The Person in Personal
(Supporting the Person
in Searching Personal Content)

Susan Dumais

Microsoft Research

http://research.microsoft.com/~sdumais
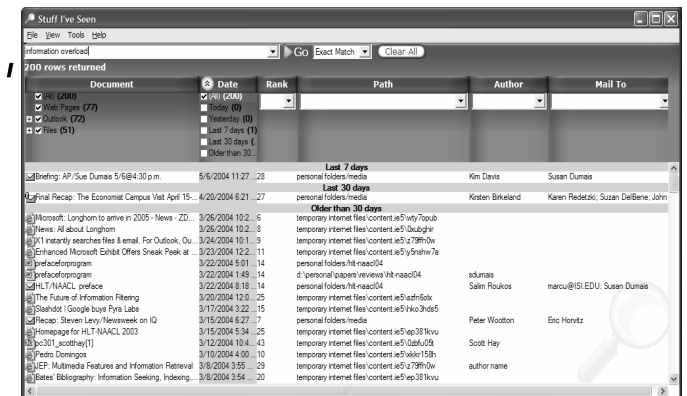
Stuff I've Seen (SIS)

- **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 ... ...C of Ben ... ...udding."*

- *"Wh... ... ...e [our team... ... ...es."*

- *"The ... ...at the Kona..."*

Importance of:
- Time, People
- Also type/source, entities, events, history of interaction
- Supporting what the searcher remembers about the information

SIS Interface

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

- Perso

- Query

  - Sho   rators)
  - But   (48%)
    - F   n),
    From.,
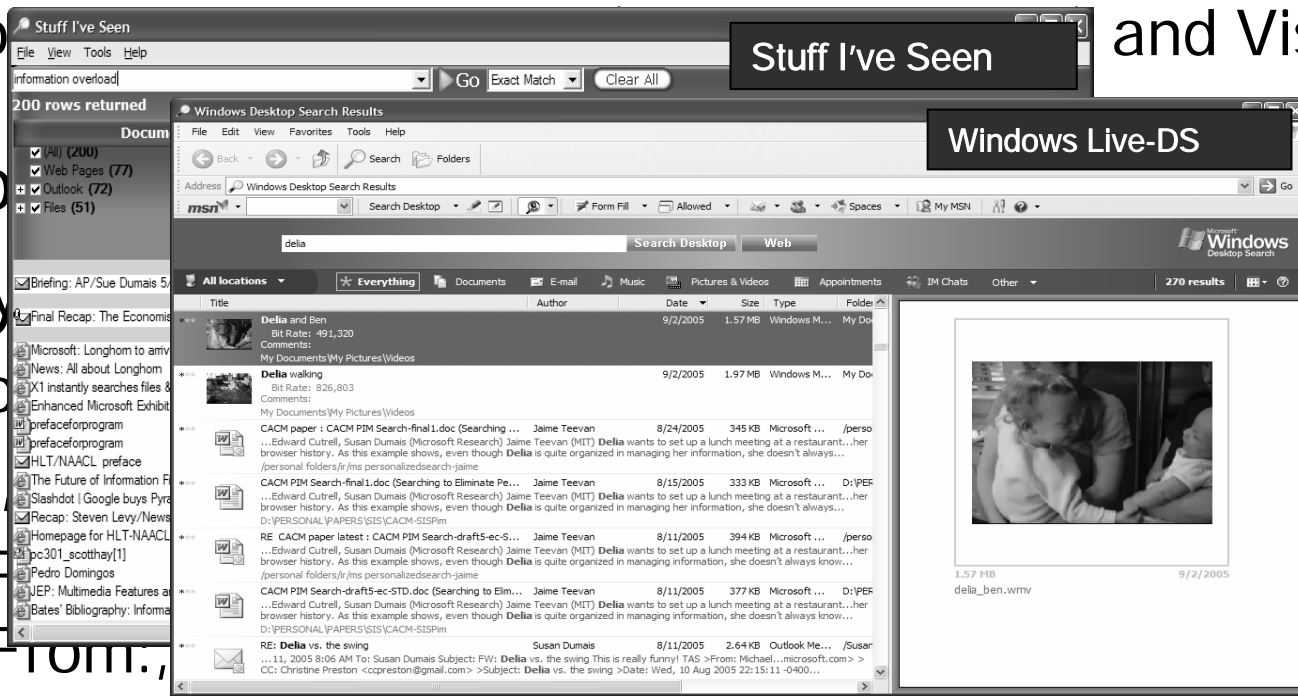
SIS Usage, cont'd

