

Stuff I've Seen:

Retrospective and Prospective

Susan Dumais



SIGIR Desktop Search Workshop

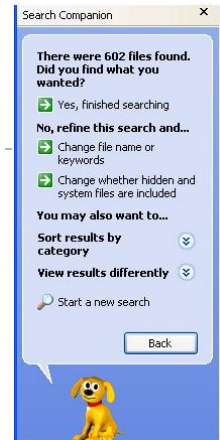
Overview

- ▶ What is Stuff I've Seen (SIS)?
 - ▶ SIS @ SIGIR 2003
 - ▶ Key findings
- ▶ What has changed?
- ▶ What is next?



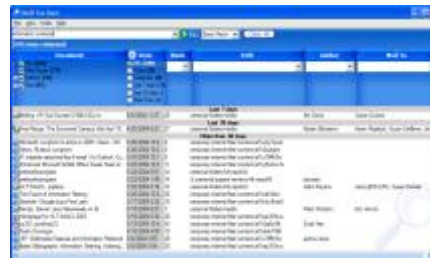
Stuff I've Seen: @ SIGIR 2003

- ▶ SIGIR 2003  
- ▶ Desktop Search in 2003
- ▶ Stuff I've Seen
 - ▶ Developed, deployed and evaluated a new system (algorithms and interface) for supporting re-finding
 - ▶ Not a typical SIGIR paper ...
 - ▶ **R1:** *The considered problem is interesting and relevant. A system like SIS would really facilitate every day's life. The collected data and the arguments drawn from it suggest the effectiveness of SIS. However, as the scientific value of the study really lies on the experiments, **somewhat more comprehensive empirical study** would have been appreciated.*
 - ▶ **R3:** *There was no reflection of the evaluation methods used. Some of the chosen criteria (variables) to evaluate the system were not motivated. The usage statistics was relevant point of departure, but e.g. **why the query characteristics or comparison between rank vs. time options?** The questions in the questionnaire were more focused evaluation measures.*
 - ▶ Yet, second most-cited paper from SIGIR 2003
 - ▶ Also, influential in Windows Search today



Stuff I've Seen: Design Motivations

- ▶ Fast, flexible search over *stuff you've seen*
 - ▶ Heterogeneous content: files, email, calendar, web, rss, IM, ...
 - ▶ Index: full-content plus metadata
 - ▶ Interface: highly interactive rich list-view
 - ▶ Sorting, filtering, scrolling
 - ▶ Grouping and previews
 - ▶ Rich actions on results (open, open folder, drag-and-drop)
 - ▶ New interface possibilities since it's your content ... re-finding
- ▶ Stuff I've Seen Demo



