

WWW Panel: Searching Personal Content

The Person in Personal

(Supporting the Person
in Searching Personal Content)

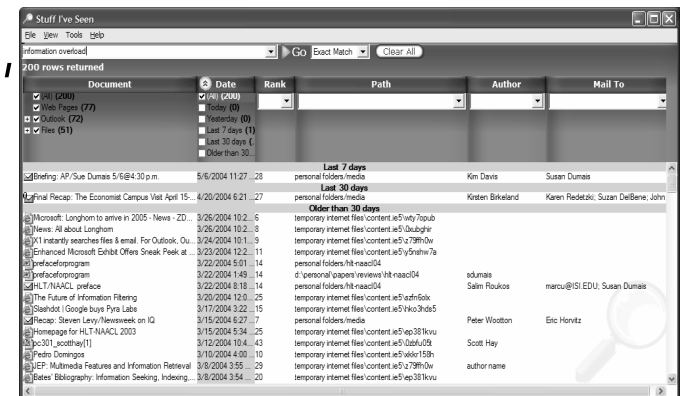
Susan Dumais

Microsoft Research

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

Stuff I've Seen (SIS)

- Personal content “stuff you've seen”
- Finding vs. Re-finding (e.g., file, web, memory)
- Unified access to *stuff you've seen*
 - Many types of info (e.g., files, email, calendar, contacts, web pages, rss, im, music, pictures)
 - Index content plus metadata (e.g., time, author, title, size, usage)
 - Rich UI possibilities, since it's your content (rich associations)



What People Do Look For?

- *“The email that I sent to Yoelle earlier this week about the format for the WWW panel.”*
- *“The pictures I took during our visit to NYC of Ben with his face covered in chocolate pudding.”*
- *“Whether the spec for user studies that we [our team] recently reviewed has been updated.”*
- *“The special price for the vacation rental at the Kona Beach Resort.”*

What People Do Look For?

- *"The email that I sent to Yoelle earlier this week about the format for the WWW panel."*
- *"The Importance of Ben C of Budding."*
 - Time, People
- *"Who are [our team] es."*
 - Also type/source, entities, events, history of interaction
- *"The Konrad at the"*
 - Supporting what the searcher remembers about the information

SIS Interface

Stuff I've Seen

File View Tools Help

information overload | Go Exact Match Clear All

200 rows returned

Document	Date	Rank	Path	Author	Mail to
Last 7 days					
☑ Briefing: AP/Sue Dumais 5/6@4:30 p.m.	5/6/2004 11:27 ...28		personal folders/media	Kim Davis	Susan Dumais
Last 30 days					
☑ Final Recap: The Economist Campus Visit April 15...	4/20/2004 6:21 ...27		personal folders/media	Kirsten Birkeland	Karen Redetzki; Suzan DelBene; John
Older than 30 days					
☑ Microsoft: Longhom to arrive in 2005 - News - ZD...	3/26/2004 10:2... 6		temporary internet files\content.ie5\wty7opub		
☑ News: All about Longhom	3/26/2004 10:2... 8		temporary internet files\content.ie5\0xubghir		
☑ Instantly searches files & email. For Outlook, O...	3/24/2004 10:1... 9		temporary internet files\content.ie5\z79ffh0w		
☑ Enhanced Microsoft Exhibit Offers Sneak Peek at...	3/23/2004 12:2... 11		temporary internet files\content.ie5\y5nshw7a		
☑ prefaceforprogram	3/22/2004 5:01 ...14		personal folders/hlt-naacl04		
☑ prefaceforprogram	3/22/2004 1:49 ...14		d:\personal\papers\reviews\hlt-naacl04	sdumais	
☑ HLT/NAACL preface	3/22/2004 8:18 ...14		personal folders/hlt-naacl04	Salim Roukos	marcu@ISI.EDU; Susan Dumais
☑ The Future of Information Filtering	3/20/2004 12:0... 25		temporary internet files\content.ie5\szfn6ox		
☑ Seshdot Google buys Pyra Labs	3/17/2004 3:22 ...15		temporary internet files\content.ie5\hko3hds5		
☑ Recap: Steven Levy/Newsweek on IQ	3/15/2004 6:27 ...7		personal folders/media	Peter Wootton	Eric Horvitz
☑ Homepage for HLT-NAACL 2003	3/15/2004 5:34 ...25		temporary internet files\content.ie5\vp381kvu		
☑ p_301_scotthay[1]	3/12/2004 10:4... 43		temporary internet files\content.ie5\0zbfu05t	Scott Hay	
☑ Pedro Domingos	3/10/2004 4:00 ...1		temporary internet files\content.ie5\wkkr158h		
☑ EP: Multimedia Features and Information Retrieval	3/8/2004 3:55 ... 29		temporary internet files\content.ie5\z79ffh0w	author name	
☑ Gates' Bibliography: Information Seeking, Indexing...	3/8/2004 3:54 ... 20		temporary internet files\content.ie5\vp381kvu		

