

Dynamic Information Environments

Susan Dumais

Microsoft Research

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

In collaboration with:

Jaime Teevan, Eytan Adar, Jon Elsas, Dan Liebling, Richard Hughes

Outline

- Web search and context
- Temporal dynamics of information
 - Characterizing change
 - Content changes over time
 - People re-visit and re-find
 - Relationships between content change and re-access
 - Improving retrieval and understanding
 - Building support for understanding change (e.g., DiffIE)
 - Leveraging dynamics for improved retrieval

Web Search at 15

What's available

How it's accessed

■ Number of pages indexed

- 7/94 Lycos -
- 95 - 10^6 millions
- 97 - 10^7
- 98 - 10^8
- 01 - 10^9 billions
- 05 - 10^{10} ...

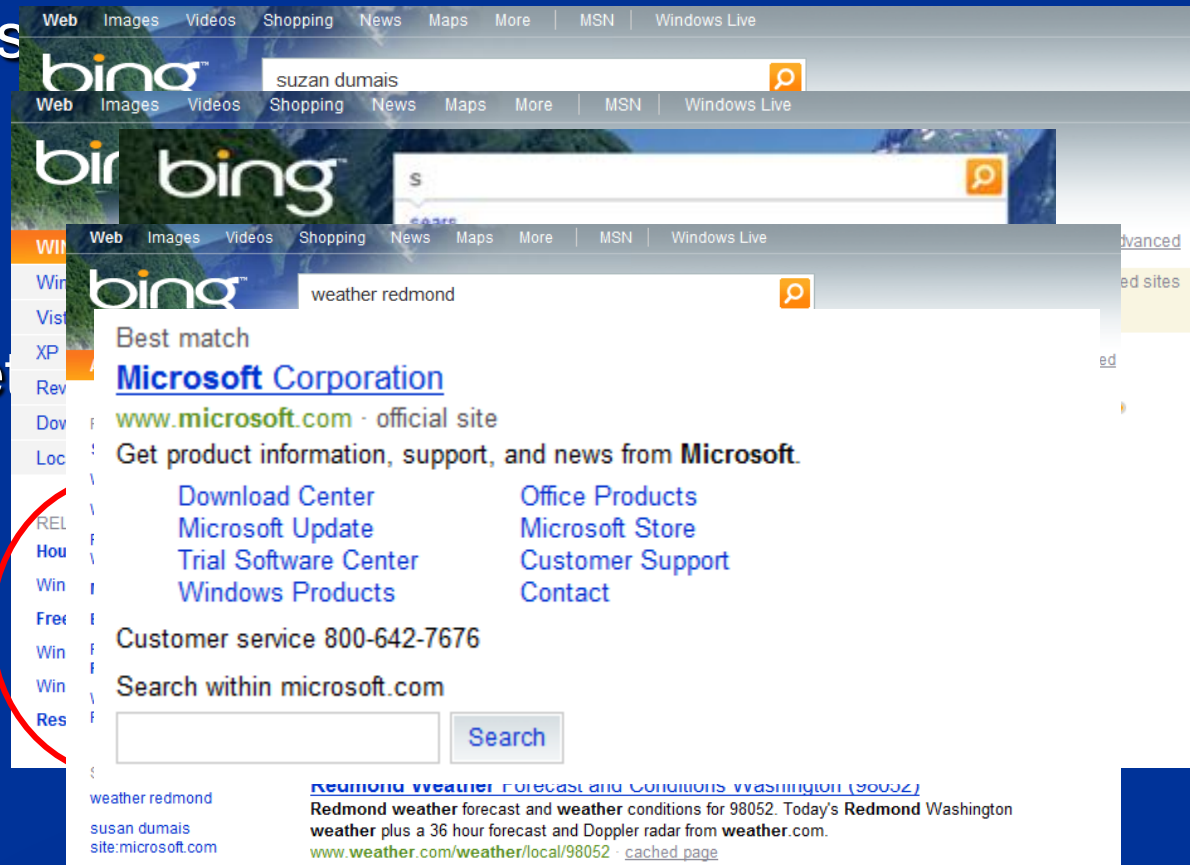
■ Types of content

- Web pages, newsgroups
- Images, videos, maps
- News, blogs, spaces
- Shopping, local, desktop
- Books, papers, many formats
- Health, finance, travel ...



Support for Searchers

- The search box
- Spelling suggestions
- Query suggestions
- Auto complete
- Inline answers
- Richer snippets
- But, we can do better



Search and Context

User
Context



Query Words



Query Words



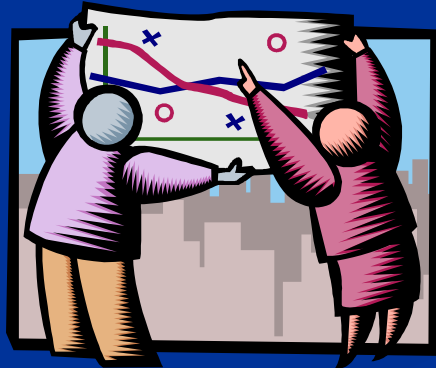
Ranked List



Document
Context



Task/Use
Context



Systems/Prototypes

- New capabilities and experiences
- Algorithms and prototypes
- Deploy, evaluate and iterate

Inter-Relationships among Documents

Categorization and Metadata

Reuters, spam, landmarks, web categories ...

Domain-specific features, time

Interfaces and Interaction

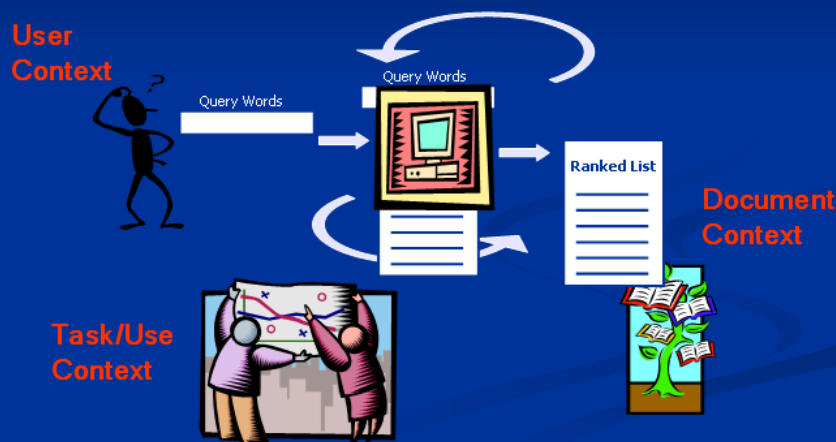
Stuff I've Seen, Phlat, Timelines, SWISH

Tight coupling of browsing and search

Redundancy

Temporal Dynamics

Search and Context



Modeling Users

Short vs. long term
Individual vs. group
Implicit vs. explicit

Using User Models

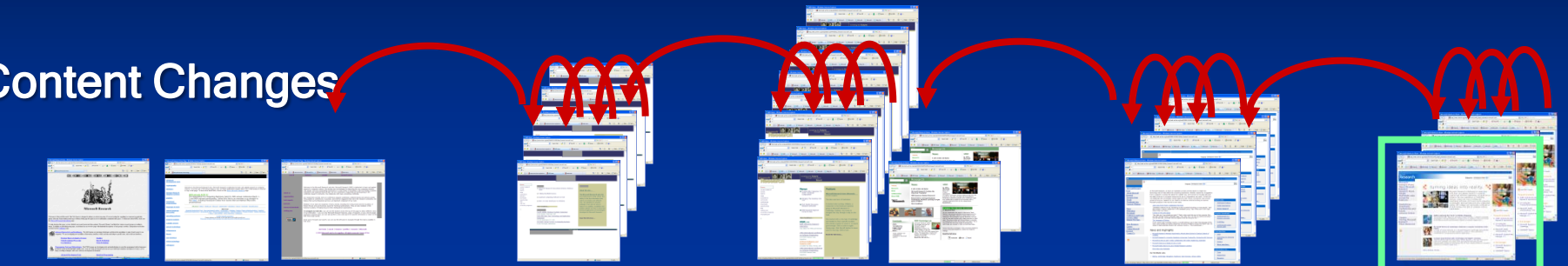
Stuff I've Seen (re-finding)
Personalized Search
News Junkie (novelty)
User Behavior in Ranking
Domain Expertise at Web-scale

Evaluation

- Many methods, scales
- Individual components and their combinations

Information Dynamics

Content Changes



1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009

1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009

User Visitation/ReVisitation

Today's Browse and Search Experiences

But, ignores ...

Digital Dynamics Easy to Capture

- Easy to capture
- Few tools support dynamics

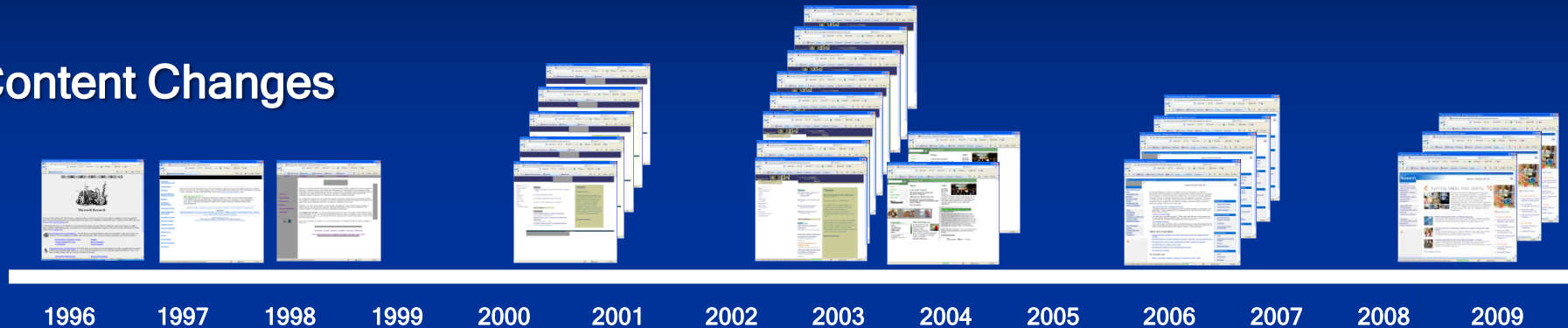


Information Dynamics

- Characterizing change
 - Content changes over time
 - People re-visit and re-find
 - Relationships between content change and re-access
- Improving retrieval and understanding
 - Building support for understanding change (e.g., DiffIE)
 - Leveraging dynamics for improved retrieval