- ## 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
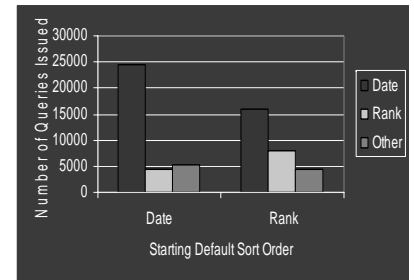
Ranked list vs. Metadata
(for personal search)

- 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
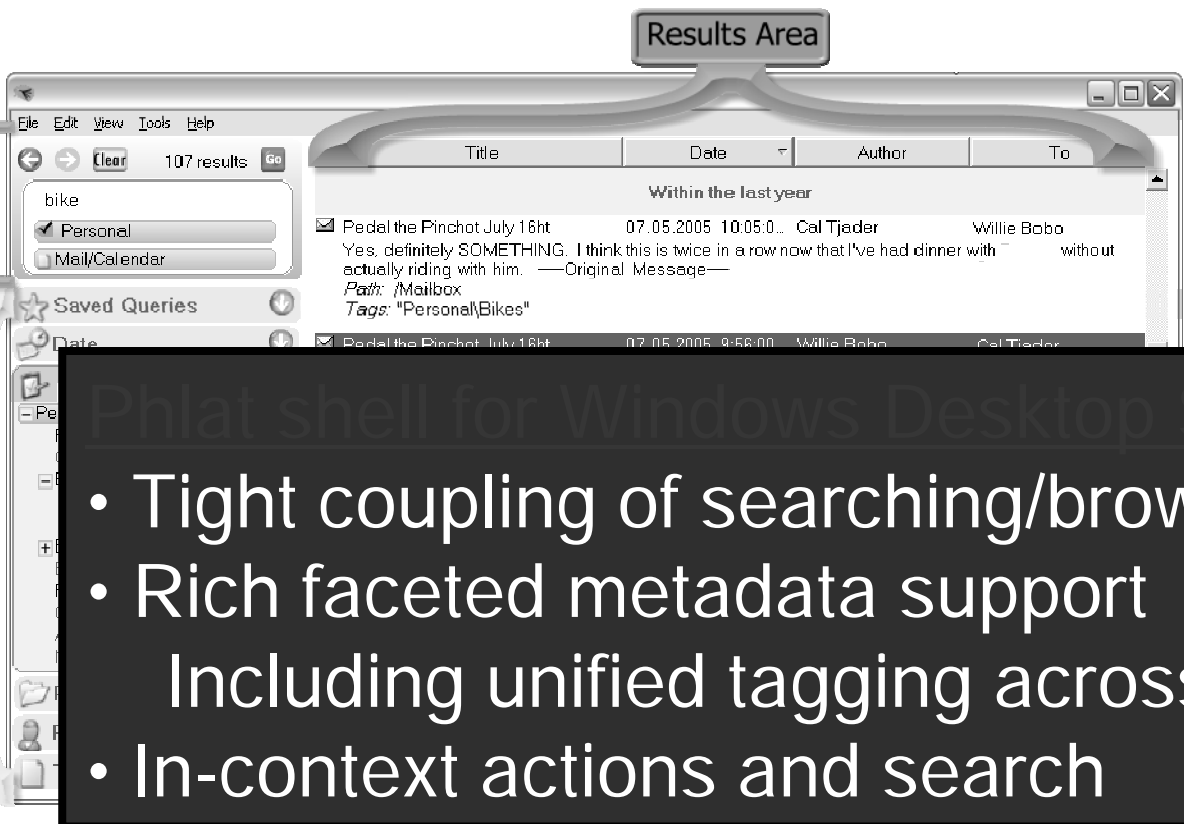
SIS -> Phlat
(richer metadata support, including tags)