WWW 2007 - Panel

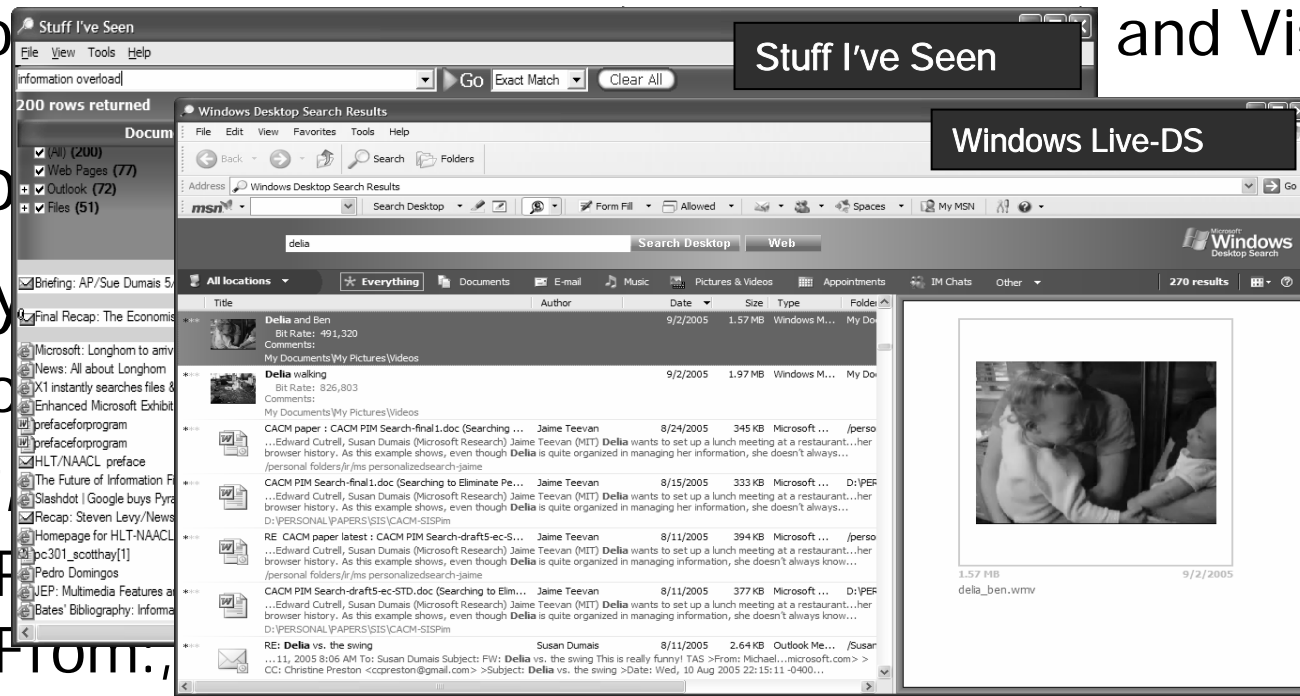
SIS Usage Experiences

- Stuff I've Seen (SIS) - Research prototype system
 - ~3000 internal Microsoft users
 - Analyzed: Free-form feedback, Questionnaires, Structured interviews, Diaries, Log analysis, UI expts, Lab expts
 - Info and Vista DS