Characterizing Change

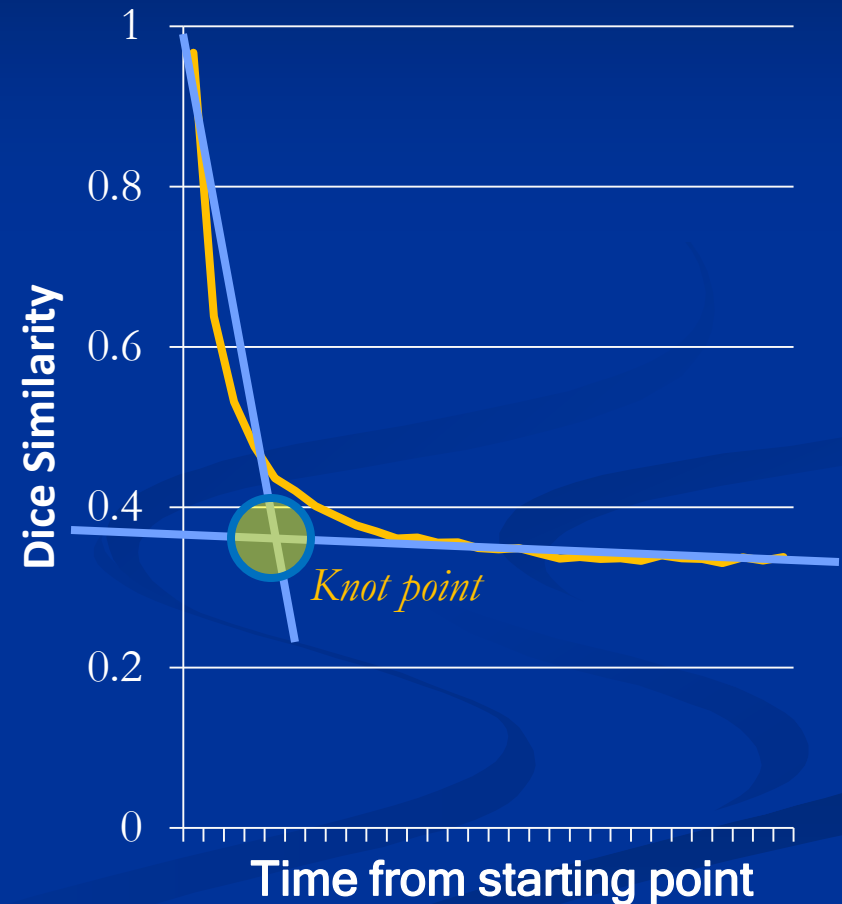
Content Changes



- Large- scale Web crawls, over time
 - Revisited pages
 - 55,000 pages crawled hourly for 18+ months
 - Unique users, visits/user, time between visits
 - Judged pages (Relevance to a query)
 - 6 million pages crawled every two days for 6 months

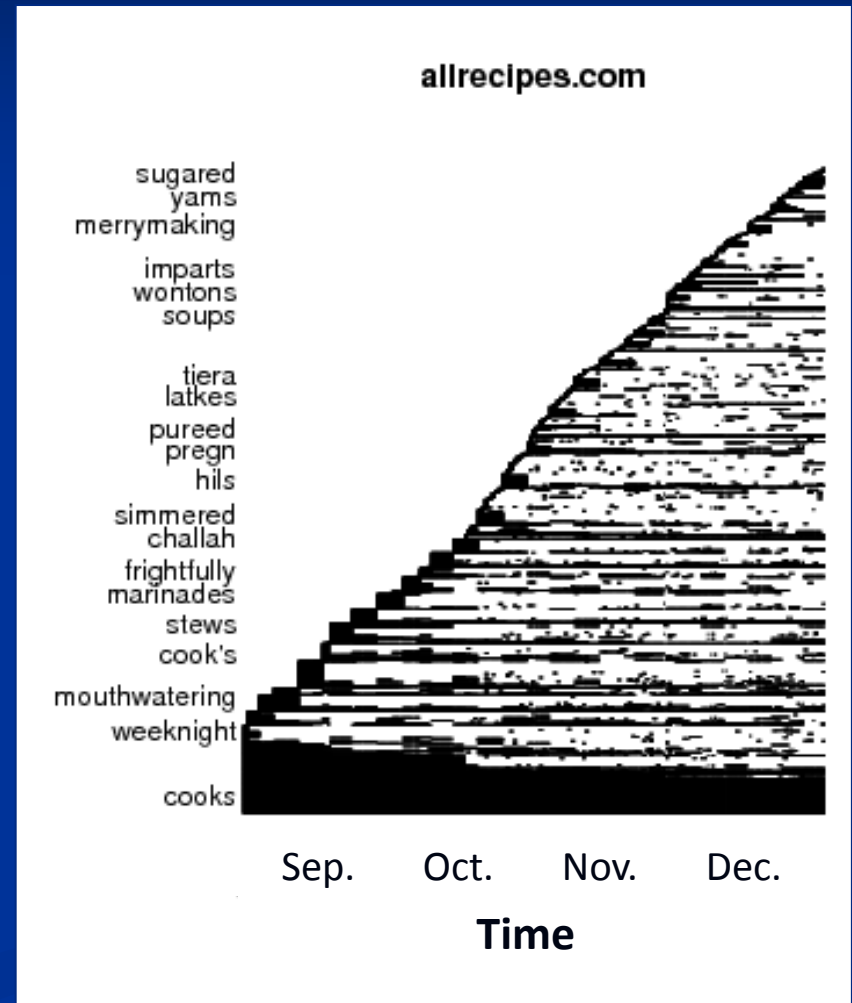
Measuring Web Page Change

- Summary metrics
 - Number of changes
 - Time between changes
 - Amount of change
- Change curves
 - Fixed starting point
 - Measure similarity over different time intervals

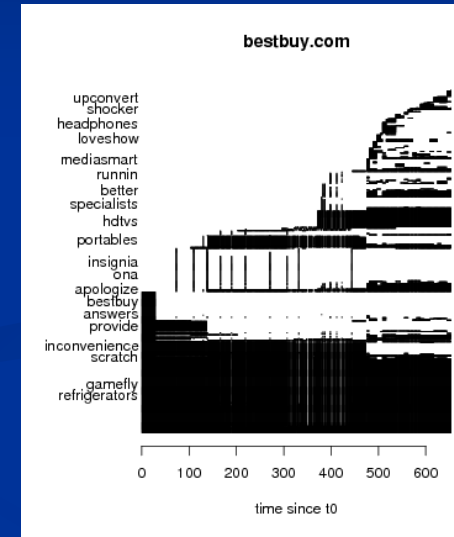
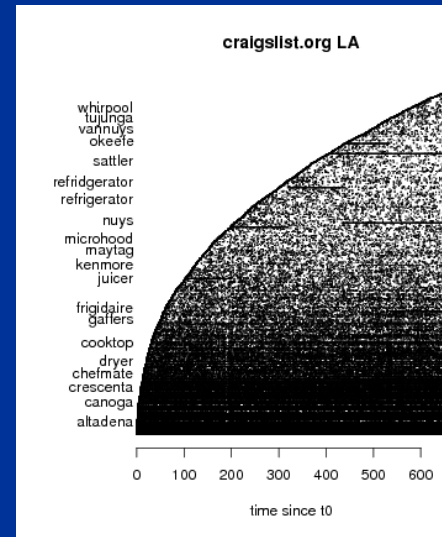
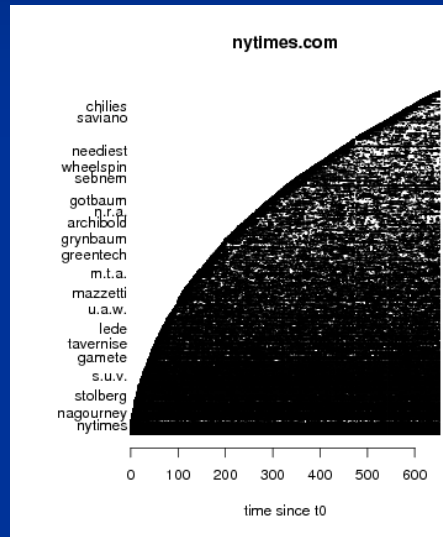
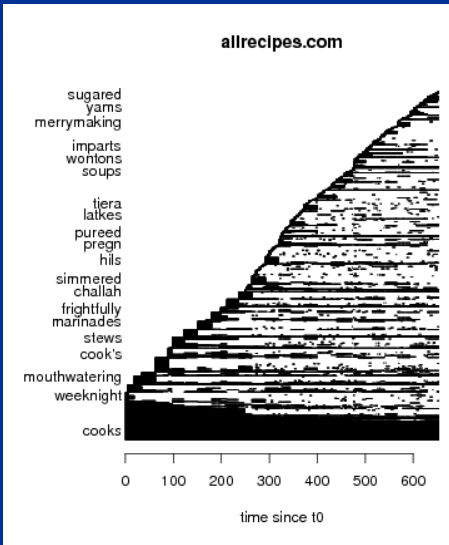


Measuring Within-Page Change

- DOM structure changes
- Term use changes
 - Divergence from norm
 - cookbooks
 - salads
 - cheese
 - ingredient
 - bbq
 - “Staying power” in page



Example Term Longevity Graphs



Revisitation on the Web

- Revisitation patterns
 - Log analyses
 - Toolbar logs for *revisitation*
 - Query logs for *re-finding*
 - User survey to understand intent in revisitations

1996

1997

1998

1999

2000

2001

2002

2003

2004

2005

2006

2007

2008

2009

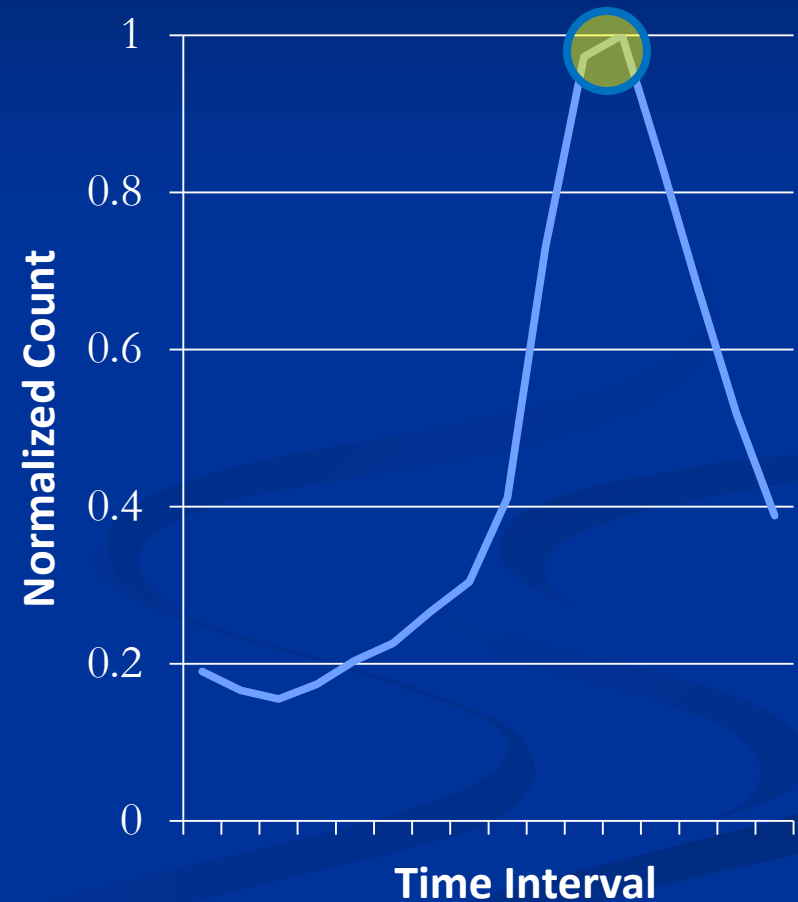


Users Revisit

What's the last Web page you visited?

