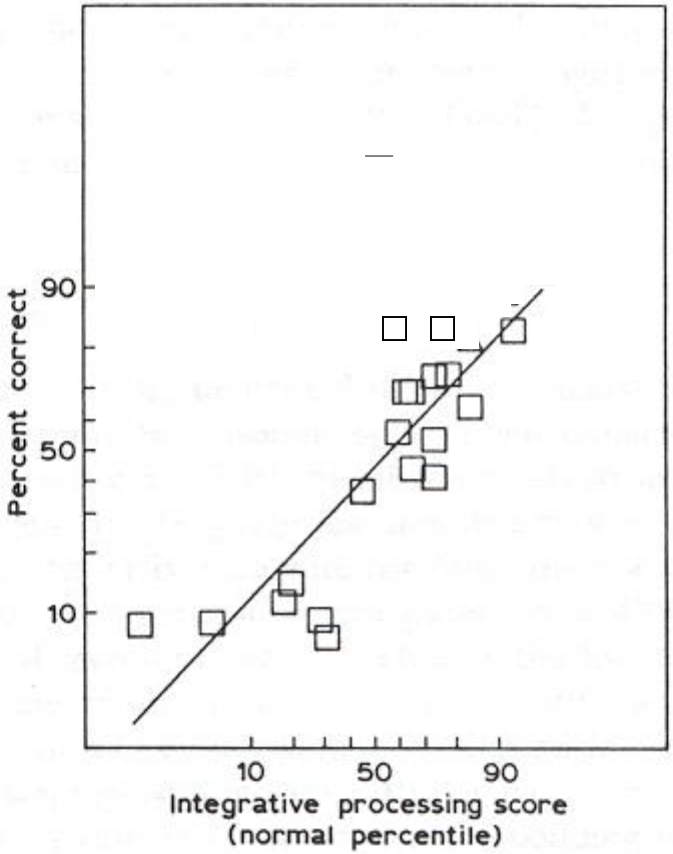
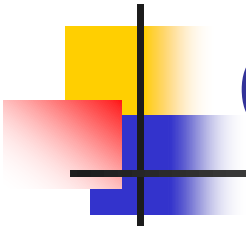


FIGURE 2. This modified partial residual plot shows percent correct as a function of integrative processing skill (a measure of reasoning ability), for two language conditions (TEBI Concrete and SQL Generation). The lines represent the predicted score holding the other individual difference variables at their means. The points represent individual subjects' scores minus the contribution of the other individual difference variables. The slopes of the lines indicate the importance of integrative processing to prediction of performance in the two language conditions. The difference between the points and the line is an indication of goodness of fit. △, TEBI; □, SQL.

"AND + OR" queries



Greene et al. (Isolate)

- Examined two possible sources of difficulties
  - Interpreting the query
  - Specifying the query in a formal notation or query language

Example TEBI Table

Employee			
Name	**Department**	**Manager**	**City**
Jones	Toy	Grant	London
Allan	Toy	Grant	Athens
Doe	Toy	Buford	London
Smith	Toy	Buford	Athens
Price	Cosmetics	Grant	London
Hall	Cosmetics	Grant	Athens
Kelso	Cosmetics	Buford	London
Lee	Cosmetics	Buford	Athens

Greene et al.

<percent correct x integrative processing>

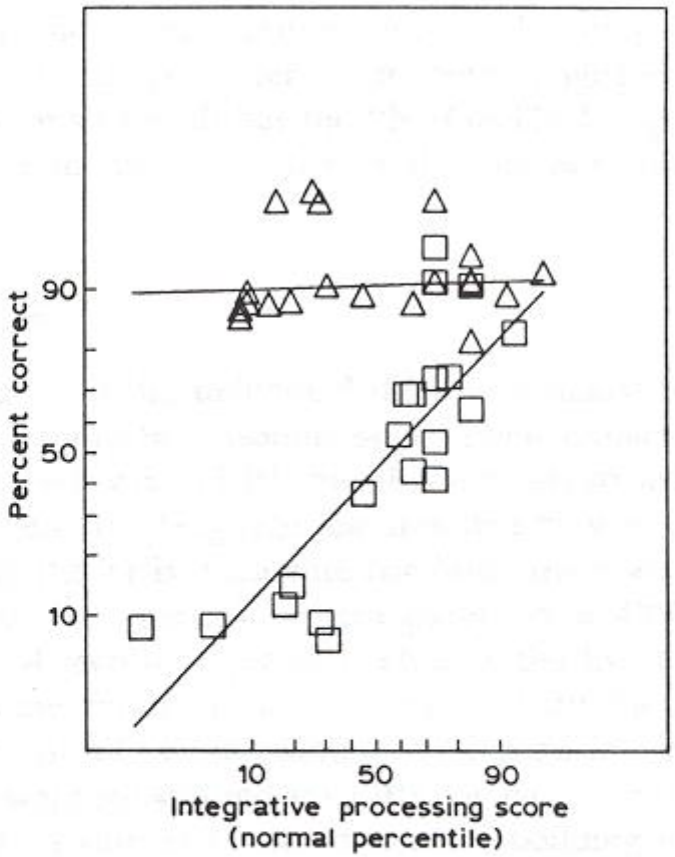
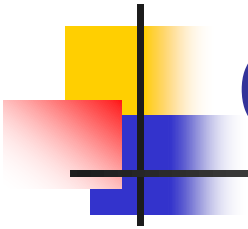


FIGURE 2. This modified partial residual plot shows percent correct as a function of integrative processing skill (a measure of reasoning ability), for two language conditions (TEBI Concrete and SQL Generation). The lines represent the predicted score holding the other individual difference variables at their means. The points represent individual subjects' scores minus the contribution of the other individual difference variables. The slopes of the lines indicate the importance of integrative processing to prediction of performance in the two language conditions. The difference between the points and the line is an indication of goodness of fit. △, TEBI; □, SQL.

"AND + OR" queries



Greene et al. (Accommodate)

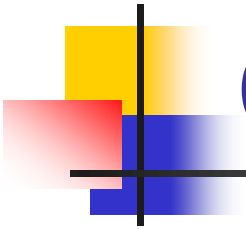
- SQL – hard, especially for some users
- TEBI – new query specification language
  - Improved performance overall
  - Reduced many dependencies on reasoning skills and age
  - “Robust interface”

How to Accommodate?

- Robust interfaces: A new design improves the performance for all
  - E.g., Greene et al.'s *TEBI* interface
  - E.g., Dumais & Schmitt's *LikeThese* interface
- Training:
- Personalization: Different interfaces/systems for different people
  - Group level - E.g., Grundy prototypes, I3R sterotypes, Expert/Novice
  - Individual level
  - Task (Info Need) level

Personalization Framework

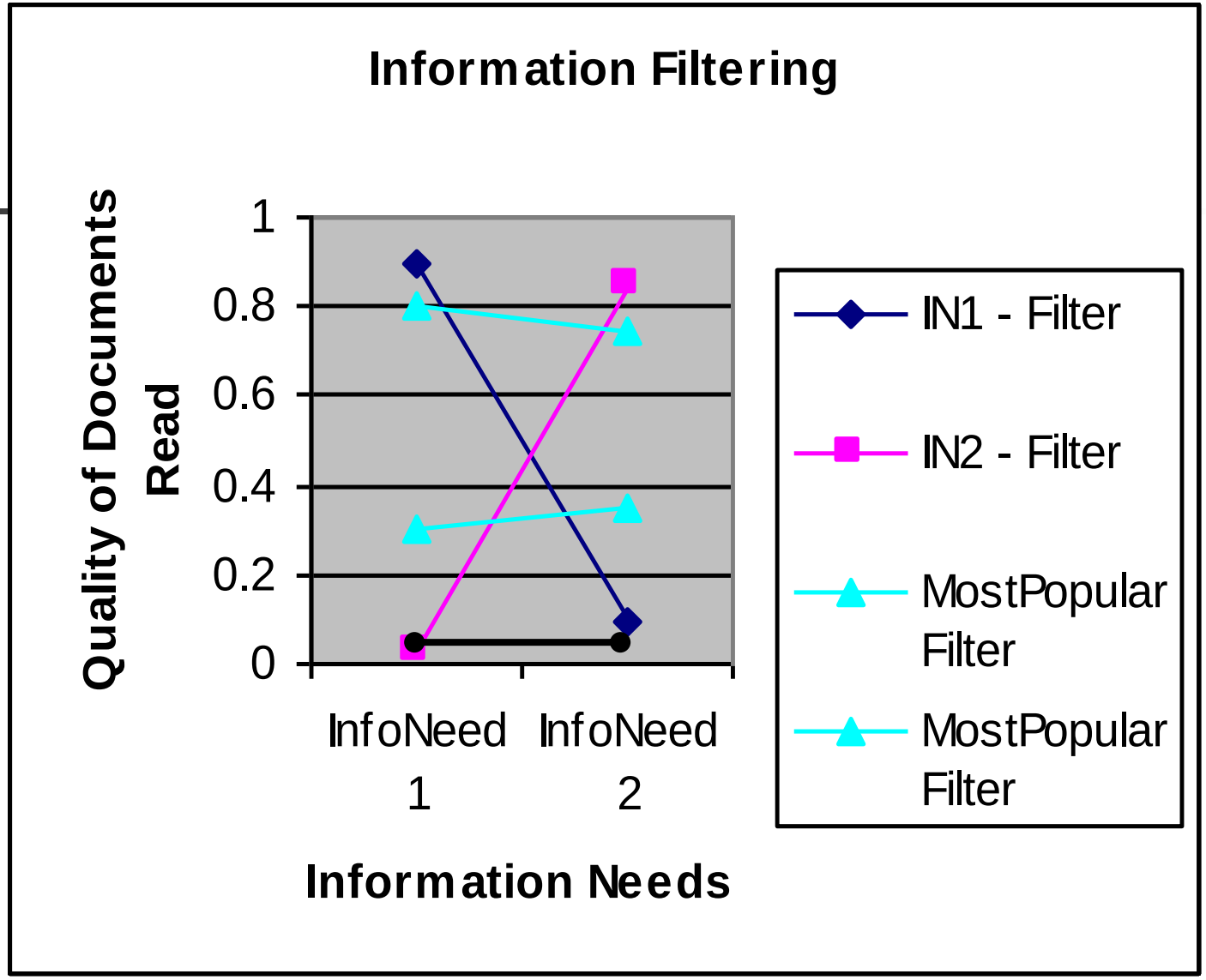
- Characteristics for personalization
  - Expertise, Task, Preferences, Cog Aptitudes, Demographics, Cog Styles, Etc.
- Assay: How specified/modeled?
  - Implicit, Explicit, Interaction
  - Stability over time?
    - Long-term, short-term
- Accommodate: What to do about it?
  - Many ways of accommodating
  - Evaluation
    - Benefits of correct assessment and accommodation
    - Costs of mis-assessment