- **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")
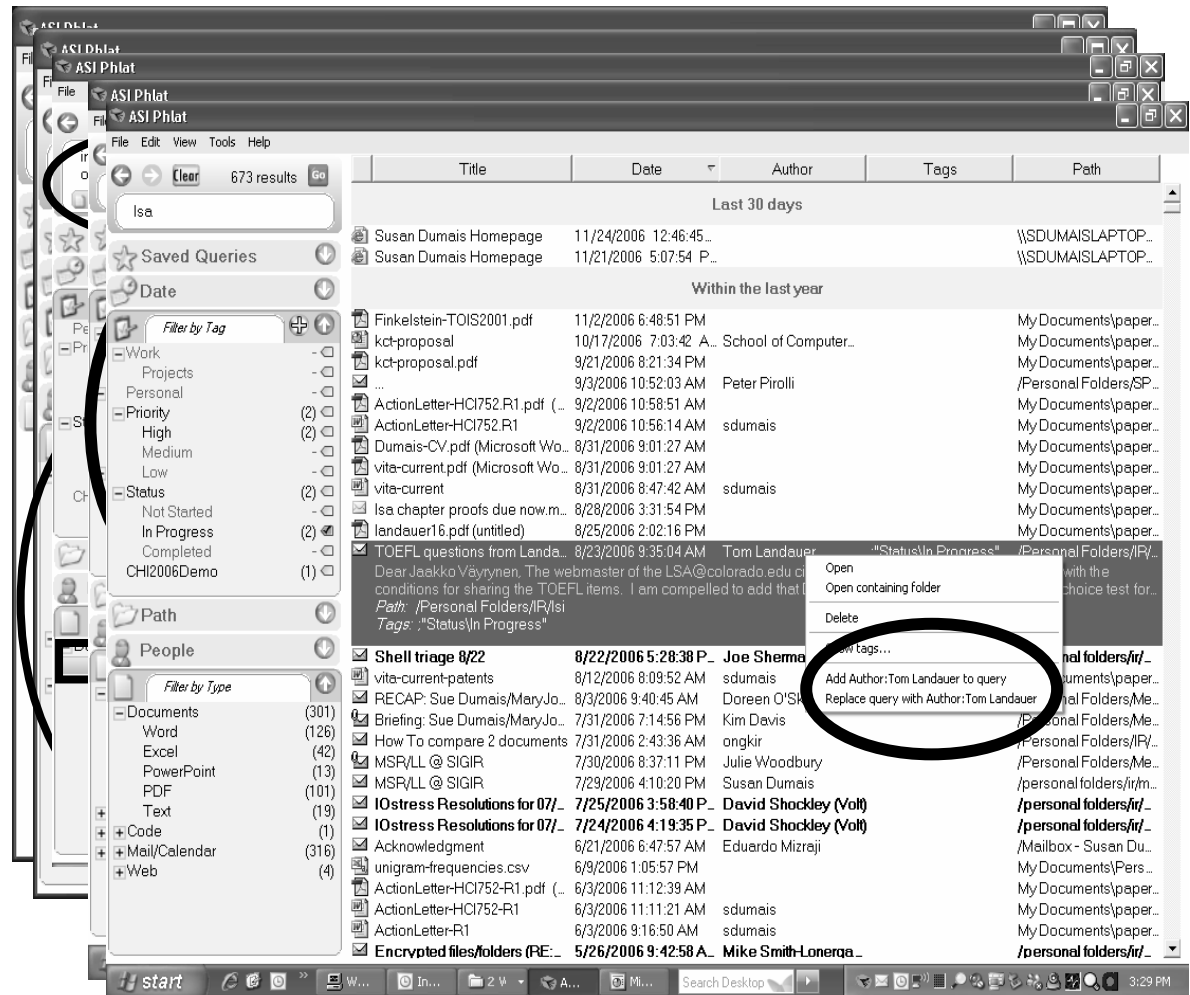
Phlat