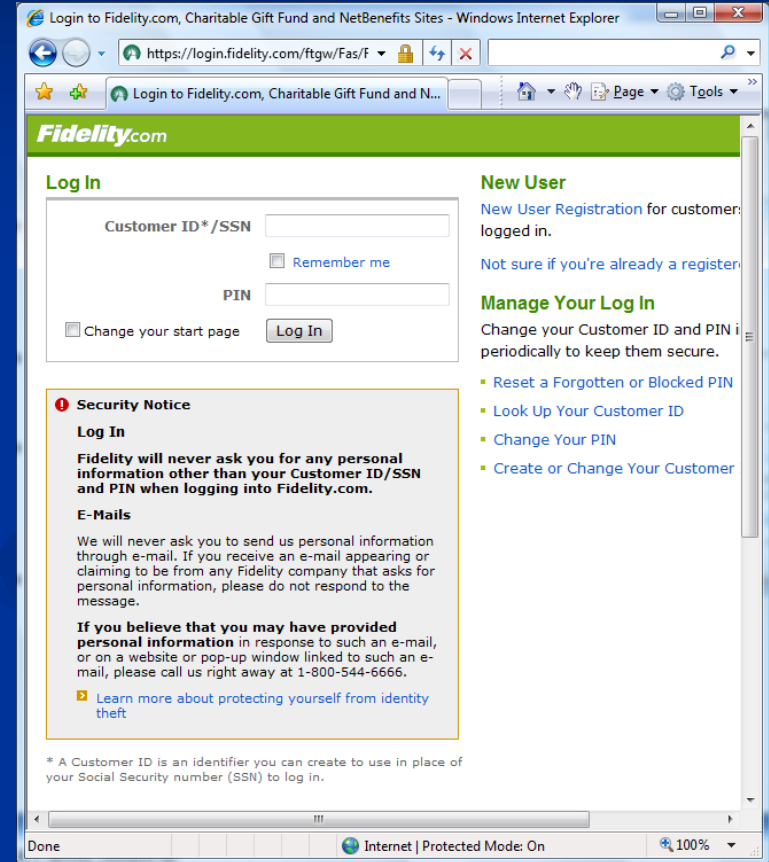
Measuring Revisitation

- Summary metrics
 - Unique visitors
 - Visits/user
 - Time between visits
- Revisitation curves
 - Histogram of revisit intervals
 - Normalized



Four Revisitation Patterns

- *Fast*
 - Hub-and-spoke
 - Navigation within site
- *Hybrid*
 - High quality *fast* pages
- *Medium*
 - Popular homepages
 - Mail and Web applications
- *Slow*
 - Entry pages, bank pages
 - Accessed via search engine



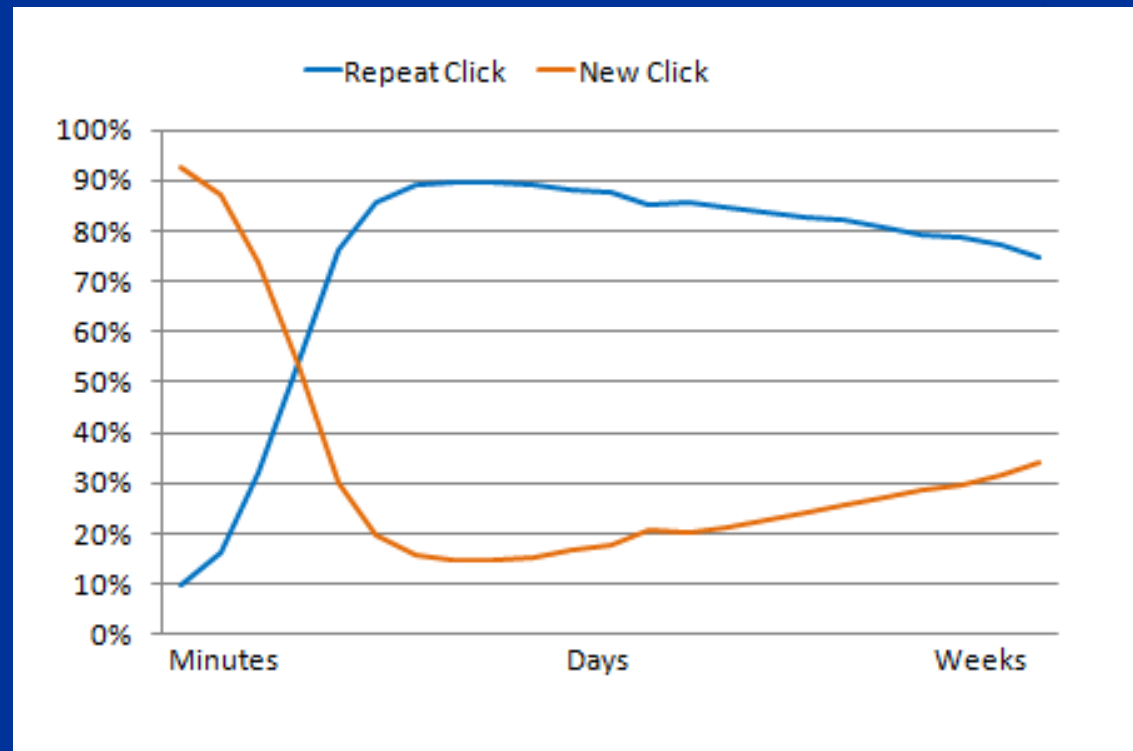
Search and Revisitation

- Repeat query (33%)
 - *microsoft research*
- Repeat click (39%)
 - <http://research.microsoft.com>
 - Q: *microsoft research, msr ...*
- Big opportunity (43%)
 - 24% “navigational revisits”

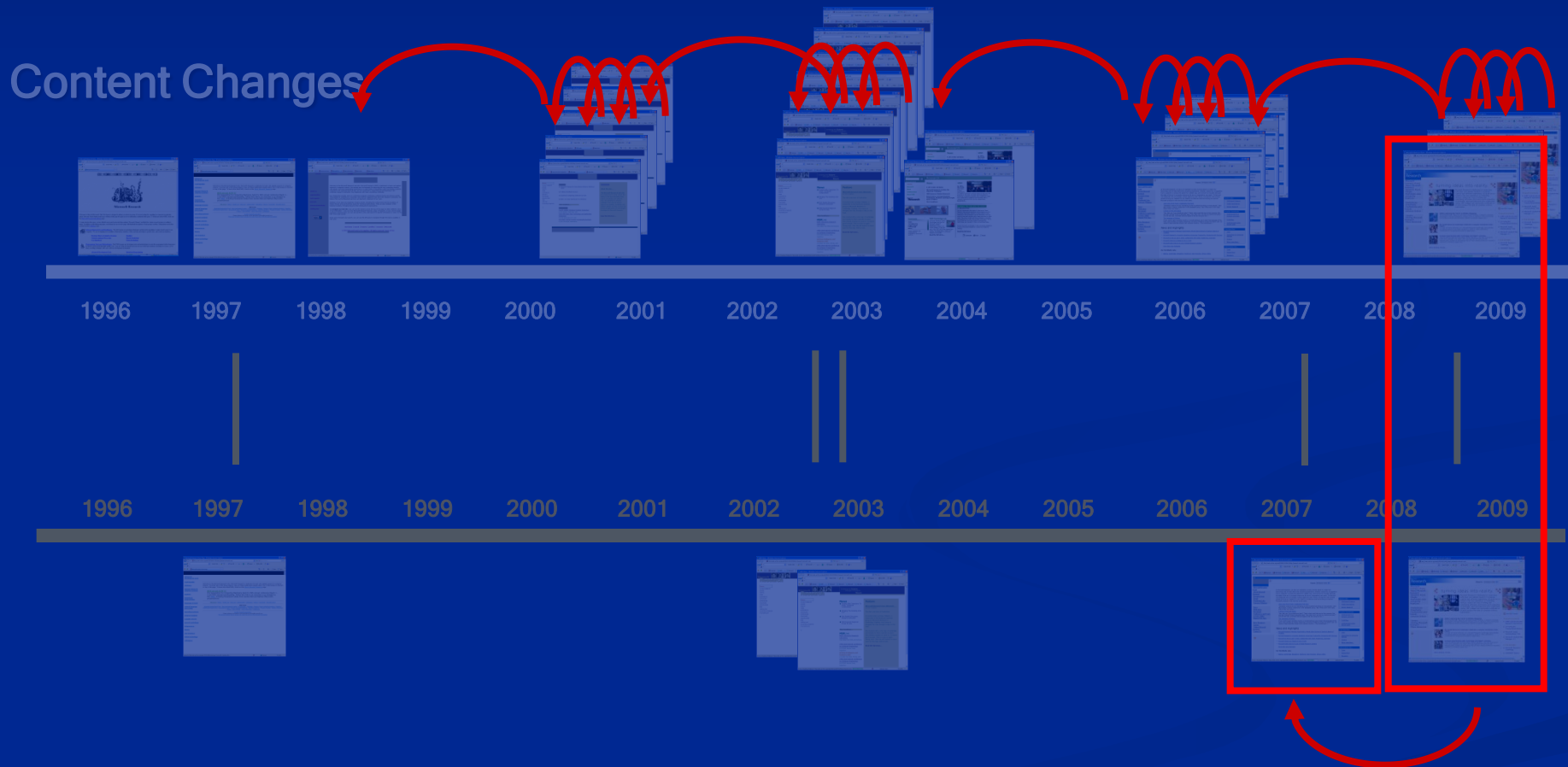
		Repeat Click	New Click
Repeat Query	33%	29%	4%
New Query	67%	10%	57%
		39%	61%

Repeat Clicks for Repeat Queries

- Within session: Repeat query -> New click
- Across sessions: Repeat query -> Repeat click



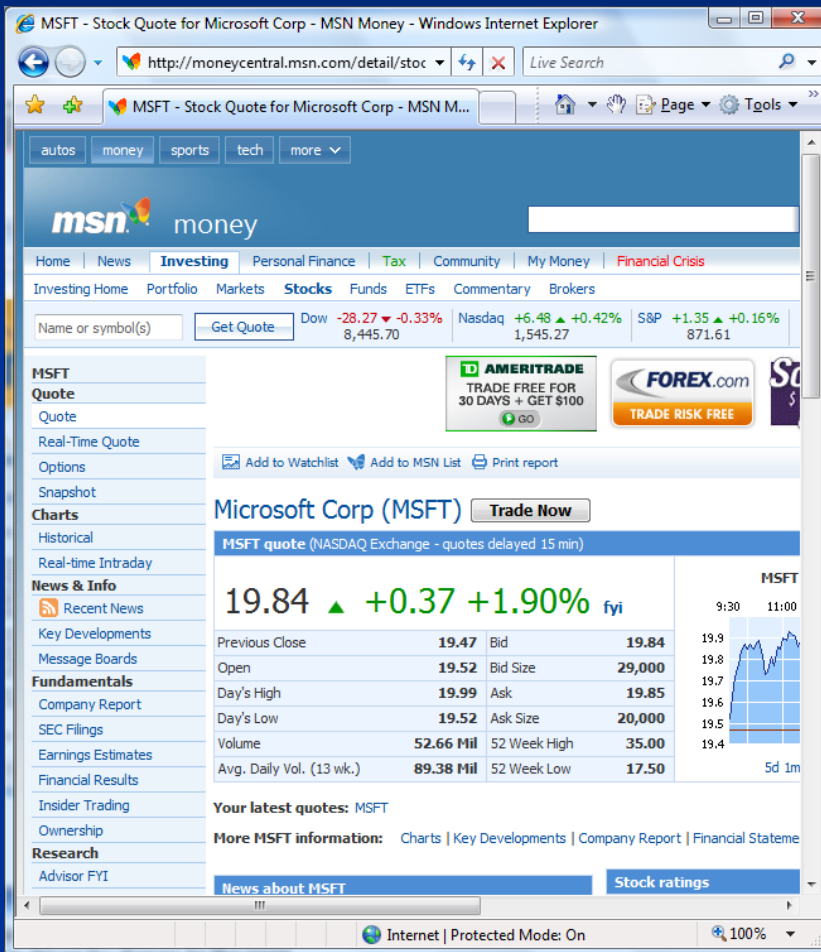
Relationships Between Revisitation and Change



Why did you revisit the last Web page you revisited?