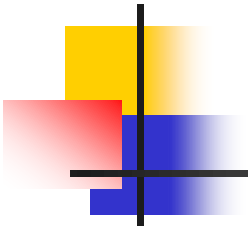


Content-Based Filtering

Match new content to standing info need

- Assay:
  - Explicit or Implicit profile specification?
    - Ongoing feedback?
  - How rapidly does profile it change?
- Accommodate:
  - Match profile against stream of new docs
    - Reduce number of docs to view
    - Return more relevant docs
  - Benefits/Costs


Information Filtering
Quality of Documents Read
1
0.8
0.6
0.4
0.2
0
InfoNeed 1
InfoNeed 2
Information Needs
IN1 - Filter
IN2 - Filter
MostPopular Filter
MostPopular Filter



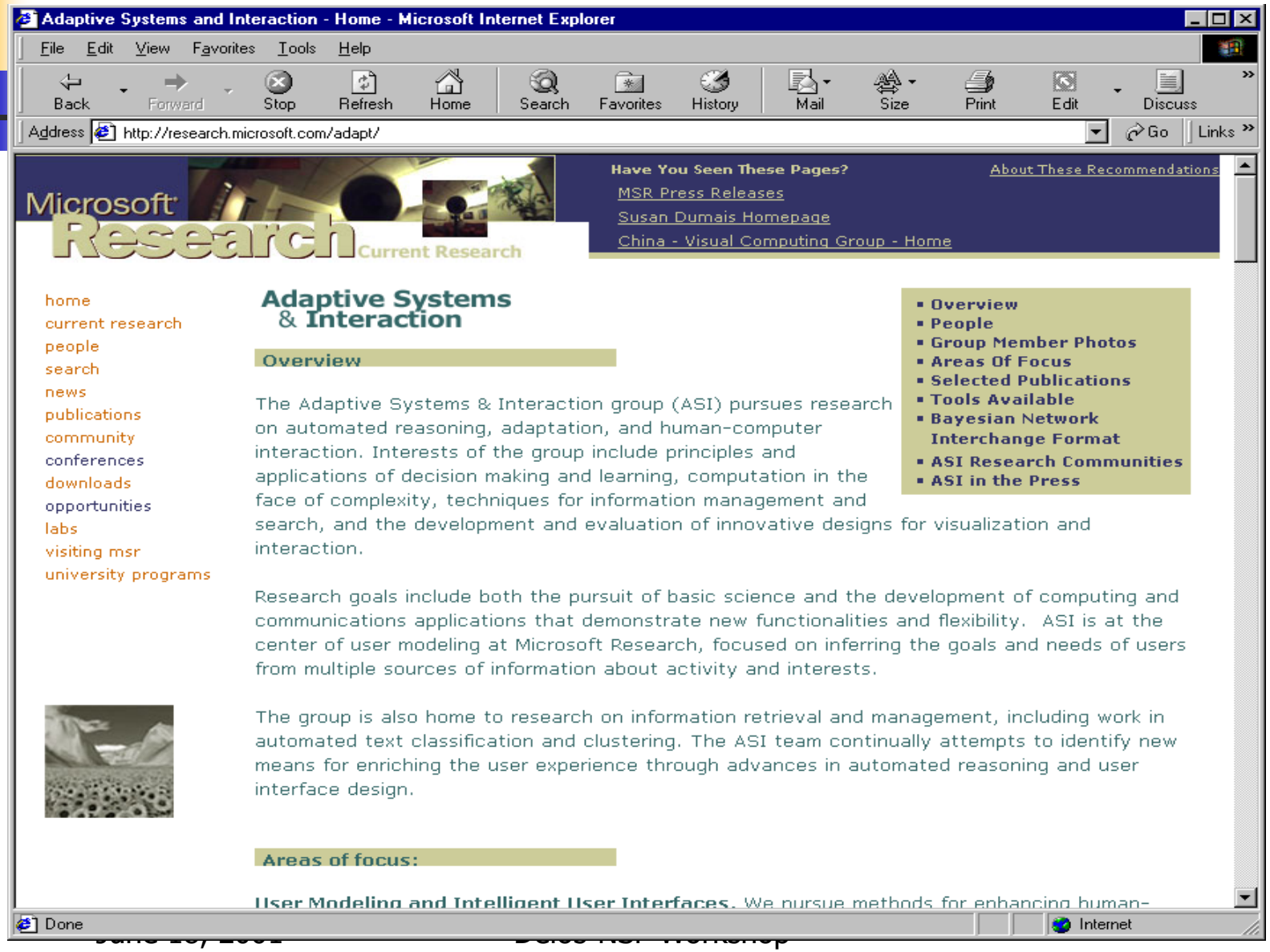
ASI Examples

- Collaborative Filtering
- Implicit/Background Query
- Lumiere
- Temporal Query Patterns

Example: *MSRweb Recommender*
<Breese, Heckerman, Kadie, UAI'98>

- Collaborative filtering algorithms
  - Bayesian network
  - Correlation+
  - Vector similarity
  - Bayesian clustering
  - Popularity
- Test collections
  - Each Movie
  - Nielsen
  - Microsoft.com
- Predicted
  - Individual scores
  - Ranked score