- Personal
- Query

- Show
- But

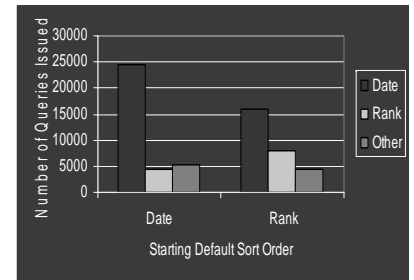
■ From...



ators)
(48%)
n),

SIS Usage, cont'd

- Importance of time (and other metadata)
 - Date by far the most common sort field, even for people who had Rank as default
 - “Best” ordering ... depends on what people remember for this search
- Human memory depends on abstractions
 - “Memorable date” is dependent on the object !
 - Appointment, when it happens
 - Picture, when it is taken
 - File, when it is changed
 - Email and Web, when it is seen
 - “Person” attribute vs. contained in text
 - To, From, Cc, Author, Artist



Ranked list vs. Metadata (for personal search)

Google Desktop Search results

The image shows a screenshot of a Windows desktop environment. In the background, a Google Desktop Search results window is visible, displaying search results for 'san diego vacation'. In the foreground, a 'Windows Desktop Search Results' window is open, showing a file named 'p212fig4.jpg' with a date of '10/10/2005' and a size of '12.5 KB'. A large, semi-transparent black box is overlaid on the search results, containing the text 'Why Rich Metadata?' and a bulleted list of points.

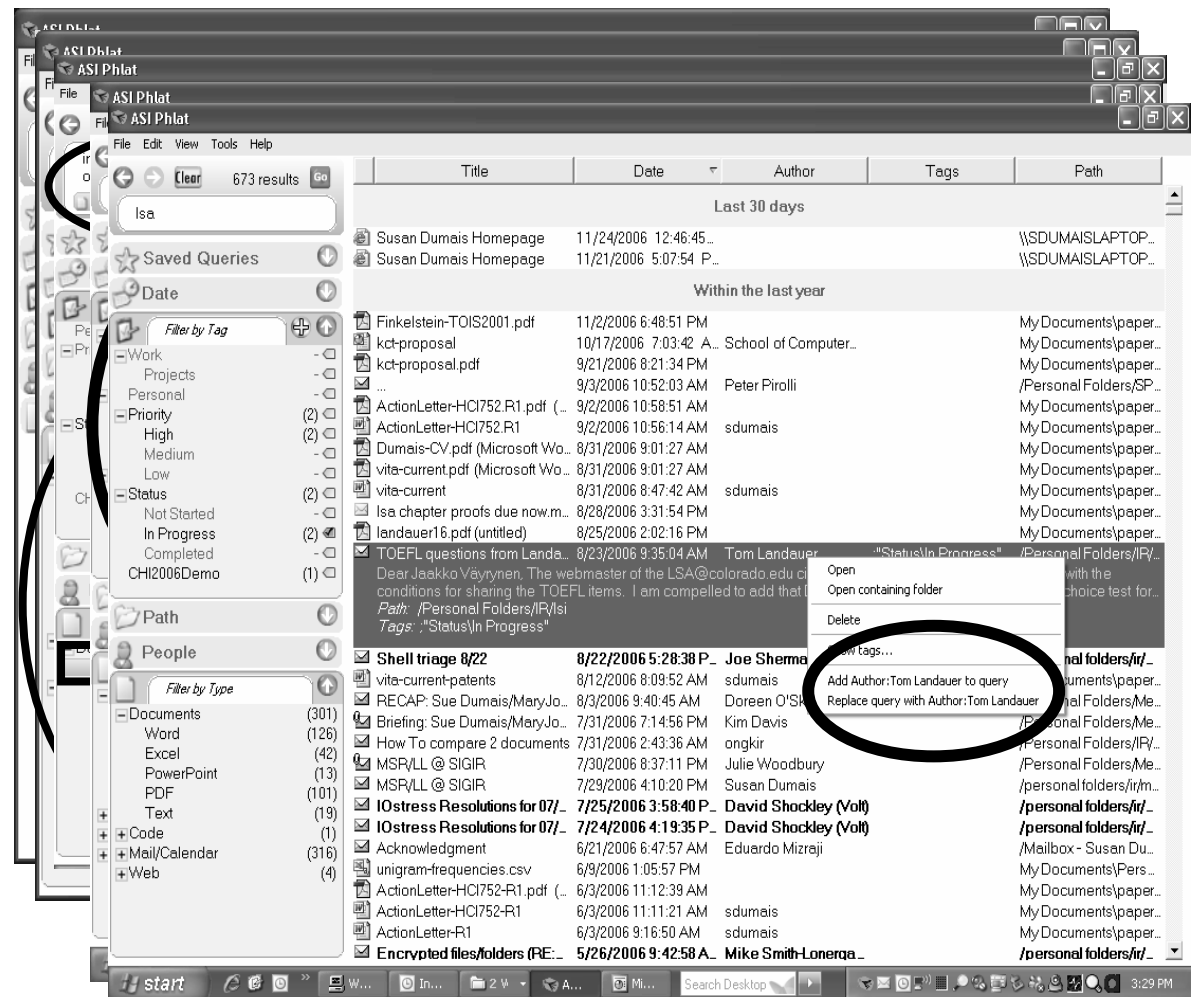
Why Rich Metadata?