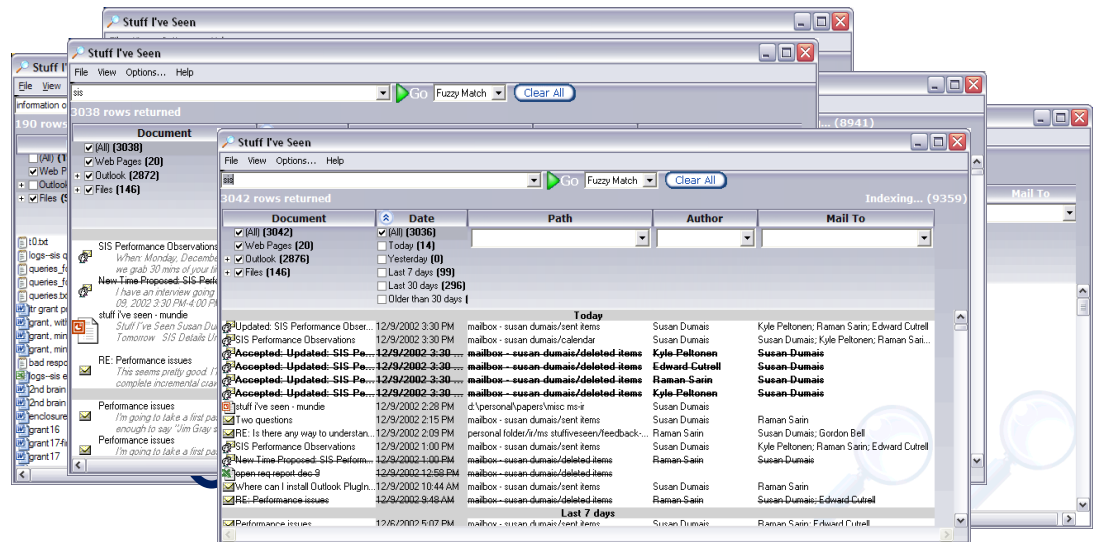
Stuff I've Seen: Evaluation

- ▶ Evaluation ... multiple methods
 - ▶ Deployed the system for 6+ weeks
 - ▶ Log data [mostly interaction data]
 - ▶ Questionnaires [pre and post]
 - ▶ Field experiments [3 variables, 6 alternative systems]

Top vs. Side

Preview vs. Not

Sort By Date vs. Rank

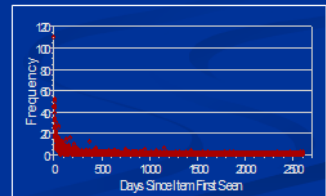


- ▶ Also: Lab studies, Interviews, etc.

Stuff I've Seen: Results

- ▶ Personal store characteristics
 - ▶ 5–500k items
- ▶ Query characteristics
 - ▶ Very short queries (1.6 words)
 - ▶ Few advanced operators in the query box (7%); many in UI (48%)
 - ▶ Filters (type, date); modify query; re-sort results
 - ▶ People are important – 25% queries involve names/aliases
- ▶ Items opened characteristics
 - ▶ Type: Email (76%), Web pages (14%), Files (10%)
 - ▶ Age: Today (5%), Last week (21%), Last month (47%)
 - ▶ 53% > one month
 - ▶ Need to support episodic access to memory

$$\text{Log(Freq)} = -0.68 * \text{log(DaysSinceSeen)} + 2.02$$



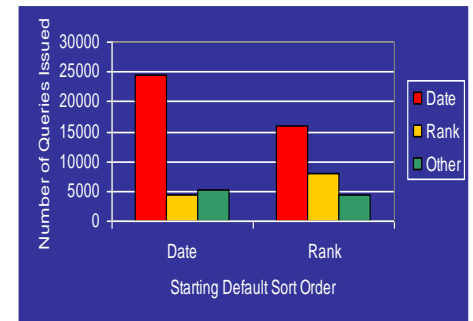
Stuff I've Seen: Results (cont'd)

▶ Interface experiments

- ▶ Small effects of Top vs. Side, or Preview vs. No Previews
- ▶ Large effect of sort order (Date vs. Rank)
 - ▶ **Date** by far the most common sort order, even for people who had best-match Rank as the default
 - ▶ Few searches for “best” matching object
 - ▶ Many other criteria – e.g., time, people

▶ Abstraction important in human memory

- ▶ “Useful date”
 - ▶ Appointment, when it happens
 - ▶ Picture, when it was taken
 - ▶ Web, when it was seen
- ▶ “People” in attribute (To, From, Author, Artist) vs. contains
- ▶ “Picture” whether jpg, tif, png, gif, pdf, ...



Example searches

Looking for: *recent email from Fedor that contained a link to his new demo*

Initiated from: Start menu

Query: from:Fedor

Looking for: *the pdf of a SIGIR paper on context and ranking (not sure it used those words) that someone (don't remember who) sent me a month ago*

