



THE WEB CHANGES EVERYTHING

Jaime Teevan, Microsoft Research

Microsoft Research - Turning Ideas into Reality - Windows Internet Explorer

http://research.microsoft.com/

Microsoft Research

Search Microsoft Research

Home | Our Research | Collaboration | Careers | About Us | News | Media Resources | Events | Community

Project Tuva's Feynman Lectures

Gates' Gift to Lifelong Learning



Feature

Cryptog
The Cry
Mathema
India has
bring to
fundame

http://research.microsoft.com/c/104

Internet | Protected Mode: On

100%

Microsoft Research Home - Windows Internet Explorer

http://research.microsoft.com/ Bing

★ Favorites Microsoft Research Home

MSR HOME MSR SEARCH MSR PUBS MSR INFO MSR TOPICS MS HOME



Microsoft Research

Welcome to Microsoft Research! This Web Space is designed to inform you about our group. If you are looking for something (or someone) in particular, please select the Search button at the top or bottom of this page for quick access to information contained in this space. To learn more about MSR, check out the [Microsoft Research Introduction](#) page.

To help structure this space we have divided our research areas into three primary "Research Themes" and listed the various research groups according to theme. In addition to individual group pages, each theme has an overview page with thumbnail descriptions of each group's activities. Independent researchers are listed on the [Colleagues](#) page.

Internet | Protected Mode: On 100%

Bill Gates - Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Bill_Gates

Bill Gates - Wikipedia, the free encyclopedia

WIKIPEDIA

The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link
- Cite this page

languages

- Afrikaans
- Alemannisch

Bill Gates

From Wikipedia, the free encyclopedia

For other persons named Bill Gates, see [Bill Gates \(disambiguation\)](#).

William Henry "Bill" Gates III (born October 28, 1955)^[2] is an [American business magnate](#), [philanthropist](#), and [chairman](#)^[3] of [Microsoft](#), the software company he founded with [Paul Allen](#). He is ranked consistently one of the [world's wealthiest people](#)^[4] and the wealthiest overall as of 2009.^[1] During his career at Microsoft, Gates held the positions of [CEO](#) and [chief software architect](#), and remains the largest individual shareholder with more than 8 percent of the [common stock](#).^[5] He has also authored or co-authored several books.

Gates is one of the best-known entrepreneurs of the [personal computer](#) revolution. Although he is admired by many, a number of industry insiders criticize his business tactics, which they consider anti-competitive, an opinion which has in some cases been upheld by the courts (see [Criticism of Microsoft](#)).^{[6][7]} In the later stages of his career, Gates has pursued a number of [philanthropic](#) endeavors, donating large amounts of money to various charitable organizations and scientific research programs through the [Bill & Melinda Gates Foundation](#), established in 2000.

Bill Gates stepped down as chief executive officer

Bill Gates



Bill Gates at the [World Economic Forum](#) in [Davos](#), 2007

Born	October 28, 1955 (age 53) Seattle, Washington, USA
Residence	Seattle, Washington, USA
Alma	Harvard University (drooped out)

Internet | Protected Mode: On

100%

Bill Gates - Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Bill_Gates

Bill Gates - Wikipedia, the free encyclopedia

WIKIPEDIA

The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link
- Cite this page

languages

- Afrikaans
- Alemannisch

Bill Gates

For other persons named Bill Gates, see [Bill Gates \(disambiguation\)](#).

William Henry Gates III (born October 28, 1955)^[2] is an [American business magnate](#), [philanthropist](#), and [chairman](#)^[3] of Microsoft, the software company he founded with [Paul Allen](#). During his career at Microsoft, Gates held the positions of [CEO](#) and [chief software architect](#), and remains the largest individual shareholder with more than 9 percent of the [common stock](#).^[5] He has also authored or co-authored several books.

Gates is one of the best-known entrepreneurs of the [personal computer](#) revolution. Although he is admired by many, a large number of industry insiders criticize his business tactics, which they consider anti-competitive, an opinion which has in some cases been upheld by the courts.^{[6][7]} In the later stages of his career, Gates has pursued a number of [philanthropic](#) endeavors, donating large amounts of money to various charitable organizations and scientific research programs through the [Bill & Melinda Gates Foundation](#), established in 2000.

Bill Gates stepped down as chief executive officer of Microsoft in January, 2000. He remained as chairman and created the position of chief software architect. In June, 2006, Gates

William Henry Gates III



Bill Gates

Born	October 28, 1955 (age 52) Seattle, Washington, USA
Residence	Seattle, Washington, USA
<i>Alma mater</i>	Harvard University (dropped out in 1975)
<i>Occupation</i>	Chairman of Microsoft

Internet | Protected Mode: On

100%

Bill Gates - Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Bill_Gates

Bill Gates - Wikipedia, the free encyclopedia

WIKIPEDIA

The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link
- Cite this page

languages

- Afrikaans
- Alemannisch

Bill Gates

From Wikipedia, the free encyclopedia

For other persons named Bill Gates, see [Bill Gates \(disambiguation\)](#).