Possible Relationships


- Interested in change
 - Monitor
- Effect change
 - Transact
- Change unimportant
 - Find new
- Change can interfere
 - Re-find



Understanding the Relationship

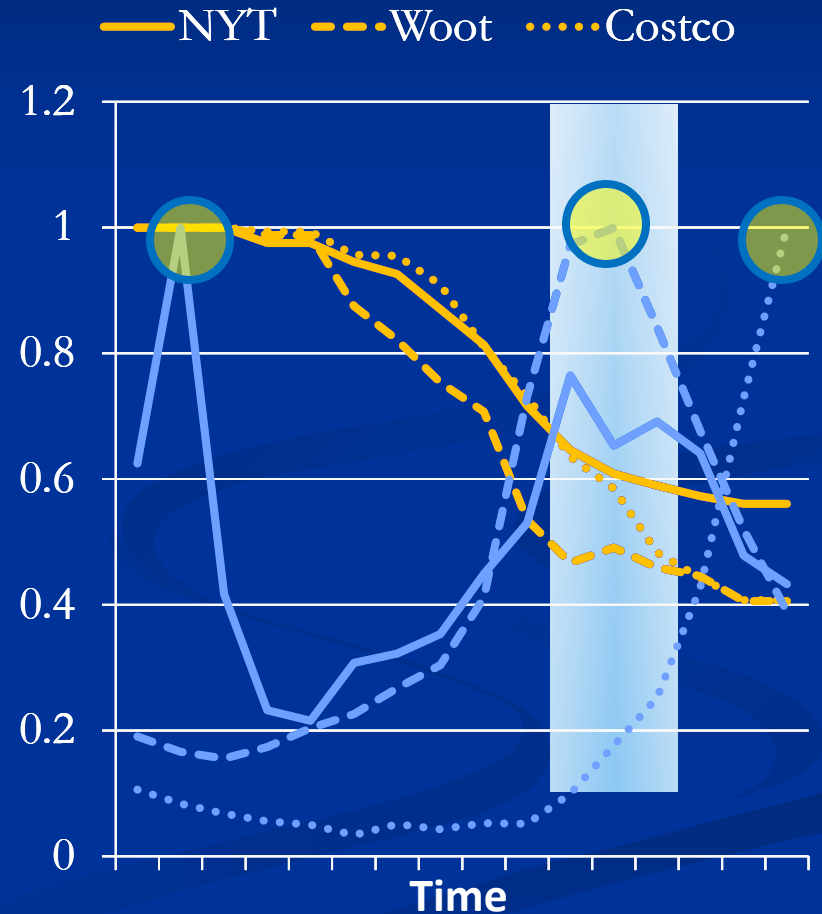
- Compare summary metrics
 - Revisits: Unique visitors, visits/user, interval
 - Change: Number, interval, Dice

	Number of changes	Time between changes	Dice coefficient
2 visits/user	172.91	133.26	0.82
3 visits/user	200.51	119.24	0.82
4 visits/user	234.32	109.59	0.81
5 or 6 visits/user	269.63	94.54	0.82
7+ visits/user	341.43	81.80	0.81



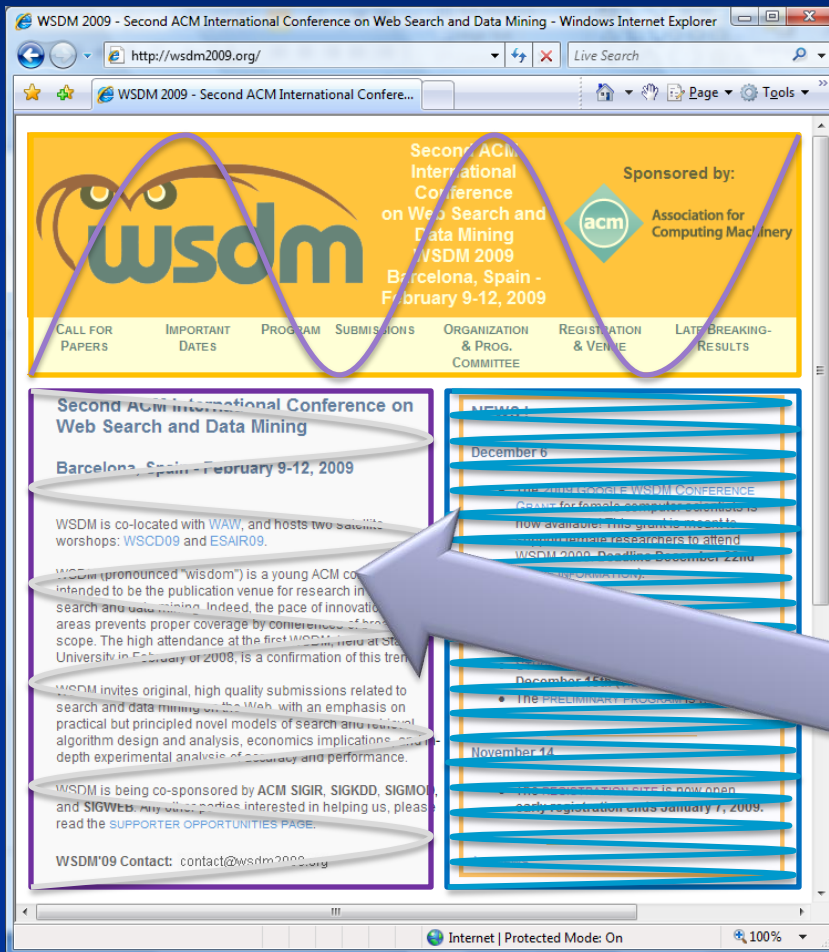
Comparing Change and Revisit Curves

- Three pages
 - New York Times
 - Woot.com
 - Costco
- Similar change patterns
- Different revisitation
 - NYT: *Fast* (news, forums)
 - Woot: *Medium*
 - Costco: *Slow* (retail)



Within-Page Relationship

- Page elements change at different rates
- Pages revisited at different rates



- “Resonance” can serve as a filter for interesting content

Protect and Provide.
HSBC Insurance

The New York Times

Thursday, August 7, 2008 Last Update: 3:02 PM ET

Instant* Term Life.
HSBC Insurance

 Search

Get Home Delivery | Personalize Your Weather

- JOBS
- REAL ESTATE
- AUTOS
- ALL CLASSIFIEDS
- WORLD
- U.S.
- Politics
- Washington
- Education
- N.Y./REGION
- BUSINESS
- TECHNOLOGY
- SPORTS
- OLYMPICS
- SCIENCE
- HEALTH
- OPINION
- ARTS
- Books
- Movies
- Music
- Television
- Theater
- STYLE
- Dining & Wine
- Fashion & Style
- Home & Garden
- Weddings/ Celebrations
- TRAVEL
- Blogs
- Cartoons / Humor
- Classifieds
- Corrections
- Crossword/ Games
- First Look
- Learning Network
- Multimedia
- NYC Guide
- Obituaries
- Podcasts

Coalition in Pakistan Moves to Impeach Musharraf

By JANE PERLEZ 12:19 PM ET
 Pakistan's usually fractious coalition government moved decisively for the first time on Thursday to impeach President Pervez Musharraf.

• Times Topics: Pervez Musharraf | Pakistan
[Post a Comment](#) | [Read \(33\)](#)

E.P.A. Won't Ease Requirements for Ethanol in Gas

By MATTHEW L. WALD 23 minutes ago
 The agency denied a request to cut the quota for the use of ethanol in cars, concluding that the goal of reducing the nation's reliance on oil trumps any effect on food prices.

In Anthrax Scientist's E-Mail, Hints of Delusions

By ERIC LIPTON
 The e-mail messages of Bruce E. Ivins provide a narrative of his paranoid episodes, dating to 2000.

- The Lede: Reactions to the Case 12:22 PM ET
- F.B.I. Presents Anthrax Case, Saying Scientist Acted Alone
- Times Topics: Anthrax

2008 BEIJING OLYMPICS



Elizabeth Dalziel/Associated Press

SPORTS OF THE TIMES

Raising the Bar and Raising Elbows

By GEORGE VECSEY 11:07 AM ET
 Increased physicality in women's basketball will not give a black eye to the entire sport. However, it most certainly gave a black eye to Penny Taylor of Australia on Tuesday.

INTERACTIVE SCHEDULE

Friday	Saturday	Sunday
1 event	4 events	18 events
1 event	1 event	23 events

⌘ Olympics Tracker

Follow the schedules of your favorite sports, day by day and event by event.

Judge Orders Detroit Mayor Jailed

By NICK BUNKLEY 39 minutes ago

MORE OLYMPICS NEWS

- Men's Soccer: U.S. 1, Japan 0
- Bush Arrives in Beijing
- Silver Keeps Her Place on U.S. Relay Team
- Federer Arrives, Soon to Lose His No. 1 Status
- Rings Blog
- Complete Coverage »

OPINION »

An Olive Branch From the Dalai Lama

By NICHOLAS D. KRISTOF

For the first time, the Tibetan leader has expressed a willingness to accept Communist Party rule over his homeland.



[Post a Comment](#) | [Read \(104\)](#)

THEATER »

With Echoes of Chekhov

A. R. Gurney's "Buffalo Gal" is about an actress who retreads the boards of her hometown.



• Times Topics: A. R. Gurney

OPINION »

- Collins: Energy Drill
- Cohen: Spanish Summer
- Egan: Nanny Nation
- Self: Vanishing Cliffs
- Editorial: Hamdan's Trial
- [Comments \(221\)](#)

AMERITRADE
\$9.99 trades. [GO](#)
 No guesswork.

MARKETS »

At 3:05 PM ET

S.&P.	Dow	Nasdaq
500	11,481.02	2,362.26
1,271.12	-175.05	-16.11
-18.07	-1.50%	-0.68%
-1.40%		

GET QUOTES

My Portfolios »

Stock, ETFs, Funds [Go](#)

With every purchase, Visa prevents, detects and resolves online fraud.

SAFE. SECURE.



Go to Visa.com/Secure for details



Hi. Are you new? [Start here.](#)

username password ([I forgot!](#))

HP Compaq 6820s 2GHz 17" Widescreen Notebook

\$529.99
+ \$5 shipping

CONDITION: **New**

PRODUCT: 1 HP Compaq 6820s Celeron 2.0GHz, 160GB, 17" Widescreen, 2GB DDR2, DL-DVDRW, 802.11a/b/g, Vista

[DISCUSS THIS PRODUCT \(176 COMMENTS\)](#)



[PRODUCT INFO](#) [PRODUCT STATS](#)

Sometimes The Past Is Hard To Let Go

"-and when she told me, I just couldn't deal. So I left, just like I always did. Just like my father."