Initiated from: Outlook

Query: SIGIR

Looking for: *meeting invite for the last intern handoff*

Initiated from: Start menu

Query: intern handoff kind:appointment

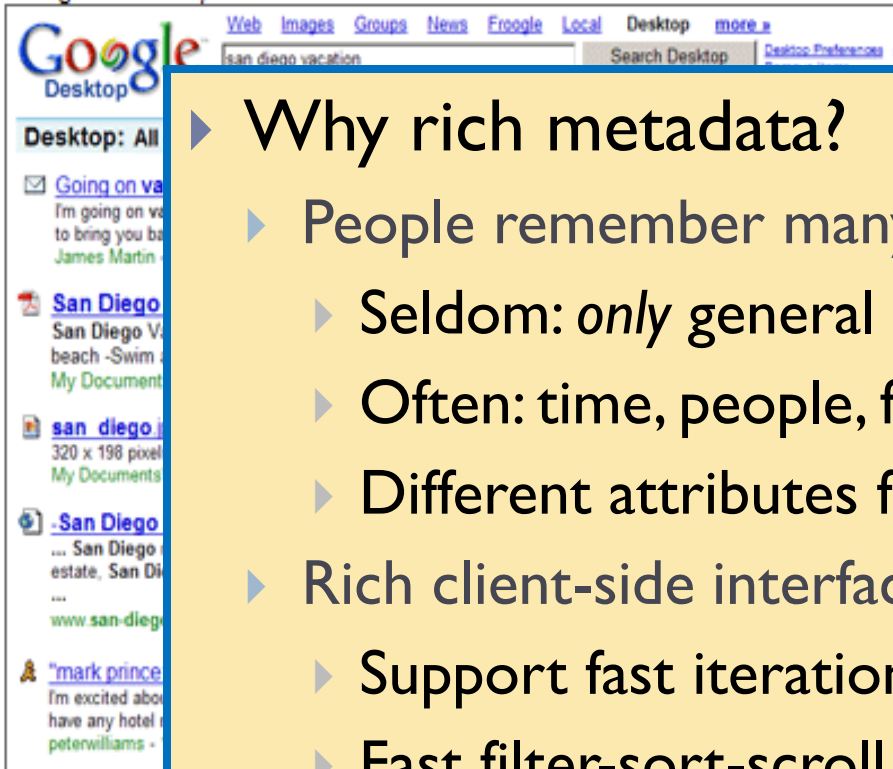
Looking for: *C# program I wrote a long time ago*

Initiated from: Explorer pane

Query: QCluster*.*

Stuff I've Seen: Ranked list vs. Metadata (for personal content)

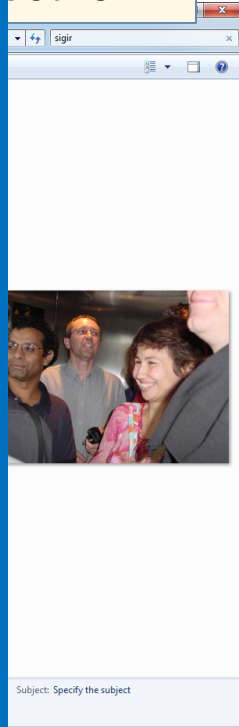
Google Desktop Search results



- ▶ Why rich metadata?
 - ▶ People remember many attributes in re-finding
 - ▶ Seldom: *only* general overall topic
 - ▶ Often: time, people, file type, etc.
 - ▶ Different attributes for different tasks
 - ▶ Rich client-side interface
 - ▶ Support fast iteration and refinement
 - ▶ Fast filter-sort-scroll vs. next-next-next
 - ▶ “Fluidity of interactions”
 - ▶ Desktop search != Web search