- People remember many attributes in re-finding
 - Often: time, people, file type, use, etc.
 - Seldom: *only* general overall topic
- Rich client-side interface
 - Support fast iteration/refinement
 - Fast filter-sort-scroll vs. next-next-next

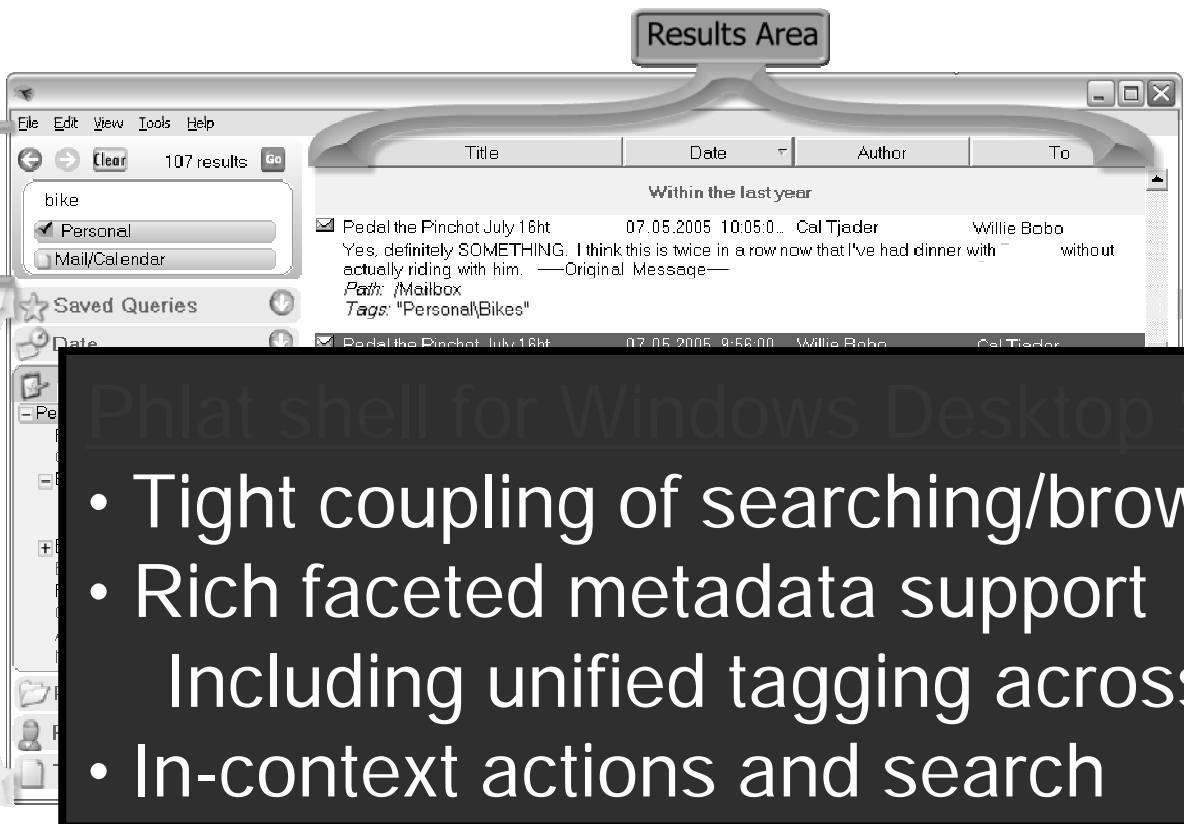
SIS -> Phlat

(richer metadata support, including tags)

- Shell for WDS
- Tight coupling of search and metadata
 - Q -> Results and associated metadata
- Tagging
 - A single set of user-generated tags to all content
- "In-context" search
 - See assoc metadata
 - Actions on results
 - Pivot on metadata ("Sideways search")



Phlat



The screenshot shows the Phlat search interface. On the left, there is a 'Query Area' with a search bar containing 'bike' and a 'Filter Area' with a 'Personal' filter selected. The 'Results Area' displays a list of search results with columns for Title, Date, Author, and To. The first result is 'Pedal the Pinchot July 16ht' from Cal Tjader to Willie Bobo, dated 07.05.2005 10:05:00. The interface also shows 'Saved Queries' and a 'Date' filter.

Phlat shell for Windows Desktop Search

- Tight coupling of searching/browsing
- Rich faceted metadata support
 - Including unified tagging across data types
- In-context actions and search

Download: <http://research.microsoft.com/adapt/phlat>

SIS ->Memory Landmarks

- Importance of time in re-finding
 - Absolute and relative time (e.g., landmarks in human memory)
- Identify and use landmarks facilitate information management and search