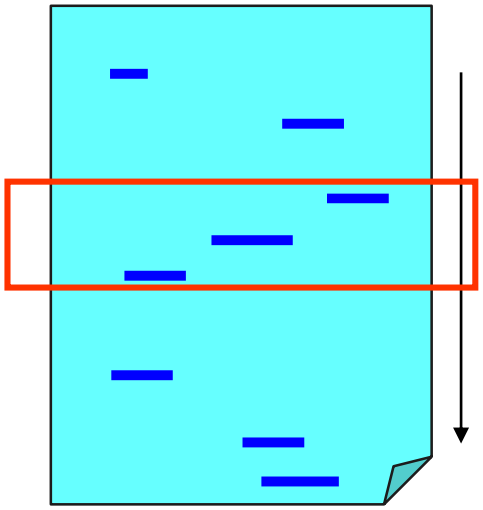
Example: *MSRweb Recommender*

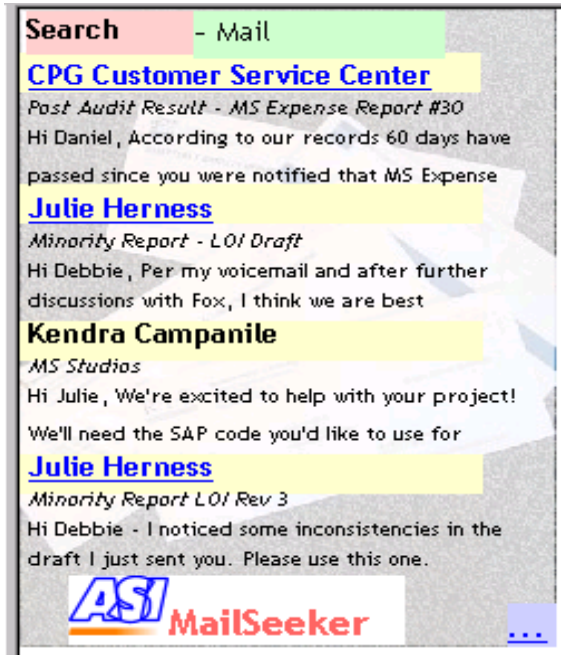


Example: *Background Query*
<Dumais et al., Horvitz et al.>

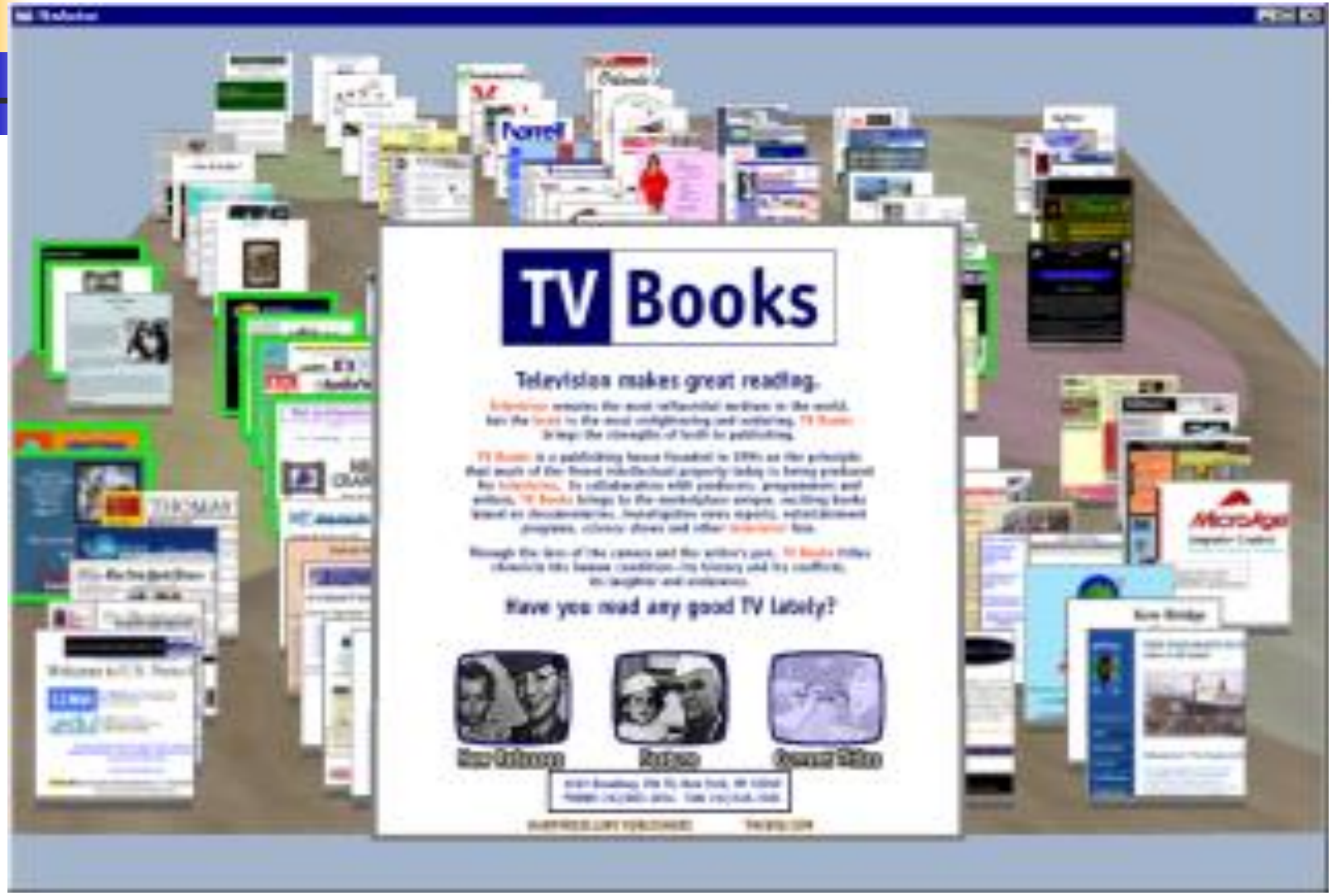
- Identify content at user's focus of attention
- Formulate query, provide related information as part of normal work flow
- Background, implicit queries

Consider doc structure, basic scroll, dwell patterns





Data Mountain with Implicit Query results (highlighted pages to left of selected page)



Implicit Query Results

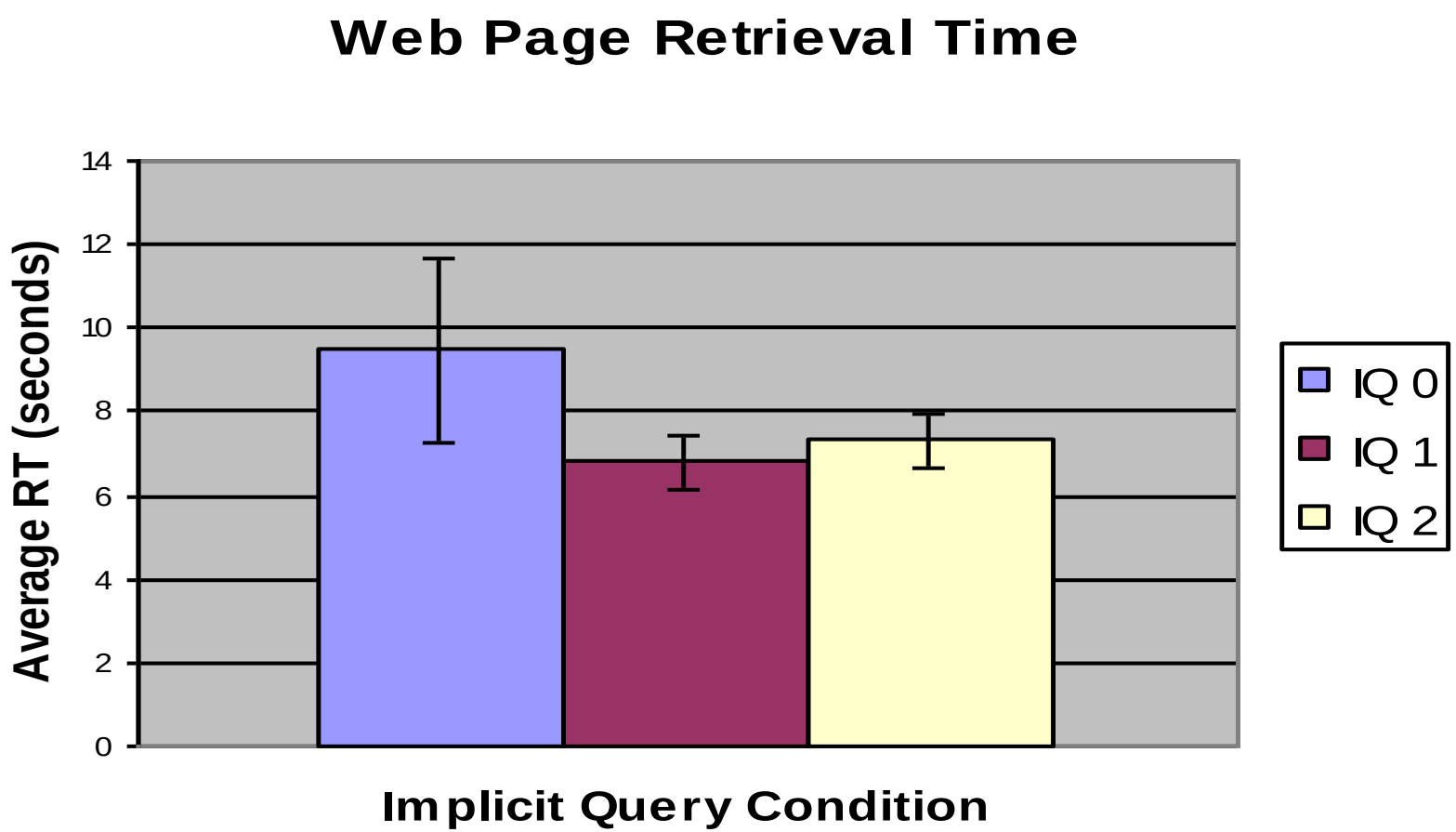
- Filing strategies

	Filing Strategy		
IQ Condition	Semantic	Alphabetic	No Org
IQ0: No IQ	11	3	1
IQ1: Co-occur based	8	1	0
IQ2: Content-based	10	1	0