Seen

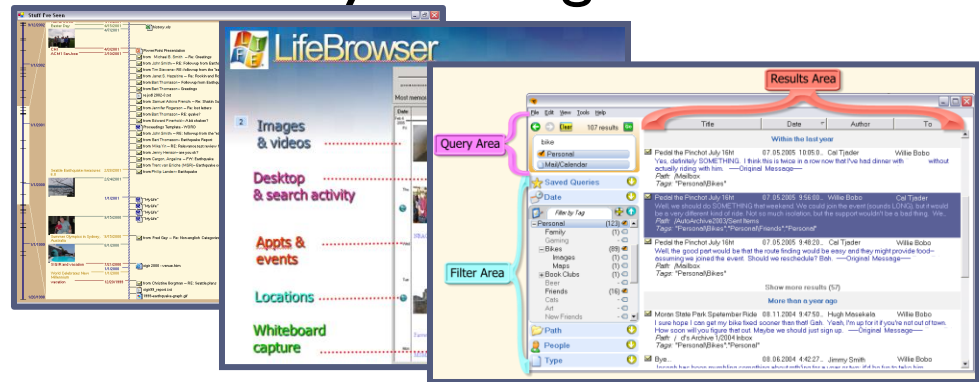
Search



Beyond Stuff I've Seen

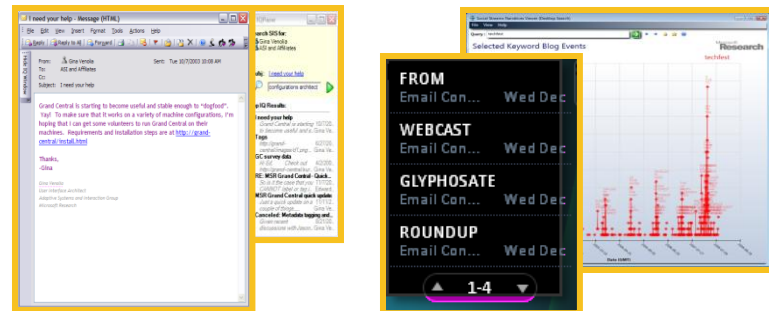
▶ Better support for human memory & integration with browsing

- ▶ Memory Landmarks
- ▶ LifeBrowser
- ▶ Phlat



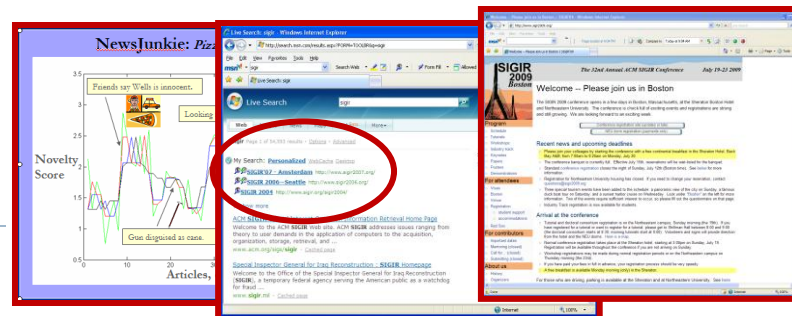
▶ Beyond search

- ▶ Proactive retrieval
 - ▶ Stuff I Should See (IQ)
 - ▶ Temporal Gadget



▶ Using desktop index as a rich “user model”