 - Timeline interface, augmented w/ landmarks
 - Bayesian models to identify memorable events
- Extensions beyond search, Life Browser

SIS, Timeline w/ Landmarks

Distribution of Results Over Time

The screenshot shows a software interface titled "Stuff I've Seen" with a search bar containing the word "earthquake" and 59 results found. The search results are displayed in a list with checkboxes, organized by a vertical timeline on the left. The timeline shows dates from 1999 to 2001. A callout box labeled "Search Results" points to the list of results. Another callout box labeled "Memory Landmarks" points to the timeline, listing "General (world, calendar)", "Personal (appts, photos)", and "<linked by time to results>".

Search Results

Memory Landmarks

- General (world, calendar)
- Personal (appts, photos)
- <linked by time to results>

Search Results

earthquake

submit 59 results found

AND OR

fuzzy match exact match

show advanced controls

zoom in zoom out

study

Stuff I've Seen

4/15/2001

4/7/2001

history.xls

4/3/2001

3/10/2001

PowerPoint Presentation

from Michael B. Smith - Re: Greetings

from John Smith - RE: Followup from Earthquake Meeting

from Tim Stevens- RE: followup from the "earthquake meeting"

from Janet S. Hazeltine - Re: Rockin and Rollin ...

from Bart Thomason- Followup from Earthquake Meeting

from Bart Thomason- Greetings

re jcdl 2002-3.txt

from Samuel Atkins French - Re: Shakin Sue ...

from Jennifer Rogerson - Re: lost letters

from Bart Thomason- RE: quake?

from Edward Finerhold- A bit shaken?