Tom, thank you for sharing all that. Have one of these tissues. Do you feel like you've made a forward step?

"I do, Jack, I really do. Can I maybe have a second tissue?"

As many as you like, Tom, you don't even need to ask. Let's hear from someone else now. HP Compaq 2GHz Celeron 17" Widescreen Notebook, tell us more about your loss.

Well, with me, it was the teasing. See, I'm in the tech industry, which is a really high stress field. And, you know, growing up, everyone in a position of authority says, you know, play the hand you're dealt. But I've got a 2.0GHz Intel Celeron 550 Processor with 1MB L2 cache and 533MHz front side bus, and most people think that's nothing. But, you know, what are they doing? Playing Bejeweled on their lunch break and writing out memos in work? Maybe looking for um... nee... oh... are... en?

WOOTCAST

"Listener Mailbag" (4 comments)

THIS WEEK ON **wine.woot!**

AUG 07 [Laura Zahtila Vineyards Napa Valley Chardonnay - Three Pack](#)

TODAY ON **shirt.woot!**

AUG 07 [Sweet Dreams Are Made of These...](#)

SONY

The new Sony® VAIO® FW notebook with extra-wide screen.



High Definition. It's in our DNA.

LOGIN

Search in All **SHOP FOR BUSINESS PRODUCTS... CLICK BUSINESS ABOVE**

OFFERS ENDING SUNDAY

NEW OFFERS

CLICK HERE



^ Select Dell Desktops & Laptops
Up to \$400 OFF



^ Q-See 16 CH Security System
\$999.99 After \$1,000 OFF



^ Select Branded TV Hot Buys
Up to \$400 OFF

USA

Sign up to receive offers by e-mail:

your e-mail address

Earn 3% cash back when you buy gas.

APPLY NOW

Terms, conditions and restrictions apply.

THE TALES OF BEEDLE THE BARD

\$7.39 DELIVERED

[CLICK HERE](#)

THE SNOWBALL: Warren Buffett and the Business of Life

\$19.99

[CLICK HERE TO PRE-ORDER](#)

for your baby

[CLICK HERE](#)



^ Sunsetter Motorized Retractable Patio Awnings 30%



^ Hoover All Terrain SteamVac
\$199.99 After \$50 OFF



^ Emergency Food Kit
\$74.99 After \$10 OFF



^ 4-pc Hardside Luggage
\$149.99 After \$50 OFF



^ Garmin nuvi 770 4.3" GPS
\$399.99 After \$100 OFF



^ Saeco Deluxe Espresso
\$699.99 After \$200 OFF

Online Savings Account

3.75% APY*

on balances of \$10,000 or more

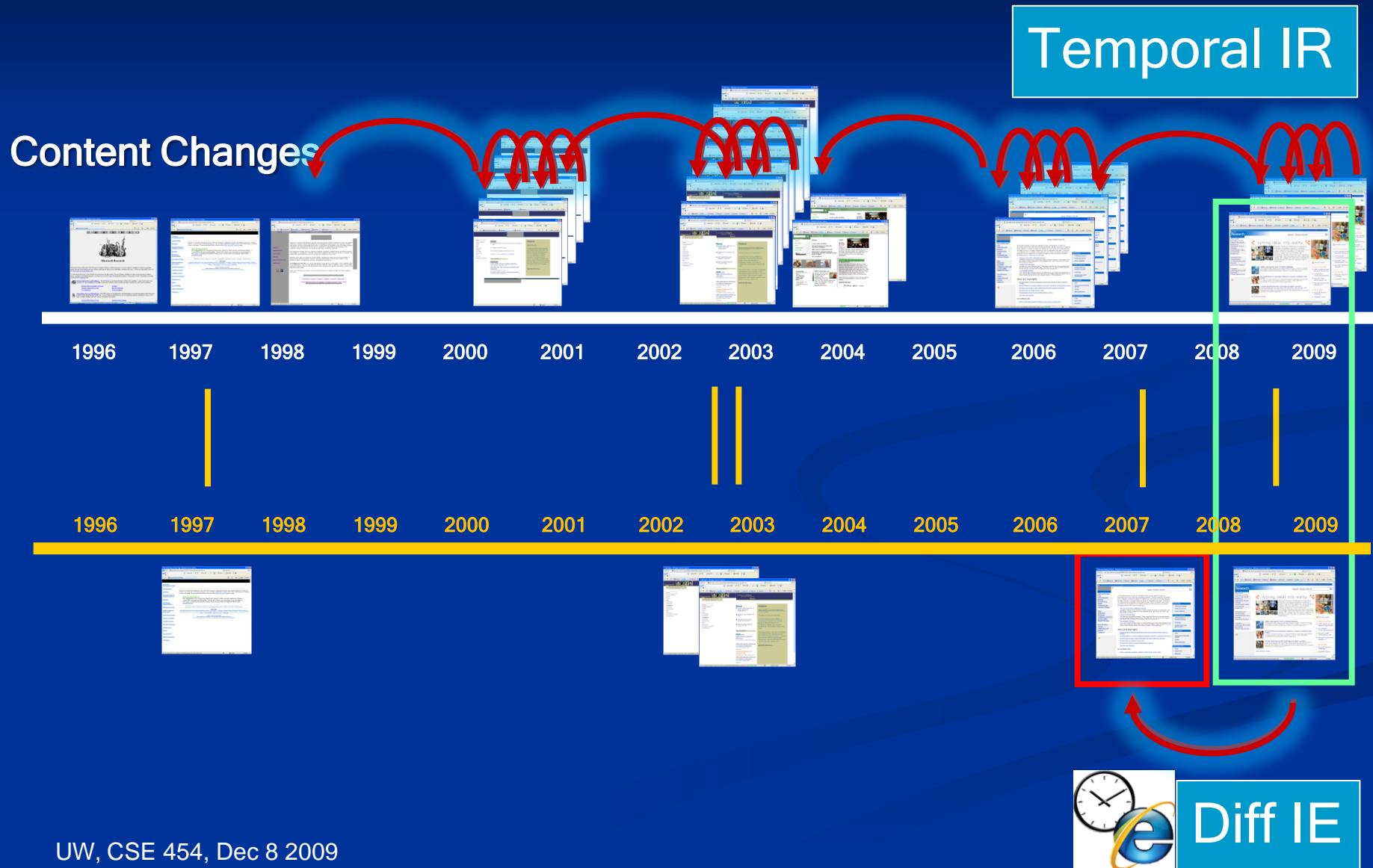
MEMBER FDIC

Give us your **Best Shot**

Dynamics of Information

- Characterizing change
 - Content changes over time
 - People re-visit and re-find
 - Relationships between content change and re-access
- Improving retrieval and understanding
 - Building support for understanding change (e.g., DiffIE)
 - Leveraging dynamics for improved retrieval

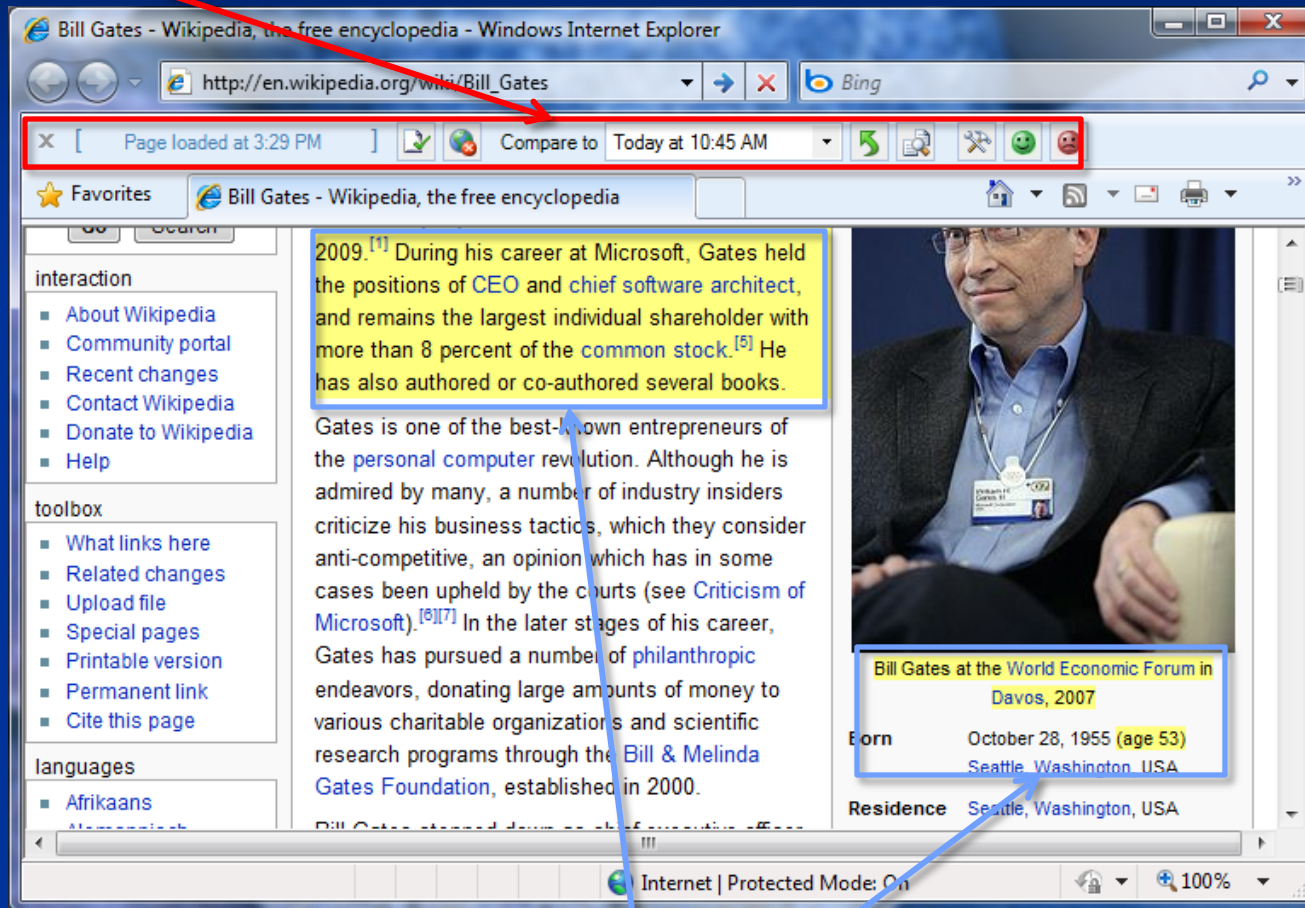
Building Support for Web Dynamics



DiffIE



DiffIE toolbar



Changes to page since
your last visit

Interesting Features of *DiffLE*



Examples of *DiffIE* in Action

Expected New Content

msnbc
Powered by Bing
MSN Home | Mail | More | Sign In
featuring TODAY | Nightly News | Dateline | Meet the Press | msnbc tv | NBC Sports

Top stories as of 10:14 a.m. Wed., Sept 30, 2009
New York, NY mostly cloudy, 57°F

World news

Thousands trapped in 2nd Pacific-area quake, official says

UPDATED At least 75 people dead as 7.6-magnitude temblor hits Indonesia; hundreds of houses reportedly damaged; panicked residents flee their homes. [Full story](#) | [Video: Western Sumatra struck](#)


Iraq resonates in claims over Iran nukes

NYT: While similarities between 2002, with its faulty IWMMD estimates, and 2009 are unmistakable, the differences are profound. [Full story](#)

China stalls action against Iran

Beijing's reluctance stems from oil, ideology. [Story](#)

[Video: World leaders meet on Iran nuclear threat](#)



Damaged boat (John Newton / AFP - Getty Images)

Samoans count cost after deadly tsunami

Undersea earthquake propels waves onto island nation and U.S. territory; toll expected to rise. [Story](#) | [Video: South Seas terror](#) | [Are you there? Send images, reports](#)

Sports

Some in Windy City hope 2016 bid is a bust

"I know it's going to cost u money somehow," one Chicago resident says. [Full story](#)

NEW Meet the Olympic contestants: Chicago, Tokyo, Rio, Madrid

[Video: Obamas and Winfrey make '16 pitch](#)

U.S. news

Guns to be allowed in Arizona bars | Vote

[Video: Summit to tackle distracted drivin](#)

advertisement | your ad here

Categories

- U.S. news
- World news
- Politics
- Business
- Sports
- Entertainment


Explore

More news | My news | Pinned

Other top stories

- Swine flu school closings could cost billions
- 1st lady in Europe for Olympics vote | [Video](#)
- GOP candidate: Too busy with family to vote
- Newsweek: What most suicide bombers have in common
- UPDATED** Church ordains sex offender as minister

Only on msnbc.com



Reuters
Slideshow: China

We do right by you and your money. ally Straightforward.