- Number of categories

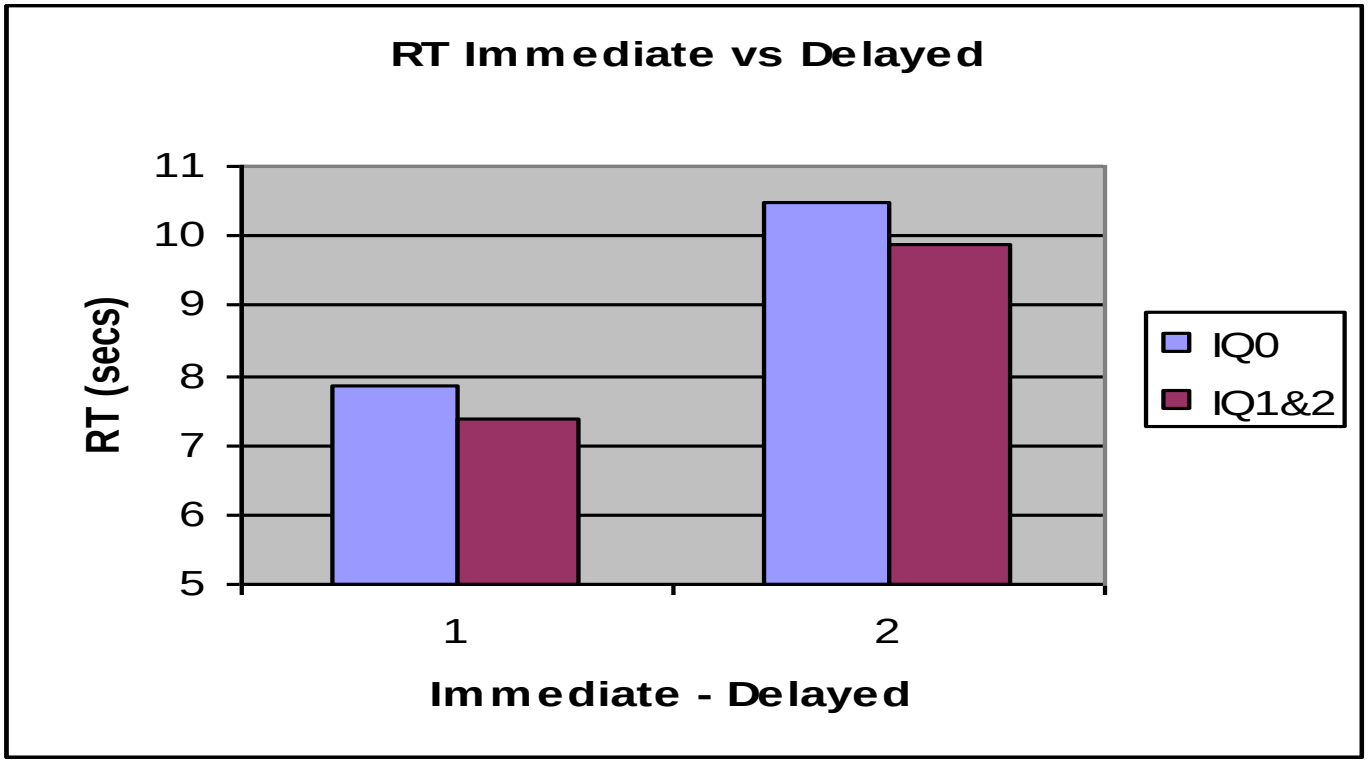
IQ Condition	Average Number of Categories (std error)
IQ0: No IQ	10.0 (3.6)
IQ1: Co-occur based	15.8 (5.8)
IQ2: Content-based	13.6 (5.9)

Implicit Query Results



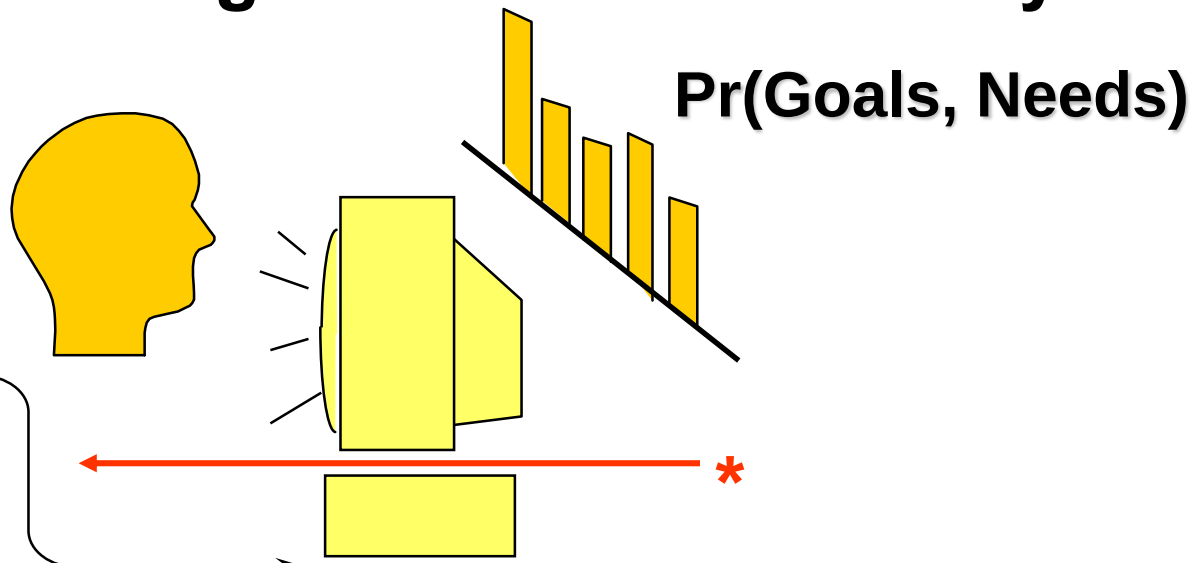
Implicit Query Results (Delayed Retrieval, 6 months)

- 17 subjects (9 IQ1, 8 IQ1&2)



Example: *Lumiere*
<Horvitz, et al.>

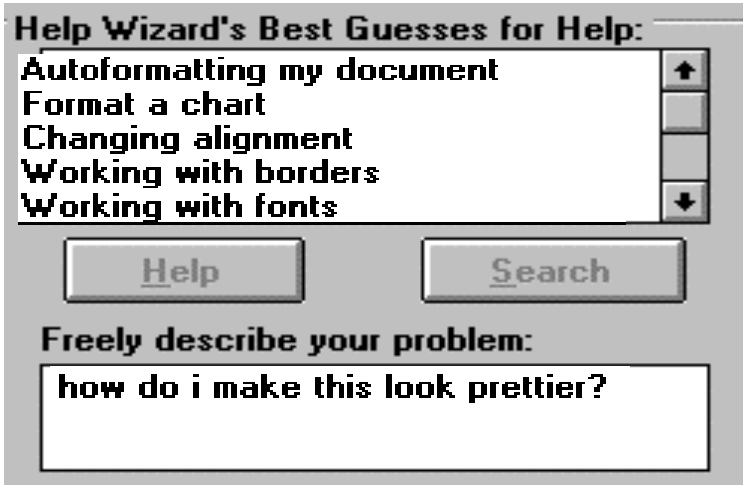
- **Inferring user's goals under uncertainty**



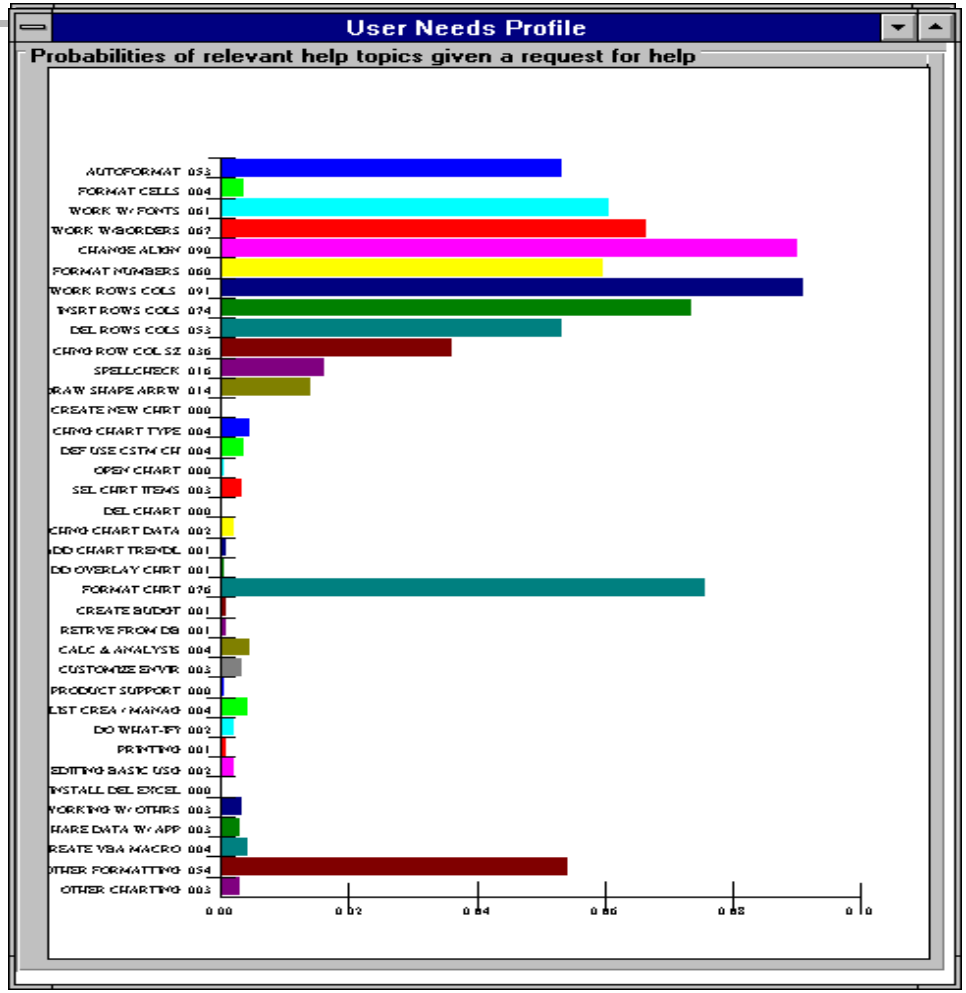
- User query
- User activity
- User profile
- Data structures