Proceedings Template - WORD

from John Smith - RE: followup from the "earthquake meeting"

from Bart Thomason- Earthquake Report

from Mike Yin - RE: Relevance test review 1

from Jenny Henson- are you ok?

from Cargon, Angeline - FW: Earthquake

from Trent van Erliche (MSR)- Earthquake coverage

from Phillip Lander- Earthquake

2/28/2001

2/24/2001

1/1/2001

9/15/2000

9/15/2000

8/1/2000

7/21/2000

1/1/2000

1/1/2000

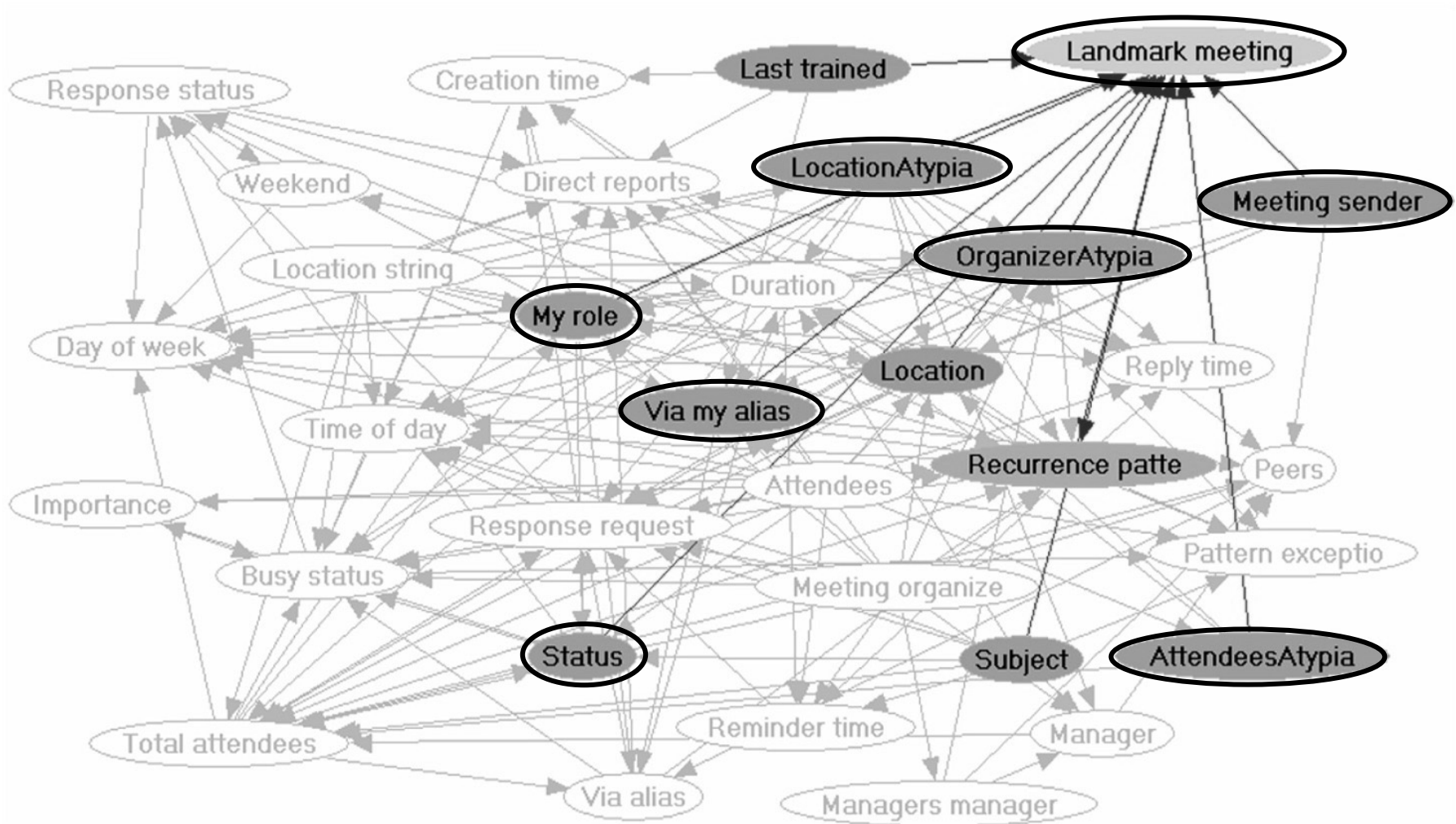
12/20/1999

from Christine Borgman - RE: Seattle plans