Mouse over tabs for details

- 9-Month No Penalty CD
- Great Rates

Monitor

The screenshot shows a Twitter profile for Susan Dumais. The top navigation bar includes 'Home', 'Profile', 'Find People', 'Settings', 'Help', and 'Sign out'. The profile header shows 'Name Susan Dumais' with statistics: 20 following, 12 followers, and 1 listed. Below this are sections for 'Tweets' (0) and 'Favorites'. The 'Following' section displays a grid of profile pictures for users being followed, including logos for CNN, Huffington Post, and NY Times. The main content area lists five followed users with their profile pictures, names, and recent tweets:

- pfromthenc** (Loren Terveen): "Just had my first Surly Furious... waited until *after* I finished working on the proposal tonight. [about 22 hours ago](#)"
- DARPA_News** (DARPA | Arlington, VA): "Team standings from DARPA Network Challenge posted on <http://bit.ly/5kdAZ1> [about 4 hours ago](#)"
- huffingtonpost** (HuffingtonPost.com): "GOP Senator: We Will Unanimously Oppose Newest Health Care Compromise <http://bit.ly/5xRBEi> [about 1 hour ago](#)"
- nytimes** (The New York Times | New York, NY): "Citi Races to Pay Back Bailout Aid <http://bit.ly/87E9Ry> [5 minutes ago](#)"
- dmrussell** (CA, USA): "@roblyons Google Office is at 1101 New York Ave (although the entrance is on I) #gtadc [about 13 hours ago](#)"

Unexpected Important Content

Welcome -- Please join us in Boston | SIGIR'09 - Windows Internet Explorer

http://www.sigir2009.org/

File Edit View Favorites Tools Help

msn Page loaded at 4:04 PM Compare to Today at 9:04 AM

Welcome -- Please join us in Boston | SIGIR'09

SIGIR 2009 Boston

The 32nd Annual ACM SIGIR Conference July 19-23 2009

Welcome -- Please join us in Boston

The SIGIR 2009 conference opens in a few days in Boston, Massachusetts, at the Sheraton Boston Hotel and Northeastern University. The conference is chock full of exciting events and registrations are strong and still growing. We are looking forward to an exciting week.

Conference registration site (updates or late)

NEU dorm registration (payments only)

Program

- Schedule
- Tutorials
- Workshops
- Industry track
- Keynotes
- Papers
- Posters
- Demonstrations

For attendees

- Visas
- Boston
- Venue
- Registration
 - student support
 - accommodations
- Red Sox

For contributors

- Important dates
- Mentoring (closed)
- Call for... (closed)
- Submitting (closed)

About us

- History
- Organizers

Recent news and upcoming deadlines

- Please join your colleagues by starting the conference with a free continental breakfast in the Sheraton Hotel, Back Bay A&B, from 7:00am to 8:20am on Monday, July 20.
- The conference banquet is currently full. Effective July 15th, reservations will be wait-listed for the banquet.
- Standard conference registration closes the night of Sunday, July 12th (Boston time). See below for more information.
- Registration for Northeastern University housing has closed. If you need to change your reservation, contact questions@sigir2009.org
- Three special tourism events have been added to the schedule: a panoramic view of the city on Sunday, a famous duck boat tour on Saturday, and a sunset harbor cruise on Wednesday. Look under "Boston" on the left for more information. Two of the events require sufficient interest to occur, so please fill out the questionnaire on that page.
- Industry Track registration is now available for students.

Arrival at the conference

- Tutorial and doctoral consortium registration is on the Northeastern campus, Sunday morning (the 19th). If you have registered for a tutorial or want to register for a tutorial, please get to Shillman Hall between 8:00 and 9:00 (the doctoral consortium starts at 8:30; morning tutorials start at 9:00). Volunteers and signs will provide direction from the hotel and the NEU dorms. [Here is a map.](#)
- Normal conference registration takes place at the Sheraton hotel, starting at 3:00pm on Sunday, July 19. Registration will be available throughout the conference if you are not arriving on Sunday.
- Workshop registrations may be made during normal registration periods or on the Northeastern campus on Thursday morning (the 23rd).
- If you have paid your fees in full in advance, your registration process should be very speedy.
- A free breakfast is available Monday morning (only) in the Sheraton.

For those who are driving, parking is available at the Sheraton and at Northeastern University. See [here](#)

Done Internet 100%


Serendipitous Encounters



One Microsoft Way
Redmond, WA 98052
(425) 421-9299
[Home](#) * [Work](#) * [Publications](#)

Organized by [Type](#) * [Subject](#) * [Date](#) * [Internal](#)

REFEREED JOURNAL ARTICLES:


 Jaime Teevan, Susan T. Dumais and Eric Horvitz. *Potential for Personalization*. To appear in the ACM Transactions on Computer-Human Interaction ([TOCHI](#)) special issue on Data Mining for Understanding User Needs, 2009. [[pdf](#)]

Jaime Teevan. *How People Recall, Recognize and Reuse Search Results*. ACM Transactions on Information Systems ([TOIS](#)) special issue on Keeping, Refinding, and Sharing Personal Information, 26(4), September 2008. [[pdf](#)]

Edward Cutrell, Susan T. Dumais and Jaime Teevan. *Searching to Eliminate Personal Information Management*. Communications of the ACM ([CACM](#)), Special Issue on Personal Information Management, 49(1), January 2006.

Diane Kelly and Jaime Teevan. *Implicit Feedback for Inferring User Preference: A Bibliography*. [SIGIR Forum](#), 37(2), 2003. [[pdf](#)]

REFEREED CONFERENCE PAPERS:

 Paul André, m. c. schraefel, Jaime Teevan and Susan T. Dumais. *Discovery is Never by Chance: Designing for (Un)Serendipity*. To appear in Proceedings of the 7th ACM Creativity and Cognition Conference ([CC '09](#)), Berkeley, CA, October 2009.

Jaime Teevan, Susan T. Dumais, Daniel J. Liebling and Richard Hughes. *Changing the Way People View Changes on the Web*. To appear in Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology ([UIST '09](#)), Victoria, BC, October 2009. [[pdf](#)]

Eytan Adar, Jaime Teevan and Susan T. Dumais. *Resonance on the Web: Web Dynamics and Revisitation Patterns*. In Proceedings of ACM Conference on Human Factors in Computing Systems ([CHI '09](#)), Boston, MA, April 2009. [[pdf](#)]

Jaime Teevan, Edward Cutrell, Daniel Fisher, Steven M. Drucker, Gonzalo Ramos, Paul André and Chang Hu. *Visual Snippets*:

Unexpected Unimportant Content

automation unit test **visual studio** vsts
Web WTT **XML**

What is a Tag Cloud? Browse All Tools

Most Downloaded


7 | 30 | All days

-  **ShowSenderPhoto** - Automatically displays a photo of every Microsoft Employee that sends you mail.
1,676 downloads • Published: 01/15/2008
★★★★★
-  **Bugger** - Bugger is a personal bug portal, designed to make your bug list easier to manage and a bit more fun.
1,093 downloads • Published: 06/13/2006
★★★★★
-  **Magic Mouse** - A "virtual" keyboard/mouse to control multiple machines, control all machines concurrently, copy/paste drag/drop text/media/files across machines.
960 downloads • Published: 06/12/2008
★★★★★
-  **Outlook Conference Room Finder** - Find an available conference room in a specific floor and building at any time from within the new appointment dialog in Outlook!
833 downloads • Published: 03/18/2005
★★★★★
-  **Source Depot Browser (SDB)** - GUI client for Source Depot
613 downloads • Published: 02/08/2005
★★★★★
-  **RDCMan** - RDCMan manages multiple remote desktop connections.
577 downloads • Published: 02/08/2005
★★★★★
-  **RichCopy 4.0.211** - RichCopy is designed to help people who copy huge number of files between storage devices

New Releases

-  **!exploitable** - A Windows Debugging Engine extension for auto-triaging of exceptions based on their security imp... more
Released: 04/15/2009
★★★★★
-  **Biztalk Msgbox Viewer** - Health Check tool for Biztalk Databases
Released: 04/15/2009
★★★★★
-  **Send-a-Bug** - Send-a-Bug allows you to easily send Product Studio bugs in Outlook. It automatically adds the ... more
Released: 04/15/2009
★★★★★
-  **Fix Build Order** - FixBuildOrder ensures the build order is correct for a SLN file that contains csproj or vbproj... more
Released: 04/15/2009
★★★★★
-  **Machine Management Tool (MMT)** - The Machine Management Tool is a simple C# application written to manage machines in the a lab e... more
Released: 04/14/2009
★★★★★

Top Rated

-  **Clip** - Command line Windows clipboard operations
5.0 Rating: [7] • Published: 02/08/2005
★★★★★
-  **ERR** - command-line-based error lookup tool
5.0 Rating: [19] • Published: 02/08/2005
★★★★★
-  **Shist** - Source history browser and function tracking tool for Source Depot and SLM

Understand Page Dynamics

The image shows a screenshot of a Bing search results page for the query "weather 98005". The page layout includes a navigation bar at the top with links for Web, Images, Videos, Shopping, News, Maps, More, MSN, and Hotmail. The Bing logo is prominently displayed on the left. The search bar contains the text "weather 98005" and a search icon. Below the search bar, the results are categorized under "ALL RESULTS" with a sub-section for "Weather". The main content area displays the current weather for "98005" as "Today 54°F (°C)" with a wind speed of 6 mph and 71% humidity. It also shows a 5-day forecast for Wednesday, Thursday, Friday, and Saturday. The page includes several search suggestions on the left, such as "Weather 98052", "Movies 98005", "xooost.com", "fox news", and "msn". The search history section lists "weather 98005" and "xooost.com". The main results list includes a "10 Day Weather Forecast for Bellevue, WA (98005) - weather.com" and "Bellevue, WA (98006) Weather Forecast - Weatherlet.com".