Example: *Lumiere*
<Inference from words and actions>

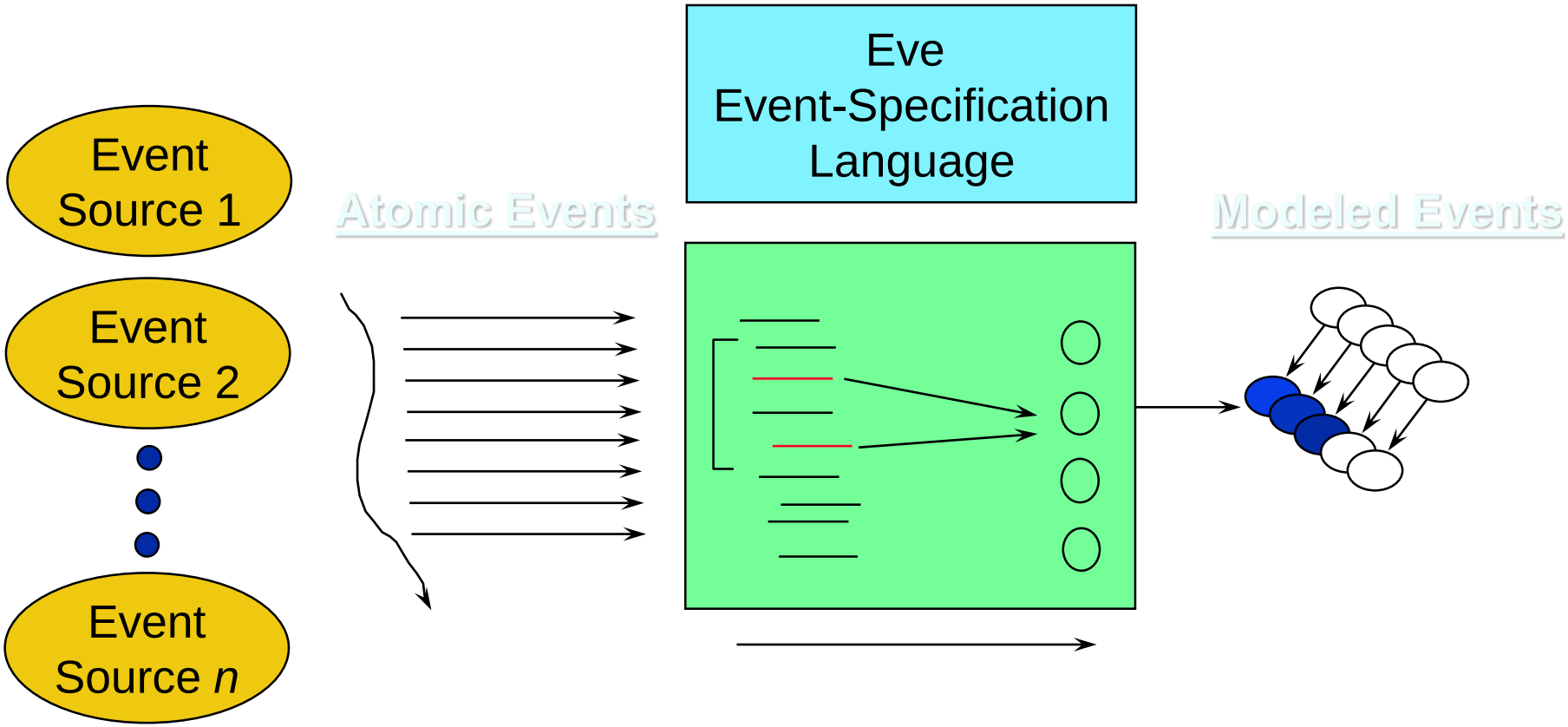
- **User's query**



- **Sensed actions**



Example: *Lumiere*
<Eve Event System>



Example: *Web Queries*

user = A1D6F19DB06BD694 ***date = 970916*** ***excite log***

150052	lion
152004	lions
152036	lions lion
152219	lion facts
[illegible]	roaring
153848	lions roaring
160232	africa lion
160642	lions, tigers, leopards and cheetahs
161042	lions, tigers, leopards and cheetahs cats
161144	wild cats of africa
161414	africa cat
161602	africa lions
161308	africa wild cats
	mane
161840	lion
161858	lion lions
163041	lion facts
163919	picher of lions
164040	lion picher
165002	lion pictures
165100	pictures of lions
165211	pictures of big cats
165311	lion photos
170013	video in lion
172131	pictureof a lioness
172207	picture of a lioness
172241	lion pictures
172334	lion pictures cat
172443	lions
172450	lions

Excite Search Results: roaring - Microsoft Internet Explorer provided by ITG

File Edit View Go Favorites Help

Back Forward Stop Refresh Home Search Favorites History Channels Fullscreen Mail Print Edit

Address ook=default&search=roaring&tsug=0&csug=10&Search=Search+Again&collection=web&clk=b-b-F1&showSummary=true&sorig=rpage&trace=b Links

excite EXCITE SEARCH SEARCH MY CHANNEL

roaring Search Again Help

Select words to add to your search..

- twenties
- forties
- fork
- basalt
- aspen
- aspenonline
- carbondale
- glenwood
- klux
- doldrums

Print coupons right from your own computer.

www.coolsavings.com

Click Here

Top **10** matches. *[24867 hits. About Your Results]*

Show Titles only List by Web site

AMAZON.COM
FIND Related Books

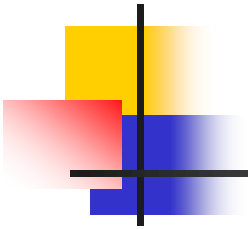
66% **Roaring Fork High School**
URL: http://www.carbondale.com/schools/rfhs/index.htm
Summary: "A school is a place to be educated about the importance of an education . . . ". In a small Colorado town lies a small Colorado high school. Wait A Second, Where Am I ?!?!. Roaring Fork Newsletter - Last updated ??/??/??. Students of the month - Last updated January 19, 1996.
More Like This: Click here to perform a search for documents like this one.

64% **Visit Morrison's Cove Home Retirement Living**
URL: http://www.mchome.org/mchvisit.htm
Summary: At the Traffic Light in Martinsburg turn right onto Market Street (Route 866). At the Traffic Light in Martinsburg turn right onto Market Street (Route 866).
More Like This: Click here to perform a search for documents like this one.