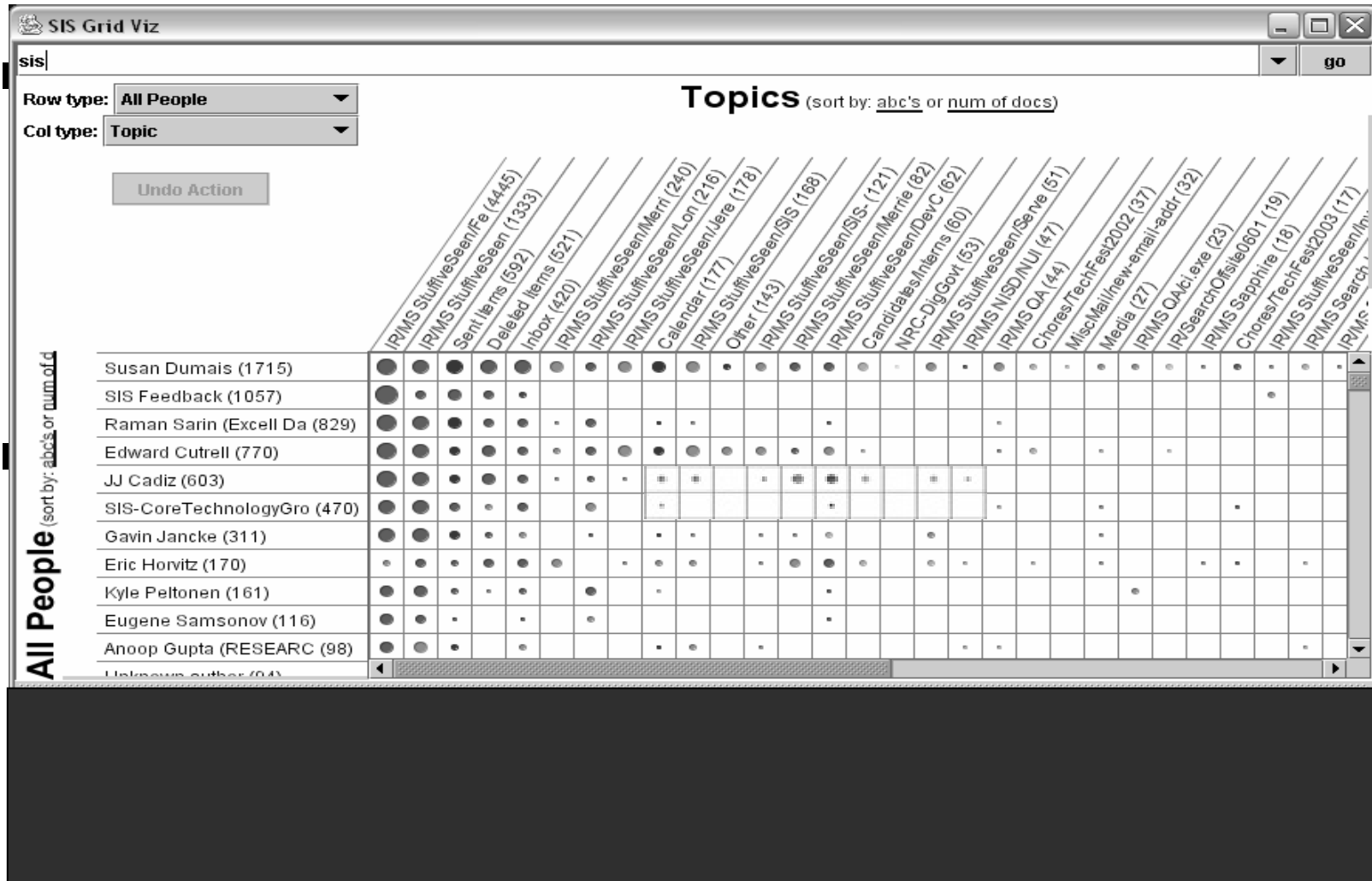
sigir99_report.txt

1999-earthquake-graph.gif

Landmarks, key dependencies (from learned graphical model)



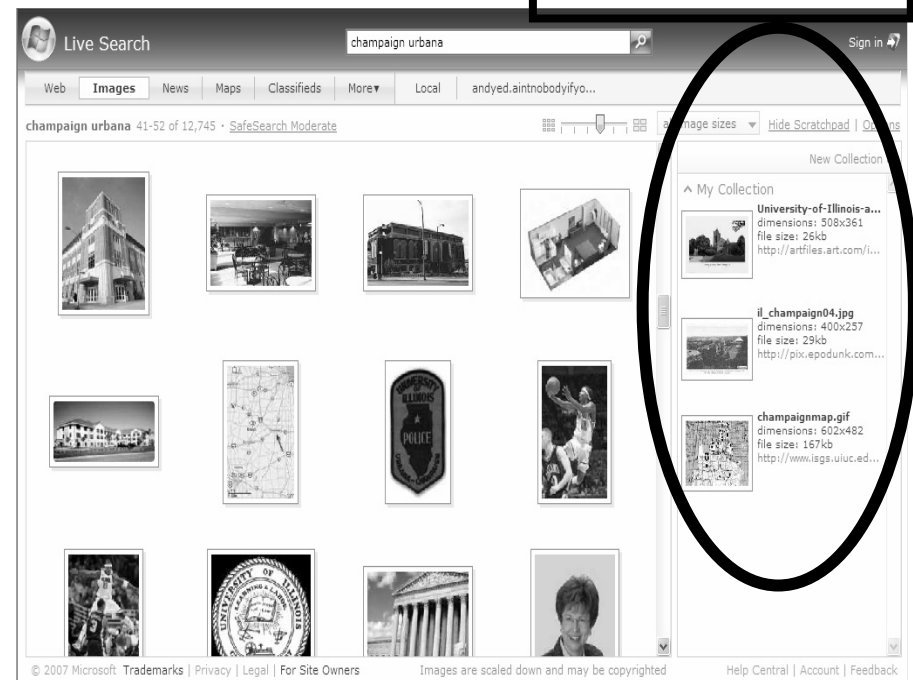
Visualizing Trends



Beyond "Search" (Supporting Tasks)

- Requires going beyond just retrieving documents
 - Retrieve -> Analyze -> Use
- Lightweight scratchpad or workspace support
 - Iterative and evolving nature of search
 - Resuming at a later time or on other device
 - Sharing with others
- Richer sense-making

ScratchPad



<http://images.live.com>

The Person in Personal

- Support the person in searching personal content
 - Rich associations, incl time, people, contexts of use
- Unified access to personal content
 - vs. information islands
- Representation and use of rich metadata
 - vs. single ranked list or single hierarchy
 - Fast and flexible access via what the person remembers, including cognitive landmarks
- Thinking outside the search box
 - In-context search, implicit queries (stuff I should see), richer task support (search is not the goal !)