Results Area

Query Area

Filter Area

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

Stuff I've Seen

Date	
4/15/2001	history.xls
4/7/2001	
4/3/2001	PowerPoint Presentation
3/10/2001	from Michael B. Smith — Re: Greetings

Search Results

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, Angelina — FW: Earthquake
from Trent van Erliche (MSR)— Earthquake coverage

2/28/2001	from Phillip Lander— Earthquake
2/24/2001	
1/1/20	

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

9/15/2	
9/15/2	
8/1/20	
7/21/2000	sigir 2000 - venue.htm
1/1/2000	
1/1/2000	
12/20/1999	from Christine Borgman — RE: Seattle plans
	sigir99_report.txt
	1999-earthquake-graph.gif

QUERY:

earthquake

submit   59 results found

AND   fuzzy match
OR   exact match

show advanced controls

zoom in   zoom out

study

Landmarks, key dependencies
(from learned graphical model)
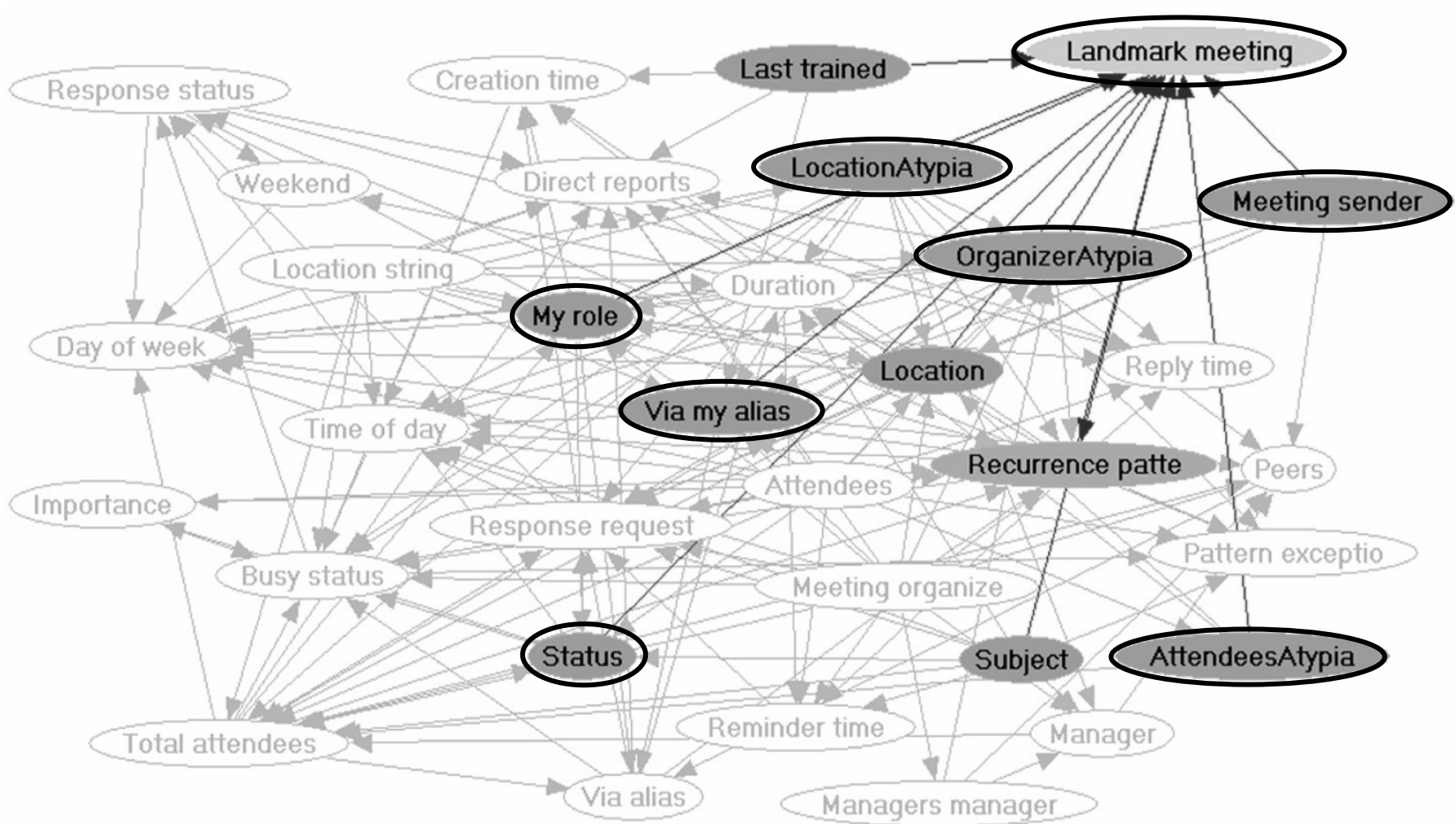
Visualizing Trends
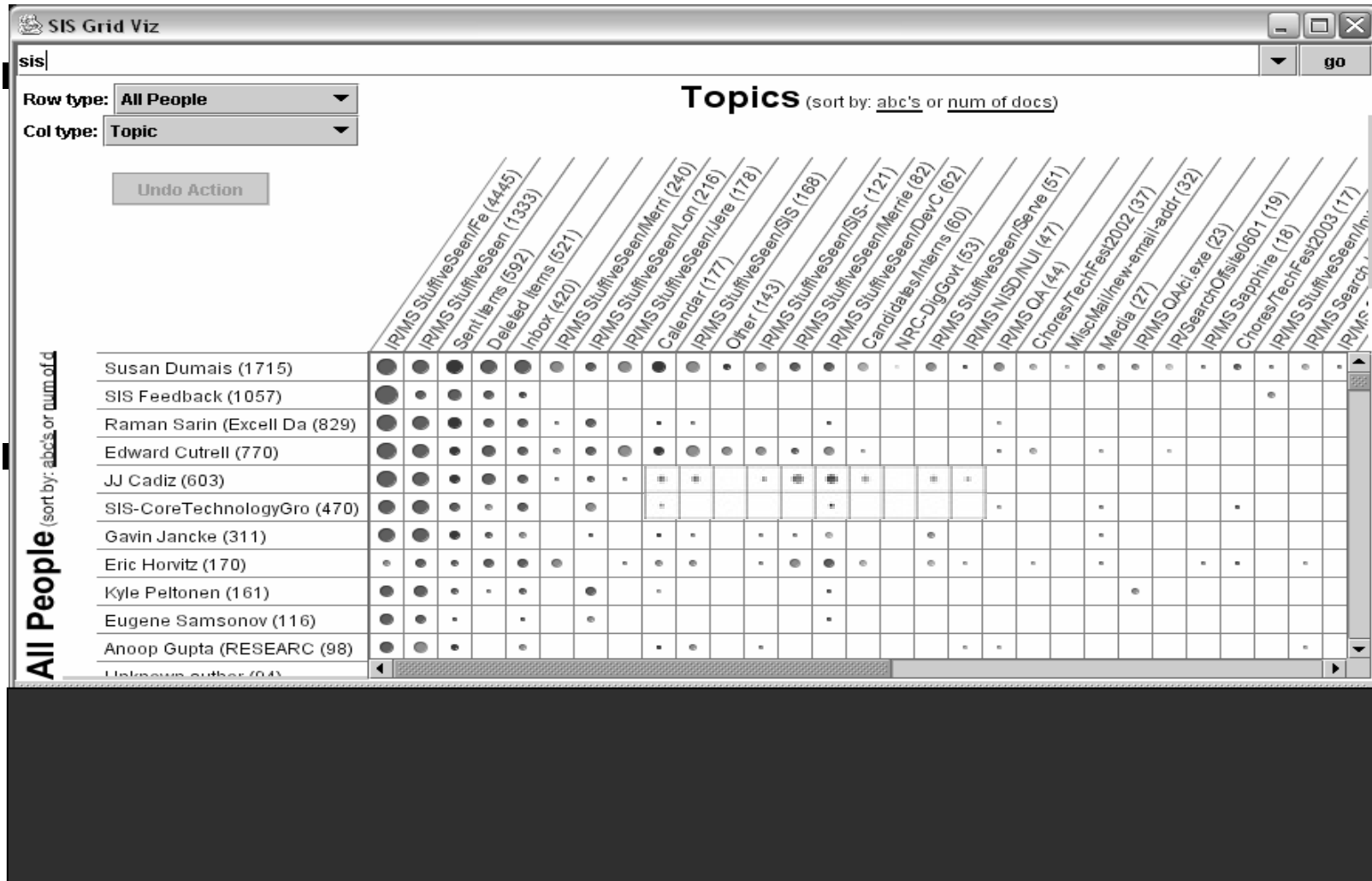
Beyond "Search"
(Supporting Tasks)