- ▶ News Junkie
- ▶ PSearch
- ▶ Diffie



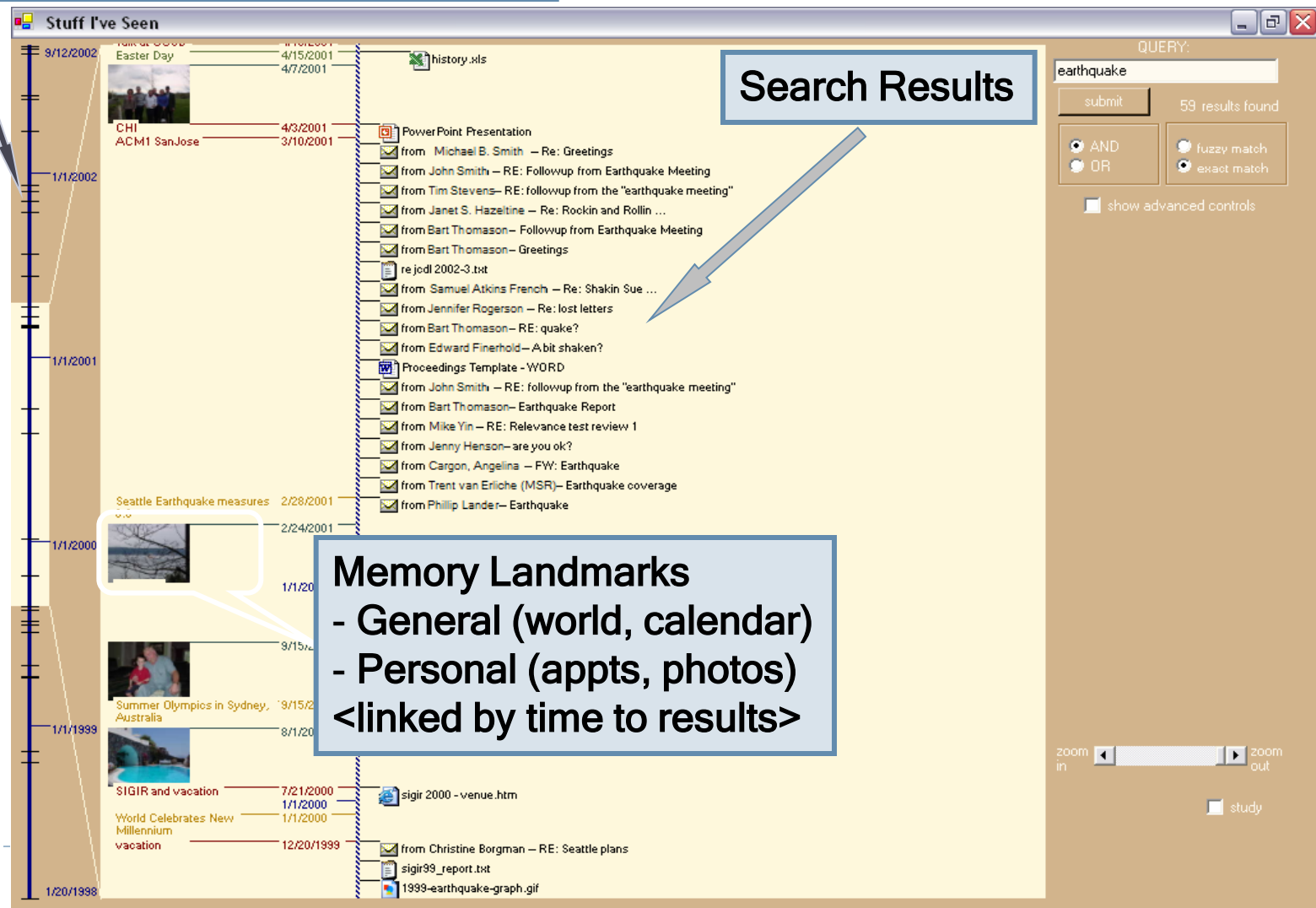
Memory Landmarks

- ▶ Importance of episodes in human memory
 - ▶ Memory organized into episodes (Tulving, 1983)
 - ▶ People-specific events as anchors (Smith et al., 1978)
 - ▶ Time of events often recalled relative to other events, historical or autobiographical (Huttenlocher & Prohaska, 1997)
- ▶ Identify and use landmarks facilitate search and information management
 - ▶ Timeline interface, augmented w/ landmarks
 - ▶ Bayesian models to identify memorable events
- ▶ Extensions beyond search, Life Browser



Memory Landmarks

Distribution of Results Over Time



key dependencies (from learned graphical model)



LifeBrowser

Images
& videos

Desktop
& search activity




Appts &
events

Locations

Whiteboard
capture

LifeBrowser interface showing a timeline of events and search results.