64% **ROARING - Ring Of Australian RoleplayING**

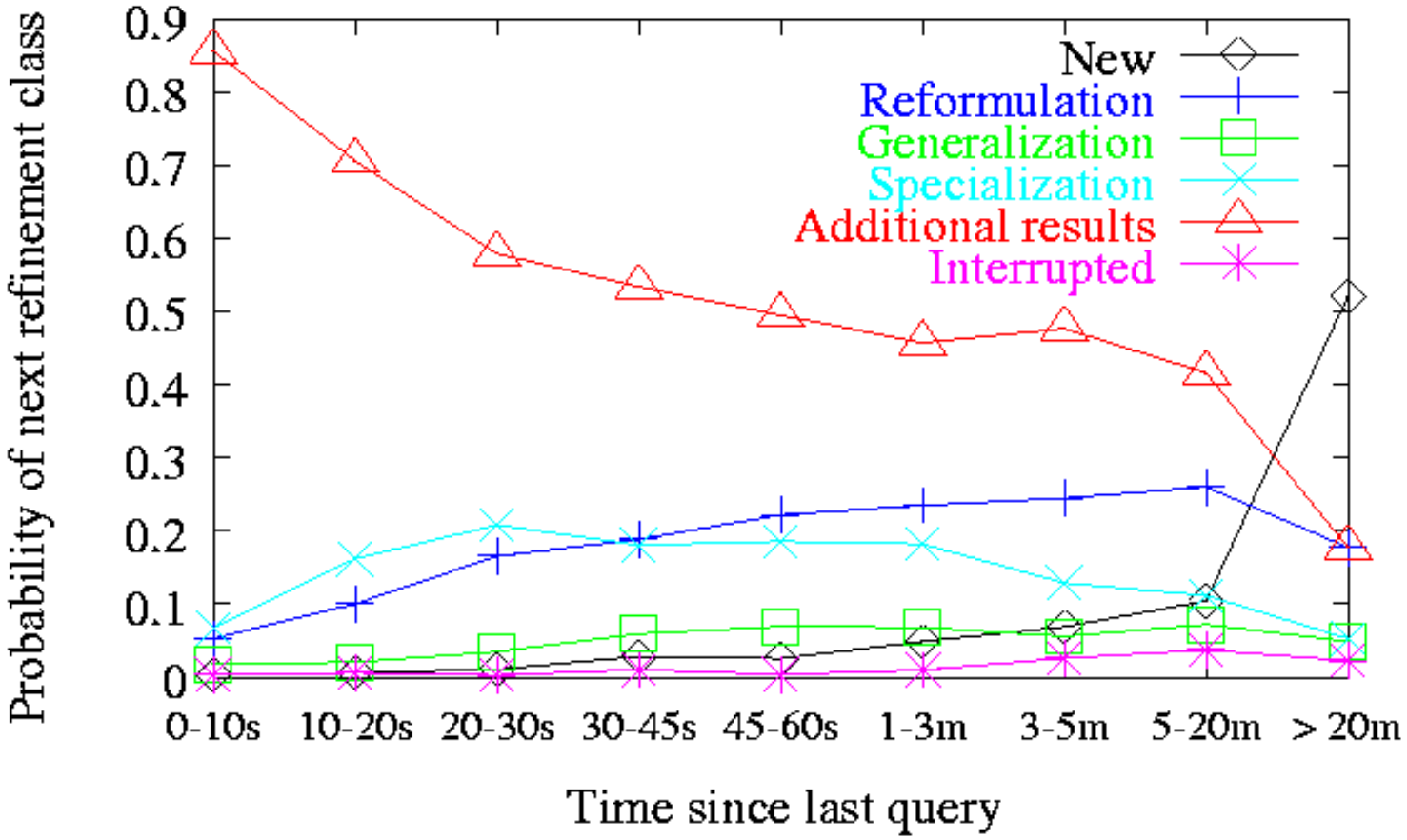
Done Internet zone

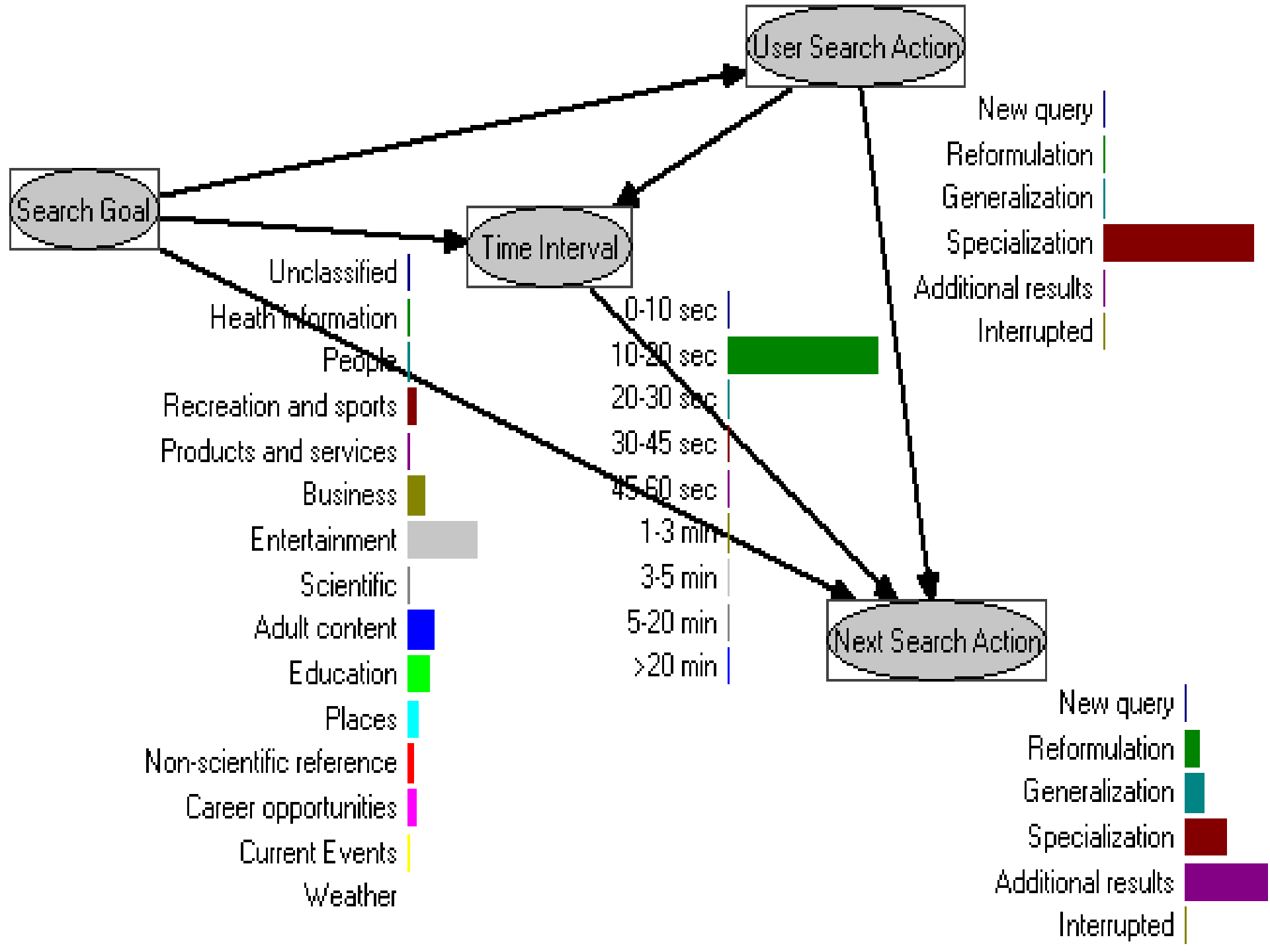
Query Dynamics & User Goals <Lau & Horvitz, UM'99>

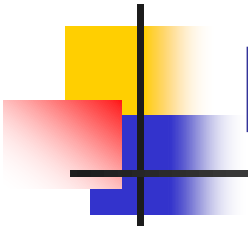


- Queries are not independent
- Consider:
  - Search goals (e.g., current events, weather)
  - Refinement actions (e.g., specialize, new)
  - Temporal dynamics
- Bayes net to predict next action, or next search goal
- Hand-tagged sample of Excite log

Temporal dynamics results

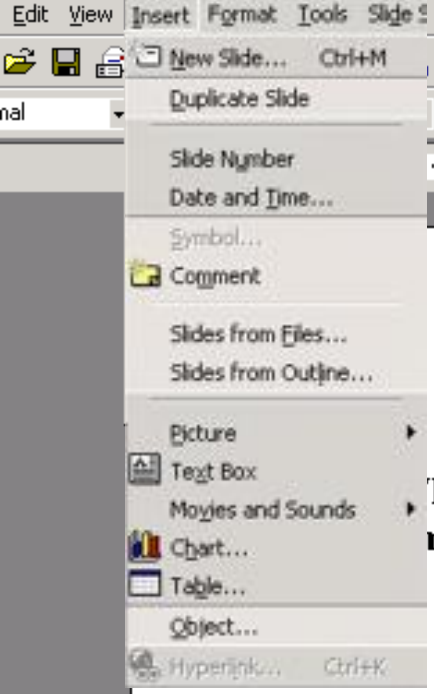



User Search Action
Search Goal
Time Interval
Next Search Action
New query
Reformulation
Generalization
Specialization
Additional results
Interrupted
Unclassified
Health Information
People
Recreation and sports
Products and services
Business
Entertainment
Scientific
Adult content
Education
Places
Non-scientific reference
Career opportunities
Current Events
Weather
0-10 sec
10-20 sec
20-30 sec
30-45 sec
45-60 sec
1-3 min
3-5 min
5-20 min
>20 min
New query
Reformulation
Generalization
Specialization
Additional results
Interrupted



Real-World Examples

- Implicit storage of history of interaction
  - Caching
  - History
  - Auto Completion
  - Dynamic Menus
- Explicit storage
  - Favorites
  - MySearch, iLOR
- Recommendations
- MyBlah …



hen do you want to go "where everybody knows your name[1]" (and nailing address, and preferences, and last 50 web pages visited)?: A framework for personalization

Susan T. Dumais
Adaptive Systems and Interaction Group
Microsoft Research
Redmond, WA 98033 USA
sdumais@microsoft.com

Abstract

Over the years, I have been involved in developing several information personalization prototypes (e.g., content-based filtering, collaborative filtering, implicit queries and visualization). Sometimes personalization works pretty well, and other times it never catches on or participation falls off after an initial spurt of interest. I think it is instructive to reflect on an even wider range of personalization efforts to begin to develop a framework (or at least articulate a set of considerations) for better understanding and developing personalized digital library services.