Web Images Videos Shopping News Maps More MSN Hotmail

bing™ weather 98005

ALL RESULTS 1-10 of 22,100 results - [Advanced](#)

Weather

RELATED SEARCHES

- Weather 98052
- Movies 98005

SEARCH HISTORY

- weather 98005
- xooost.com
- fox news
- msn
- [See all](#)
- [Clear all](#) | [Turn off](#)

ALL RESULTS: 1-10 of 22,100 results - [Advanced](#)

Weather in 98005

Today **54°F** (°C)
Wind: 6 mph
Humidity: 71%
61° / 50°

Wed 59° / 51°

Thu 61° / 50°

Fri 58° / 39°

Sat 60° / 37°

[10 Day Forecast](#) - [Hourly Forecast](#) - [Weather Maps](#) - Data provided by WDT

[10 Day Weather Forecast for Bellevue, WA \(98005\) - weather.com](#)
10 day **weather** forecast for Bellevue, WA (98005) with **weather** conditions, high and low temperatures, and precipitation.
[www.weather.com/weather/tenday/98005?day=0&vert=WeatherCity](#) - [Cached page](#)

[Bellevue, WA \(98006\) Weather Forecast - Weatherlet.com](#)
Bellevue, WA (98006) **Weather** Forecast - Weatherlet ... Auburn, WA (98002)
Weather Forecast Federal Way, WA (98003) **Weather** Forecast Bellevue, WA (98005) **Weather** Forecast
[www.weatherlet.com/?locid=98006&cityname=Bellevue%2C+today...](#) - [Cached page](#)

[Bellevue, Washington \(98005\) Conditions & Forecast : Weather ...](#)
Weather Underground provides **weather** information for worldwide locations, including current conditions, hourly forecasts, radar and satellite maps. Specialized **weather** products ...

Attend to Activity

13: Eastside Mothers & More Expo. See Chapter News and Events for more information. Yeah, baby!!

Eastside Mothers & More Message Board | Discussion Forums | Sequencing

1] 2 3 ... 9

Mark Topics as Read for this Board | Notify of replies | Start new topic | Post

Subject	Started by	Replies	Views	Last post ▼
Small Business « 1 2 »	Della Dawn	21	206	February 17, 2009, 09:09 PM by Lori Poythress
Defining Sequencing	Miti Morris	0	93	November 10, 2005, 09:09 PM by Miti Morris
Back to work Part Time! <i>new</i>	Sue Monahan	8	43	Today at 05:36:34 PM by Lynn Woods
Aldersgate Christian Preschool - HIRING <i>new</i>	Heather Ross	0	17	Today at 03:01:42 PM by Heather Ross
Do you know anyone in Italy or Sweden?	Virginia Menting	3	70	Today at 07:00:21 AM by Virginia Menting
Got an interview and I'm freaking out... « 1 2 ... 6 7 »	Trina Sooy	94	962	June 08, 2009, 06:54:00 PM by Kyla Carlson
Academia or Industry? Finding the Right Fit	Jaime Teevan	7	105	May 26, 2009, 09:26:00 PM by Trina Sooy
Any members work at Amazon.com?	Kathryn Pizzo	10	188	May 08, 2009, 08:20:00 PM by JJ Salley
Time Magazine on SAHMs getting back into workforce	Kathryn Pizzo	2	75	April 17, 2009, 07:34:00 PM by Erin Proctor
Freelance Resource	Kathryn Pizzo	0	41	April 01, 2009, 06:54:00 PM by Kathryn Pizzo
Graphic designer or Admin?	Heather Ross	0	45	March 20, 2009, 07:55:00 PM by Heather Ross

Edit



Installing DiffIE

- System requirements: Windows XP or Vista, Internet Explorer (32-bit) version 7, Outlook 2007.
- Installation location: `\\asibuild\Drops\CLUES\latest\DiffIE\DiffIE.msi`
Note: Save this MSI to your machine and stop IE before installing DiffIE.
- Announcement alias: If you download DiffIE, please [join](#) the "DiffIE Announcements" distribution list so that we can keep you informed about new versions and other important changes.
- Feedback alias: diffiefb@microsoft.com

Using DiffIE

Once installed, you should see the DiffIE toolbar in IE (shown below on Vista):



Expected



Expected
New Content



Monitor



Unexpected
Important Content

33: Outside Mothers & More Expo... Use Chapter News and Events for more information. Yeah, baby!

Click & Drag Message Board | Discussion Forums | Responses

1 2 3 ... 9

Subject	Started by	Replies	Views	Last post
Small Business + ...	Della Davis	21	216	Mon Feb 23, 2009, 9:04 AM
Confusing Sequencing	MIS Morris	0	83	Mon Feb 23, 2009, 9:04 AM
Back to work Part II	Bob Moulton	5	43	Mon Feb 23, 2009, 9:04 AM
Aldergate Christian Preschool - HIKING	Heather Reese	0	17	Mon Feb 23, 2009, 9:04 AM
do you know anyone in Italy or Sweden?	Virginia Manning	3	30	Mon Feb 23, 2009, 9:04 AM
Get an interview and the recruiting ...	Thom Dyer	64	662	Mon Feb 23, 2009, 9:04 AM
Academics or Industry? Finding the Right Fit	Janet Taylor	7	105	Mon Feb 23, 2009, 9:04 AM
Any members work at Amazon.com?	Kathryn Hesse	10	118	Mon Feb 23, 2009, 9:04 AM
Time Magazine on 9/11 - getting back into work	Kathryn Hesse	1	75	Mon Feb 23, 2009, 9:04 AM
Frederick Rosecrance	Kathryn Hesse	0	41	Mon Feb 23, 2009, 9:04 AM
Congrat designer or Admin?	Heather Reese	0	45	Mon Feb 23, 2009, 9:04 AM

Attend to Activity



Serendipitous
Encounter



Unexpected
Unimportant Content

Unexpected



Edit



Understand
Page Dynamics

Methods for Studying *DiffIE*

■ Feedback buttons



■ Survey

- Prior to installation
- After a month of use

Response	Always	Often	Sometimes	Rarely	Never
News pages	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Software tools, forums, newsgroups	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Company homepages	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Personal homepages of people you know	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pages with product information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reference pages (dictionaries, other pages)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Weblogs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Blog you read	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Search engine results	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In situ

Representative

Experience

Longitudinal

■ Logging




- URLs visited
- Amount of change when revisited






■ Experience interview



People Revisit More

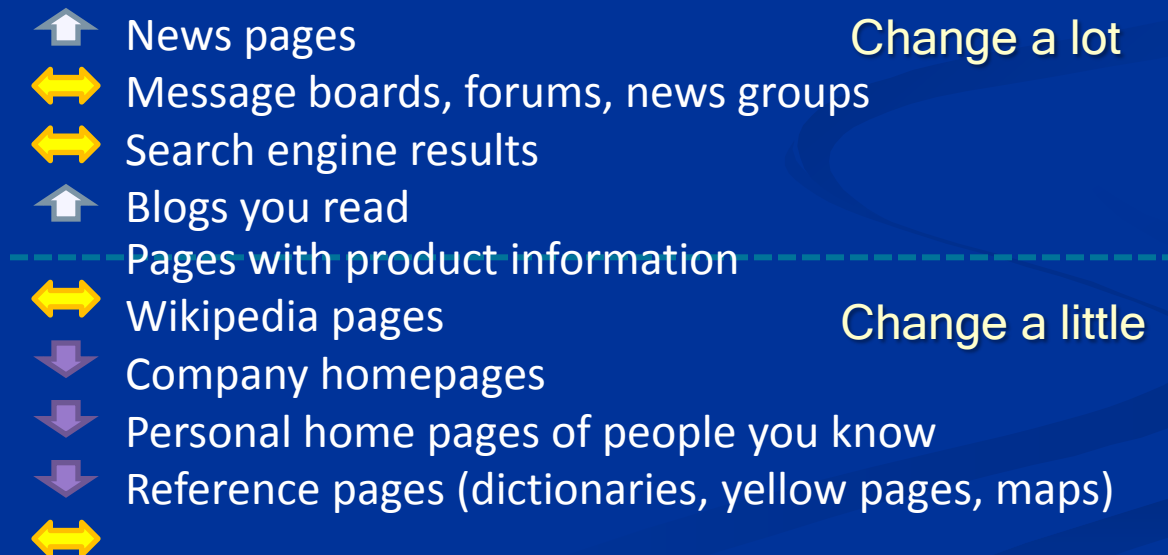
- Perception of revisitation remains constant
 - How often do you revisit? 
 - How often are revisits to view new content? 
- Actual revisitation increases
 - First week: 39.4% of visits are revisits
 - Last week: 45.0% of visits are revisits 
- Why are people revisiting more with *DiffIE*?

Revisited Pages Change More

- Perception of change increases
 - What proportion of pages change regularly?  8%
 - How often do you notice unexpected change?  17%
- Amount of change seen increases
 - First week: 21.5% revisits changed by 6.2%  51+%
 - Last week: 32.4% revisits changed by 9.5%
- *DiffIE* is driving visits to changed pages

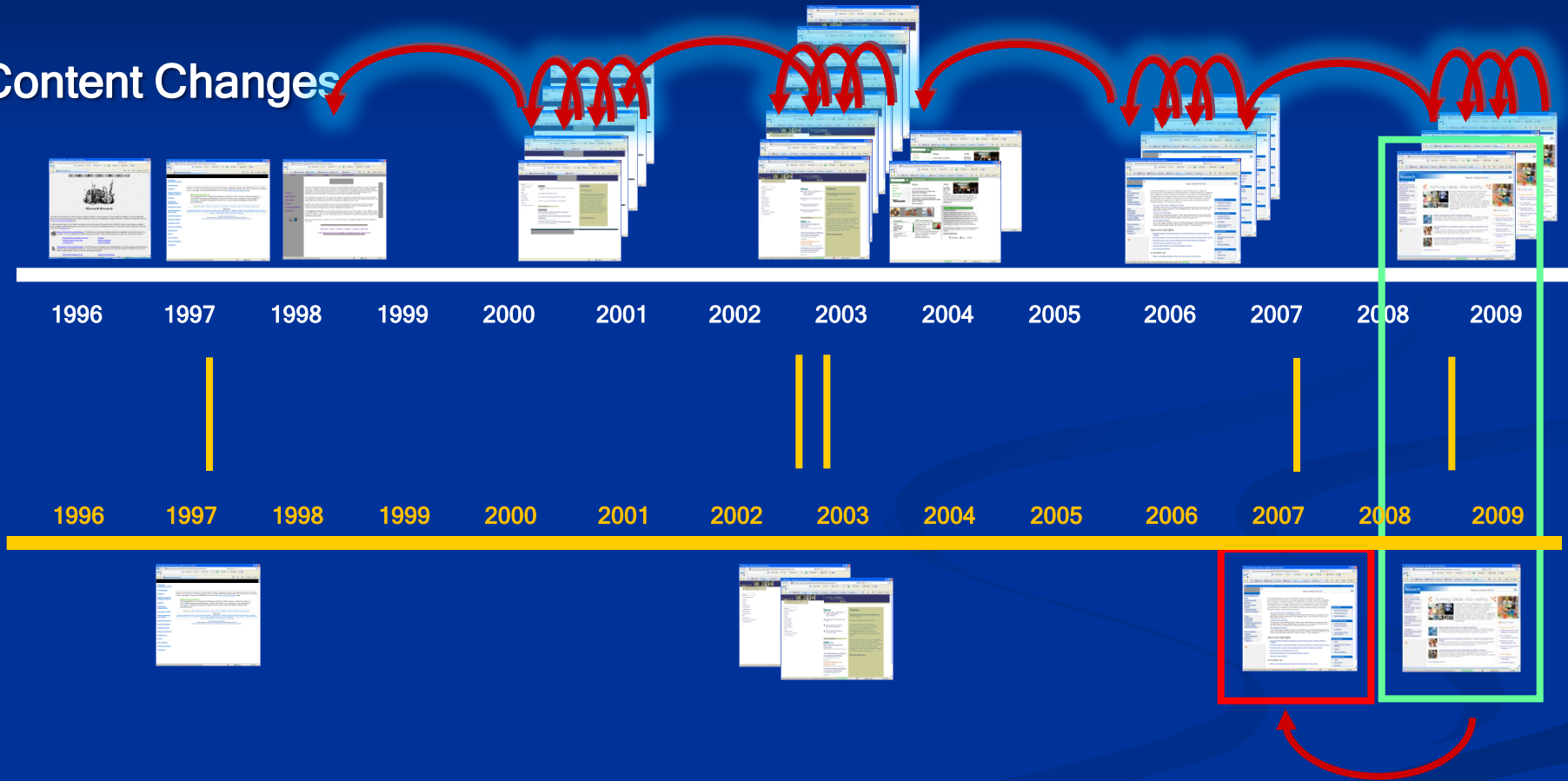
Change by Page Type

- Perceptions of change reinforced
- Pages that change a lot → change more
- Pages that change a little → change less