Query: Smartphlow

Date	Events	Results by Time
Feb 4 2005 Fri	 <p>privacy talk 2005.ppt priv_out.ppt</p>	<p>Micronews 5 to 1</p> <ul style="list-style-type: none"> JamBayes Discussion and Demo JamBayes Discussion and Demo
	<p>Micronews 5 to 1.doc</p>	
Thu	 <p>lake-central-visualize.ppt C4 AUV.doc</p> <p>NRAC Italy Meeting</p>	<ul style="list-style-type: none"> Inrix folks 6 month check up... :> Video... smartphlow.mpg
Wed	<p>isometric_volumes_study.p LC_visual.ppt techfest.doc</p>	<ul style="list-style-type: none"> Phlat c# filter & query controls Phlat c# filter & query controls Smartphlow/Caravel meeting recap
Tue	 <p>Gates Invitation.doc</p>	<ul style="list-style-type: none"> Smartphlow cannot start? Smartphlow cannot start? 6 month check up... :> smartphlow_exclamation.bmp
	<p>20040525_C10.xls</p> <p>Farewell Party for David Weise</p>	
Mon	<p>MSMLS coordination</p>	<ul style="list-style-type: none"> Asi dev status week of 1/17/2005 Questions about SmartPhlow

LifeBrowser – Selective Memory

Events

TAB Meeting - October 21-22, 2002
 Amir Aczel, 10/21, "Entanglement: The G
 TAB 20' - Nuria Oliver: The Challenge o
 Brainstorm speed from 802.11
 IA working group
 Lunch Eric/BryanMi RE: InfoAgent

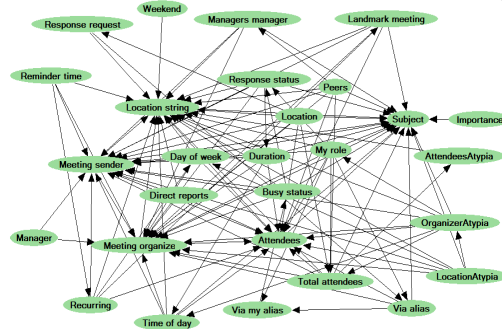
TAB Meeting - October 21-22, 2002
 Follow-up on Notification research
 UW CSE Colloquium - 10/22/02 - "Trac

Isat Irvine
 ISAT Fall Meeting 10/23-24/02 *REMIND
 IA Schema Design Group
 Amy Bruckman-- Talk about the design
 Media & Machine Learning seminar: Li D
 Ejowi Nuwere, 10/23, "Hacker Cracker: A
 Dan Ling's Direct Report Meeting
 Dan Ling's Extended Staff Meeting
 NLPRead - Basu
 Dan Ling's Group Managers Meeting - Oct

Recurrent ASI focus meetings
 Eric Horvitz x62127 "Meeting"
 Weekly get togethers, catch up
 ASI Perception & Interaction meetings
 Eric Horvitz x62127 "Recurring mtgng"
 Horvitz/Ling 1:1

Building Large-scale Ontologies by Unsup
 ,

TOCHI paper
 10/25/2002 A Machine Learning Approac
 1:1 w/ Gina
 RE: MSN Content Filtering Brownbag talk



Date	Events	Desktop Activity
Oct 21 2002		
Mon		Local Bestcom for Smart cacm-picture-fodder.ppt
Tue		cacm_final.doc
Wed	Isat Irvine ISAT Fall Meeting 10/23-24/02	prefetchbib.txt readme_priorities.txt smart ambient companion smart ambient companion overlay_UI.ppt MS177850.1v3redline1-f abstract.doc
Thu	Horvitz/Ling 1:1	junk_filter.ppt icmi2002.ppt strobe.ppt
Sat	DEF	2. ShurninEyeGaze-V[4] Organizing Against Spam
Sun		new pat proj.txt Organizing Against Spam Organizing Against Spam
Oct 28 2002		
Mon		itai.txt
Tue	David Hovel	bestcom_telephony talk.ppt
Wed	Bestcom presentation	telephone.ppt Copy of telephony.ppt

What's Changed ?