As a working definition, I consider a personalized system to be one that responds differently to two people who issue the same proximal request, depending on their pervious history of interaction with the system. Examples of such systems include content-based and collaborative filtering as well as caching, history lists, dynamic menus, automatic completion, customized content selection, and implicit queries. These systems vary in many respects – e.g., the characteristics that are used for personalization (e.g.,

Amazon.com: Buying Info: Deception Pass - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss Messenger

Address http://www.amazon.com/exec/obidos/ASIN/0345400690/qid=992301990/sr=1-2/ref=sc_b_2/104-2127090-2446302 Go Links

Share your thoughts

write a review

e-mail a friend about this item

Mass Market Paperback Reprint edition (July 1998)
Ballantine Books; ISBN: 0345400690 ; Dimensions (in inches): 0.83 x 6.87 x 4.20

Other Editions: Audio Cassette

Amazon.com Sales Rank: 57,009

Average Customer Rating: ★★★★½ Based on 8 reviews. Write a review.
Rate this item to get personal recommendations.

RECENTLY VIEWED ITEMS

Pattern Classification by Richard O. Duda, et al

Undercurrents by Ridley Pearson

See more in The Page You Made

Featured Item:

An Introduction to Support Vector Machines
by Nello Cristianini, John Shawe-Taylor

Visit the DVD Store

Customers who bought this book also bought:

- *The Million-Dollar Tattoo (Thomas Black Series , No 9)* by Earl W. Emerson
- *The Vanishing Smile* by Earl W. Emerson
- *Poverty Bay* by Earl W. Emerson
- *Catfish Cafe* by Earl W. Emerson

Explore similar items

Editorial Reviews

From Kirkus Reviews

What dire secret could make Lainie Smith, Seattle's well- heeled answer to Mother Teresa, vulnerable to blackmail? Whatever it is, it's something she's been paying $2,000 a week to keep quiet--and something she doesn't want to share with her lawyer, Kathy Birchfield, or Kathy's husband, private eye Thomas Black (The Million-Dollar Tattoo, 1996, etc.). Thomas doesn't insist on knowing Lainie's secret, but as he gets deeper into the case- -trailing the two men who pick up the latest two grand, searching the lair he's tracked one of them to, dispensing his trademark similes (one craven suspect has ``an alibi prepared like a frozen dinner in the freezer'')--he can hardly help finding out what it is. And it's dynamite, something it's no wonder Lainie didn't want him to know: Seventeen years ago she was present with recently executed Charlie Groth when four

Amazon.com Books Search Results - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss Messenger

Address http://www.amazon.com/exec/obidos/search-handle-form/104-2127090-2446302 Go Links

amazon.com.

VIEW CART | WISH LIST | YOUR ACCOUNT | HELP

TODAY'S FEATURED STORES: BOOKS | ELECTRONICS | CELL PHONES & SERVICE | BABY | OUTDOOR LIVING

WELCOME | STORE DIRECTORY | BOOKS

SEARCH | BROWSE SUBJECTS | BESTSELLERS | NEW & FUTURE RELEASES | BARGAIN BOOK OUTLET | E-BOOKS | RARE & USED

Search: Books earl emerson GO! Browse: Choose a subject: GO!

Most popular matches for "earl emerson"

- **The Dead Horse Paint Company** -- by Earl W. Emerson; Mass Market Paperback
  Our Price: $5.39 -- Or buy used from $1.35
- **Undercurrents** -- by Ridley Pearson; Mass Market Paperback
  Our Price: $6.29 -- Or buy used from $4.40
- **First Avenue** -- by Lowen Clausen; Mass Market Paperback
  Our Price: $6.29 -- Or buy used from $1.50

All 38 matches for "earl emerson"

Sort by: Featured Items GO!

1.
The Million-Dollar Tattoo (Thomas Black Series , No 9)
by Earl W. Emerson (Mass Market Paperback - November 1997)
Average Customer Review: ★★★★★
Usually ships in 24 hours

List Price: ~~$5.99~~
Our Price: $5.39
You Save: $0.60 (10%)

Add to cart
Or buy used from $1.75

2.
Deception Pass
by Earl W. Emerson (Mass Market Paperback - July 1998)

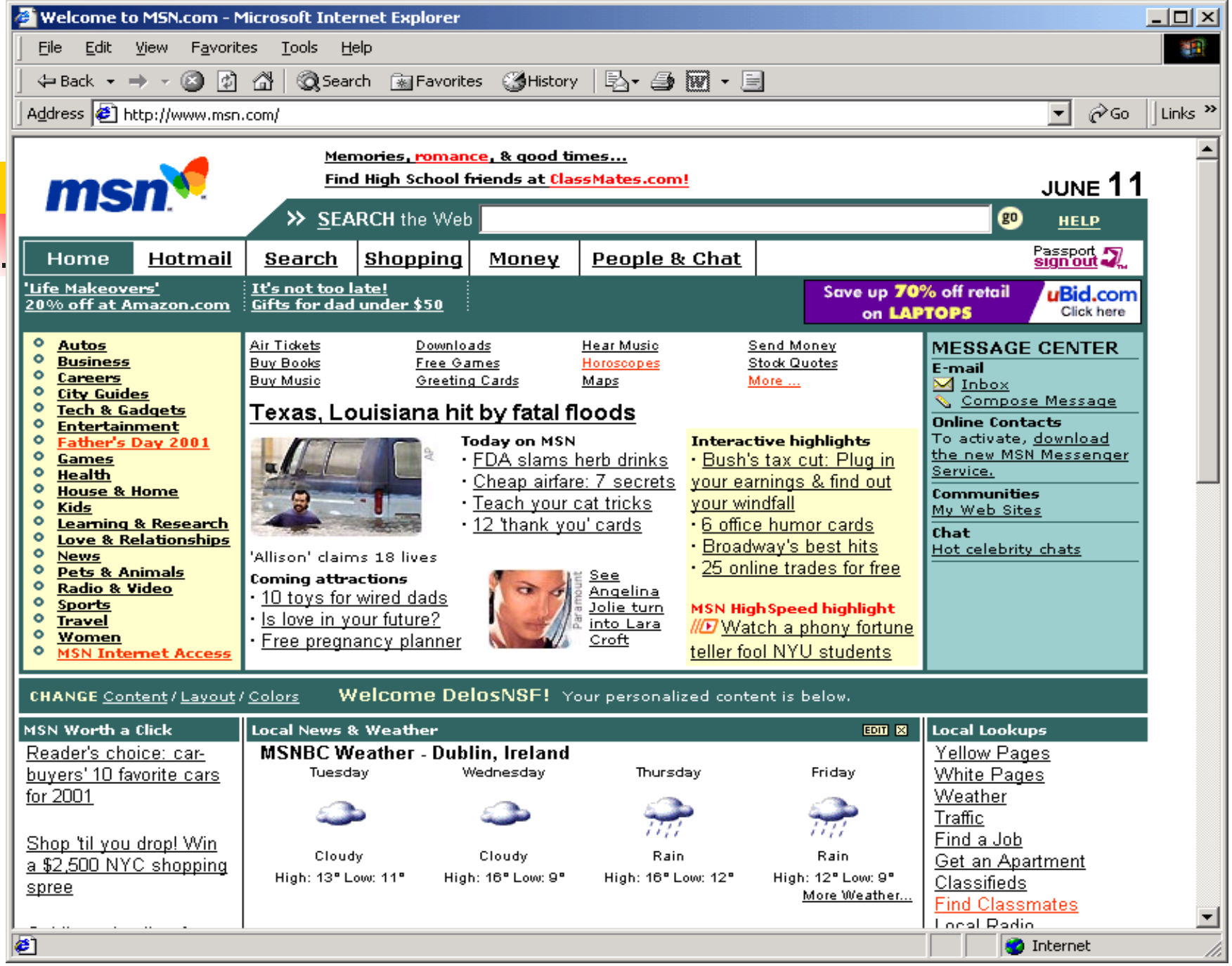
Need the perfect gift?

Send an Amazon.com gift certificate

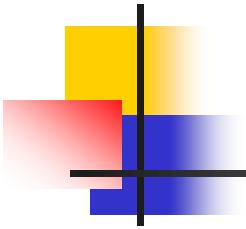
Promote your product here

Listmania!