Leveraging Dynamics for Retrieval

Content Changes

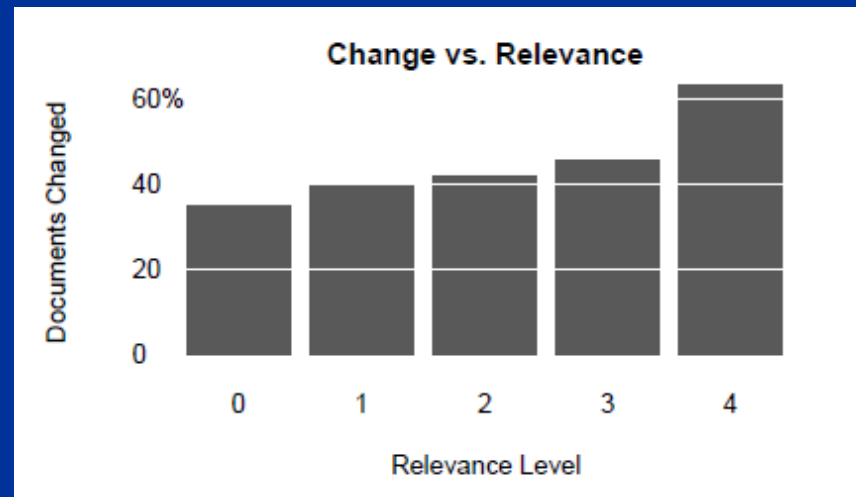


Temporal Retrieval Models

- Current IR algorithms look only at a single snapshot of a page
- But, Web pages change over time
- Can we can leverage this for improved retrieval?
 - Pages have different *rates of change*
 - Different priors (using change vs. link structure)
 - Terms have *different longevity (staying power)*
 - Some are always on the page; some transient

Temporal Retrieval Models

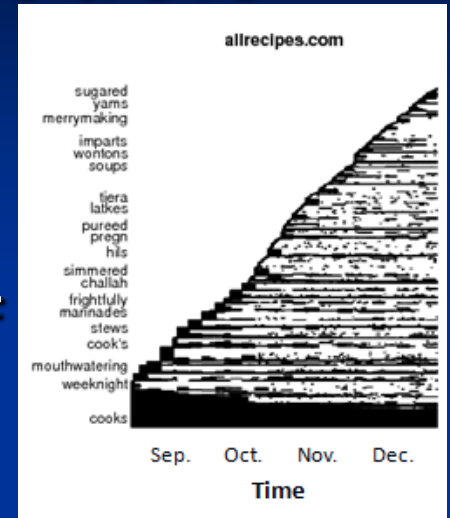
- Page change is related to relevance judgments
 - Human relevance judgments
 - 5 point scale - Bad/Fair/Good/Excellent/Perfect
 - Rate of Change -- 30% Bad pages; 60% Perfect pages



- Use change rate as a document prior (vs. priors based on links, Page Rank)

Temporal Retrieval Models

- Terms patterns vary over time
- Represent a document as a mixture of terms with different “staying power”
 - Long, Medium, Short



$$P(Q | D) = \lambda_L P(Q | D_L) + \lambda_M P(Q | D_M) + \lambda_S P(Q | D_S)$$

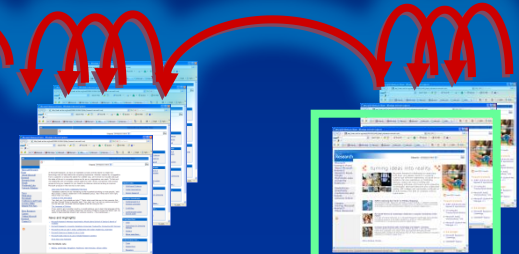
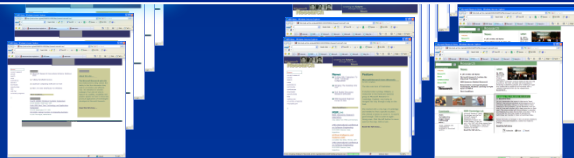
- Improves retrieval accuracy

Temporal Retrieval Models nDCG@5 (2k Navigation Queries)			
	Single Slice	Dynamic	
Term Dynamics	0.21	0.22	4.4%
TD + Change Prior	0.24	0.25	
	13.1%	11.2%	19.0%

Summary

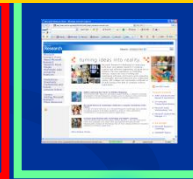
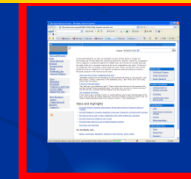
Temporal IR:
Leverage change for improved IR

Web content changes: page-level, term-level



Relating revisitation and change allows us to

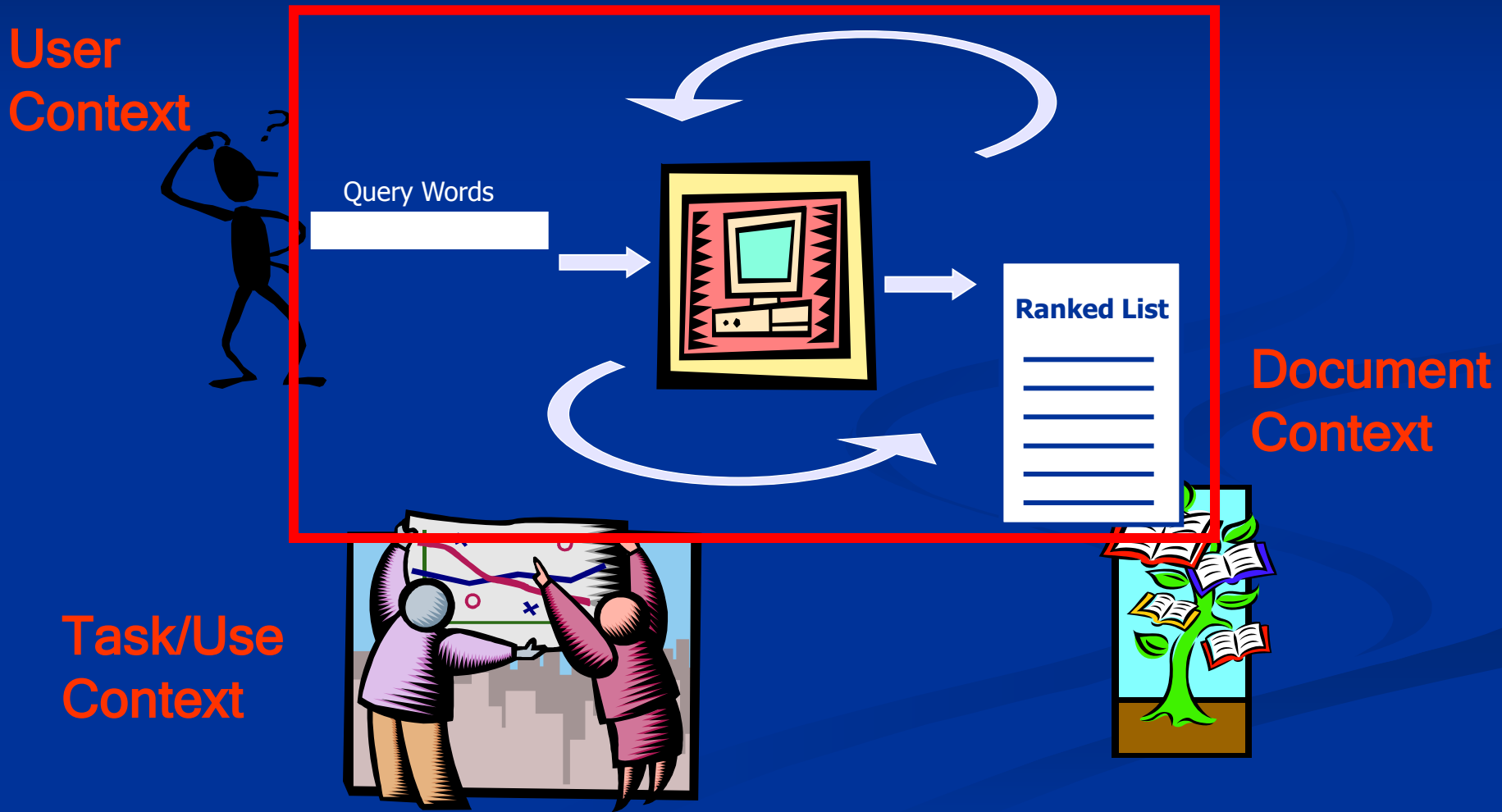
- Identify pages for which change is important
- Identify interesting components within a page



People revisit and re-find Web content

DiffIE: Support influences interaction and understanding

Think Outside the (Search) Box



Thank You !

- Questions/Comments ...
- More info,
<http://research.microsoft.com/~sdumais>

References

Change and Revisitation:

- Adar, Teevan, Dumais & Elsas. The Web changes everything: Understanding the dynamics of Web Content. WSDM'09.
- Adar, Teevan & Dumais. Large scale analysis of Web revisitation patterns. CHI'08.
- Teevan, Adar, Jones & Potts. Information re-retrieval: Repeat queries in Yahoo's logs. SIGIR'07.
- Tyler & Teevan. Large scale query log analysis of re-finding. WSDM'10.
- Adar, Teevan & Dumais. Resonance on the Web: Web dynamics and revisitation patterns. CHI '09.

DiffIE and Temporal IR:

- Teevan, Dumais, Liebling & Hughes. Changing how people view changes on the Web. UIST '09.
- Teevan, Dumais & Liebling. A longitudinal study of how highlighting Web content change affects people's Web interactions. CHI '10.
- Elsas & Dumais. Leveraging temporal dynamics of document content in relevance ranking. WSDM '10.