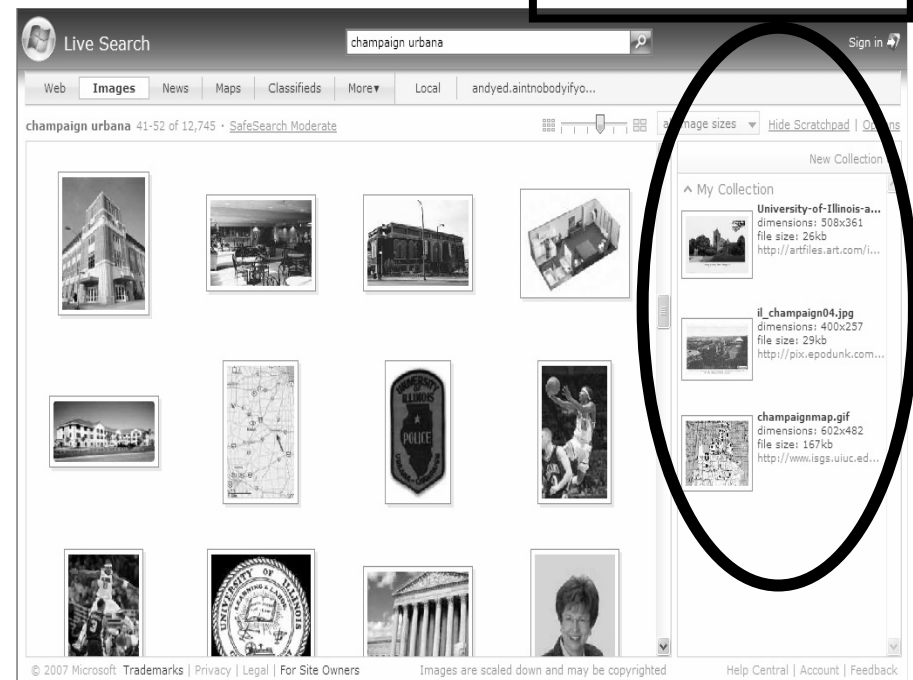- **Requires going beyond just retrieving documents**
  - Retrieve -> Analyze -> Use
- **Lightweight scratchpad or workspace support**
  - Iterative and evolving nature of search
  - Resuming at a later time or on other device
  - Sharing with others
- **Richer sense-making**

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 !)