Seattle Mysteries: A list by captain1701d, Bookseller
(7 item list)

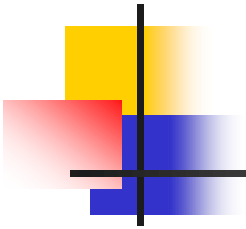

Welcome to MSN.com - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://www.msn.com/
Go Links
msn
Memories, romance, & good times...
Find High School friends at ClassMates.com!
JUNE 11
SEARCH the Web
go
HELP
Home Hotmail Search Shopping Money People & Chat
Passport sign out
'Life Makeovers' 20% off at Amazon.com
It's not too late! Gifts for dad under $50
Save up 70% off retail on LAPTOPS
uBid.com Click here
Autos
Business
Careers
City Guides
Tech & Gadgets
Entertainment
Father's Day 2001
Games
Health
House & Home
Kids
Learning & Research
Love & Relationships
News
Pets & Animals
Radio & Video
Sports
Travel
Women
MSN Internet Access
Air Tickets
Buy Books
Buy Music
Downloads
Free Games
Greeting Cards
Hear Music
Horoscopes
Maps
Send Money
Stock Quotes
More ...
Texas, Louisiana hit by fatal floods
'Allison' claims 18 lives
Today on MSN
· FDA slams herb drinks
· Cheap airfare: 7 secrets
· Teach your cat tricks
· 12 'thank you' cards
Coming attractions
· 10 toys for wired dads
· Is love in your future?
· Free pregnancy planner
See Angelina Jolie turn into Lara Croft
Interactive highlights
· Bush's tax cut: Plug in your earnings & find out your windfall
· 6 office humor cards
· Broadway's best hits
· 25 online trades for free
MSN HighSpeed highlight
Watch a phony fortune teller fool NYU students
MESSAGE CENTER
E-mail
Inbox
Compose Message
Online Contacts
To activate, download the new MSN Messenger Service.
Communities
My Web Sites
Chat
Hot celebrity chats
CHANGE Content / Layout / Colors
Welcome DelosNSF! Your personalized content is below.
MSN Worth a Click
Reader's choice: car-buyers' 10 favorite cars for 2001
Shop 'til you drop! Win a $2,500 NYC shopping spree
Local News & Weather
EDIT
MSNBC Weather - Dublin, Ireland
Tuesday Cloudy High: 13° Low: 11°
Wednesday Cloudy High: 16° Low: 9°
Thursday Rain High: 16° Low: 12°
Friday Rain High: 12° Low: 9°
More Weather...
Local Lookups
Yellow Pages
White Pages
Weather
Traffic
Find a Job
Get an Apartment
Classifieds
Find Classmates
Internet

Welcome to MSN.com - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Forward Stop Refresh Home Search Favorites History Mail Size Print Edit
Address http://www.msn.com/
Go Links
CHANGE Content / Layout / Colors
Welcome DelosNSF! Your personalized content is below.
MSN Worth a Click
Get 2 months of WebTV service, FREE!
Meet NASCAR star Tony Stewart & win $1,800
New! Expedia Bargain Fares: save up to 60%
Summer Travel
Hike & camp trip deals
Travel guides
Visit Europe, buy a car
'Foot & mouth' risk?
Road tunes
Travel with pets
US cities: what to see
Driving directions
Links & Resources
From MSN
Calendar & reminders
Fast Internet Access
Get MSN Companion
MSN on your cell phone
MSN Messenger Service
Try MSN Explorer!
From Microsoft
microsoft.com
Small Business Services
Windows Update
Office Tools
Pocket PC
Xbox.com
MSN Worldwide
Local News & Weather
MSNBC Weather - Dublin, Ireland
Sunday Monday Tuesday Wednesday
Mostly Cloudy High: 19° Low: 12°
Windy High: 20° Low: 8°
Partly Cloudy High: 20° Low: 11°
Partly Cloudy High: 19° Low: 6°
More Weather...
Computing & Web
Wired
Spam Might Taste OK
A Brief History of SPAM, and Spam
Jeffords Bolts, Tech Yawns
News
New York Times
Congress Passes Tax Cut, With Rebates This Summer
War in Sierra Leone Loosens Its Grip on Child Soldiers
Russia Sees Payoff in Storing Nuclear Waste From Around the World
Travel
Expedia.com Travel Deals
Cruise deals: We deliver the best picks to your inbox
Be unique: Reserve a room in a French Bungalow
Travel deals: Check out the latest specials and save
Personal Finance
MoneyCentral Stock Quotes
Datek Online Ameritrade
Local Lookups
Yellow Pages
White Pages
Weather
Traffic
Find a Job
Get an Apartment
Classifieds
Find Classmates
Local Radio
Event Tickets
Sports Scores
TV Listings
More...
MSN Spotlight
Ultimate home improvement how-to guide
Easy ideas for every budget
MSN Shopping
Hot new products:
Bose Wave/PC
NBA Playoff Gear
Office XP
More...
Search eShop for:
Search
Shop for Summer:
Internet



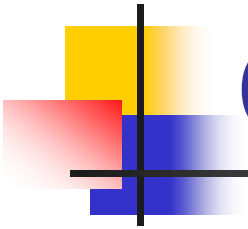
Personalization Success

- Effectively Assay and Accommodate:
  - Easy to specify relevant information
    - Explicitly: profile changes slowly
    - Implicitly: capture automatically, esp short time
  - We know what to do about it
    - Algorithmic and application levels
    - And, the user can see the benefit
    - And, there are few big failures

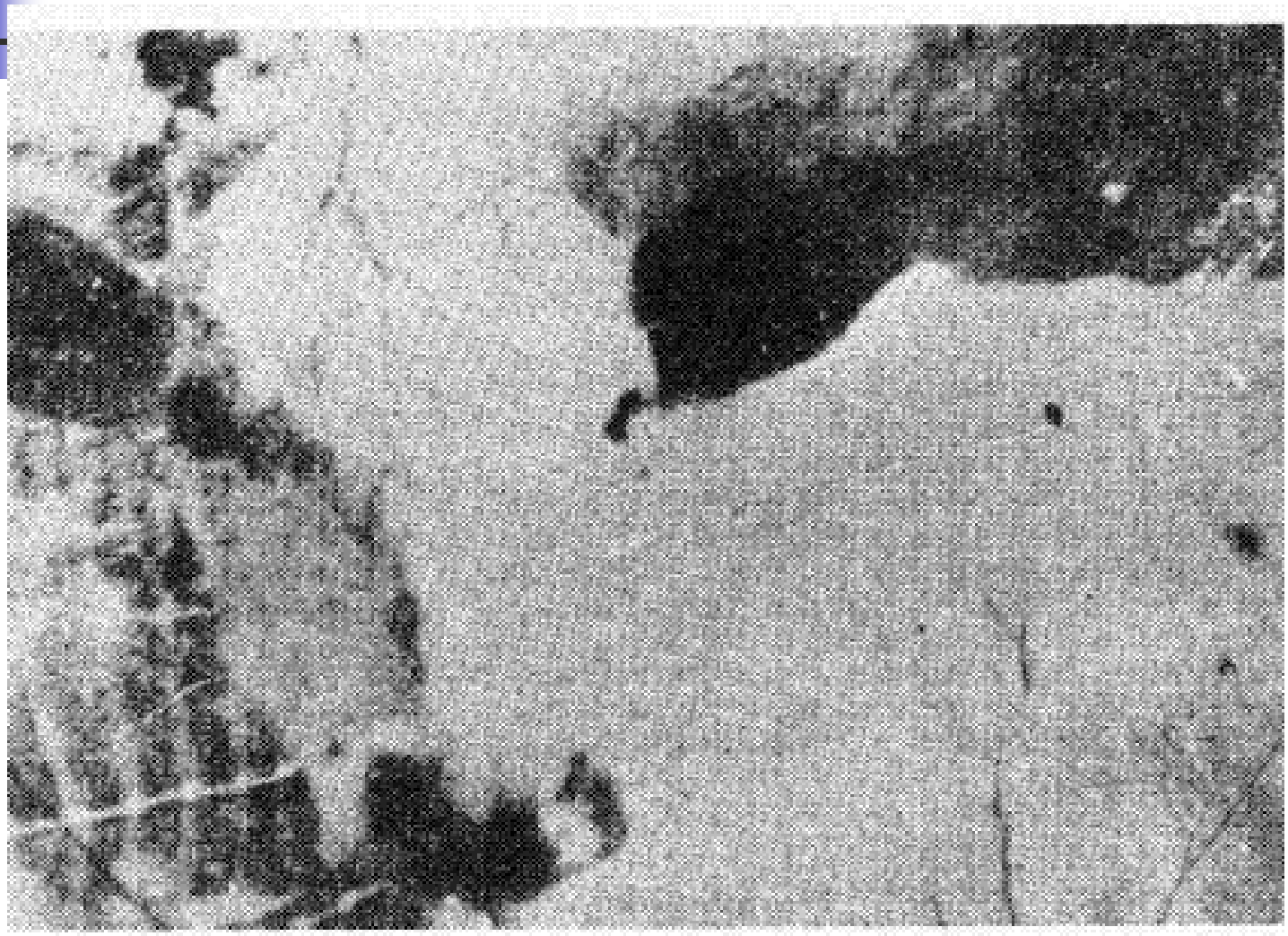


Personalization Opportunities

- Geo-coding
- Query history
- Query plus usage context
- Keeping found things found



Open Issues

- Evaluation … difficult for personalized systems
  - Components, easier
  - End-to-end applications, harder
    - Questionnaires
    - Pre-Post assessment
    - Algorithmic issues in situ
- Privacy, security …

The End …