- ▶ Desktop search is prevalent
 - ▶ Ships in Windows, OS X, GDS ... and it is widely used
- ▶ E.g., Windows Search
 - ▶ LOTS of engineering – efficiency, coverage, robustness, etc.
 - ▶ Multiple entry points – start menu, explorer, applications (e.g., Outlook)
 - ▶ New features and capabilities
 - ▶ Real-time results as you type (“word-wheel”)
 - ▶ Search to launch programs (in addition to finding content)
 - ▶ Context-specific options (filters, presentation)
 - ▶ Natural language search – e.g., mail from ryen sent this week
 - ▶ Tight coupling of navigation and search
 - ▶ Federation



What's Changed ? (cont'd)

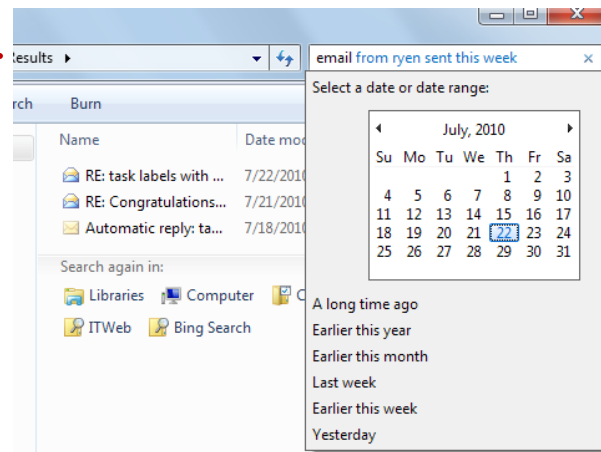
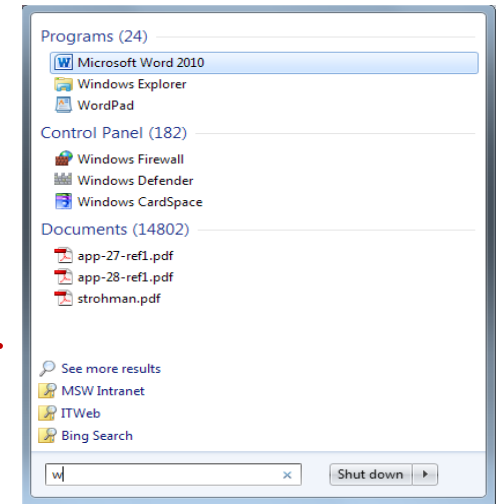
Ex: Real-time results (and search to launch programs)

Ex: Context and natural-language search

▶ E.g., Windows Search

▶ New features and capabilities

- ▶ Real-time results as you type (“word-wheel”)
- ▶ Search to launch programs (in addition to finding content)
- ▶ Context-specific options (filters, presentation)
- ▶ Natural language search – e.g., mail from ryen sent this week
- ▶ Tight coupling of navigation and search
- ▶ Federation



Ongoing Challenges

- ▶ Retrieval failures w/ desktop search
 - ▶ Vocabulary mismatch, can mitigate via metadata
 - ▶ Over specification
- ▶ Re-finding on the desktop vs. Web
 - ▶ Few navigational queries (except for commands)
 - ▶ Same query can have many intents (e.g., *from:Eric*)
- ▶ Evaluation
 - ▶ Individuals must make their own relevance judgments
 - ▶ Ranking vs. interaction
 - ▶ There is much more than a single ranking
 - ▶ Interaction – transparency, control and predictability matter
 - ▶ *In situ* vs. in simulation
 - ▶ Need to evaluate *in situ* – not enough to optimize a measure (or component) without seeing how it influences interaction

What's Next?

▶ Universal or specialized search?

- ▶ One flexible UI vs. many special purpose tools?
 - ▶ E.g., Email vs. photo vs. file search
- ▶ General entry point, w/ context-specific features
- ▶ Plus, application-specific access to same index

▶ Federation

- ▶ Multiple “desktops” [PCs, mobile, other devices]
 - ▶ Mobile especially interesting
- ▶ Desktop -> Cloud-based services (e.g., Twitter, Facebook, Mail)
 - ▶ More siloed? Where should the index live?
 - ▶ Web services vs. Web pages. What to index?
 - ▶ Personal vs. Social
 - Social aggregation – “spindex” (<http://fuse.microsoft.com/projects-spindex.html>)



Thanks!

- ▶ Questions / Comments?

- ▶ Additional info

sdumais@microsoft.com

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