William Henry "Bill" Gates III (born October 28, 1955)^[2] is an [American business magnate](#), [philanthropist](#), and [chairman](#)^[3] of Microsoft, the software company he founded with [Paul Allen](#). **He is ranked consistently one of the [world's wealthiest people](#)^[4] and the wealthiest overall as of 2009.**^[1] During his career at Microsoft, Gates held the positions of [CEO](#) and [chief software architect](#), and remains the largest individual shareholder with more than **9.8** percent of the [common stock](#).^[5] He has also authored or co-authored several books.

Gates is one of the best-known entrepreneurs of the [personal computer](#) revolution. Although he is admired by many, a **large** number of industry insiders criticize his business tactics, which they consider anti-competitive, an opinion which has in some cases been upheld by the courts (see [Criticism of Microsoft](#)).^{[6][7]} In the later stages of his career, Gates has pursued a number of [philanthropic](#) endeavors, donating large amounts of money to various charitable organizations and scientific research programs through the [Bill & Melinda Gates Foundation](#), established in 2000.

Bill Gates stepped down as chief executive officer

William Henry Bill Gates III



Bill Gates at the [World Economic Forum in Davos, 2007](#)

Born	October 28, 1955 (age 53 52) Seattle, Washington, USA
Residence	Seattle, Washington, USA
Alma	Harvard University (dropped out)

Internet | Protected Mode: On

100%

Bill Gates - Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Bill_Gates

Bill Gates - Wikipedia, the free encyclopedia

WIKIPEDIA
The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

Bill Gates

From Wikipedia, the free encyclopedia

For other persons named *Bill Gates*, see *Bill Gates (disambiguation)*.

William Henry "Bill" Gates III (born October 28, 1955)^[2] is an American business magnate, philanthropist, and chairman^[3] of Microsoft, the software company he founded with Paul Allen. He is ranked consistently one of the world's wealthiest people^[4] and the wealthiest overall as of 2009.^[1] During his career at Microsoft, Gates held CEO and chief software architect, largest individual shareholder with 12.1 percent of the common stock.^[5] He has also written or co-authored several books.

Gates is one of the best-known entrepreneurs of the personal computer revolution. Although he is often cited as a **large** number of industry best-selling authors on business tactics, which they have been largely upheld by the courts (see *Microsoft v. AT&T* and *Microsoft v. Shut-Down*).^{[6][7]} In the later stages of his career, he has pursued a number of philanthropic endeavors, donating large amounts of money to various charitable organizations and programs through the **Bill & Melinda Gates Foundation**, established in 2000.

Gates stepped down as chief executive officer of Microsoft in 2014.

William-Henry Bill Gates III



Bill Gates at the World Economic Forum in Davos, 2007

Born October 28, 1955 (age 53 52)
Seattle, Washington, USA

Residence Seattle, Washington, USA

Alma Harvard University (dropped out)



The Web Changes Everything

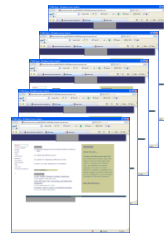
Content Changes



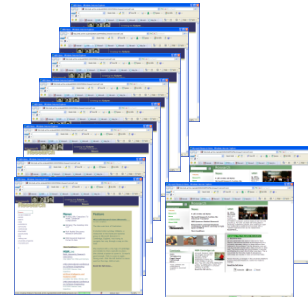
January

February

March



April



May

June



July

August



September

The

Content Ch

January

September

Bill Gates - Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Bill_Gates

Bill Gates - Wikipedia, the free encyclopedia

WIKIPEDIA
The Free Encyclopedia

navigation

- Main page
- Contents
- Featured content
- Current events
- Random article

search

Go Search

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link
- Cite this page

languages

- Afrikaans
- Alemannisch

Bill Gates

From Wikipedia, the free encyclopedia

For other persons named *Bill Gates*, see *Bill Gates (disambiguation)*.

William Henry "Bill" Gates III (born October 28, 1955)^[2] is an American business magnate, philanthropist, and chairman^[3] of Microsoft, the software company he founded with Paul Allen. He is ranked consistently one of the world's wealthiest people^[4] and the wealthiest overall as of 2009.^[1] During his career at Microsoft, Gates held the positions of CEO and chief software architect, and remains the largest individual shareholder with more than 9.8 percent of the common stock.^[5] He has also authored or co-authored several books.

Gates is one of the best-known entrepreneurs of the personal computer revolution. Although he is admired by many, a large number of industry insiders criticize his business tactics, which they consider anti-competitive, an opinion which has in some cases been upheld by the courts (see Criticism of Microsoft).^{[6][7]} In the later stages of his career, Gates has pursued a number of philanthropic endeavors, donating large amounts of money to various charitable organizations and scientific research programs through the **Bill & Melinda Gates Foundation**, established in 2000.

Bill Gates stepped down as chief executive officer

William Henry Bill Gates III



Bill Gates at the World Economic Forum in Davos, 2007

Born October 28, 1955 (age 53 52)
Seattle, Washington, USA

Residence Seattle, Washington, USA

Alma Harvard University (dropped out)

Internet | Protected Mode: On

100%



The Web Changes Everything

Content Changes



January

February

March

April

May

June

July

August

September

January

February

March

April

May

June

July

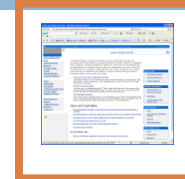
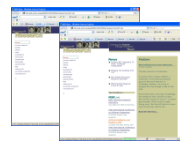
August

September

People Revisit

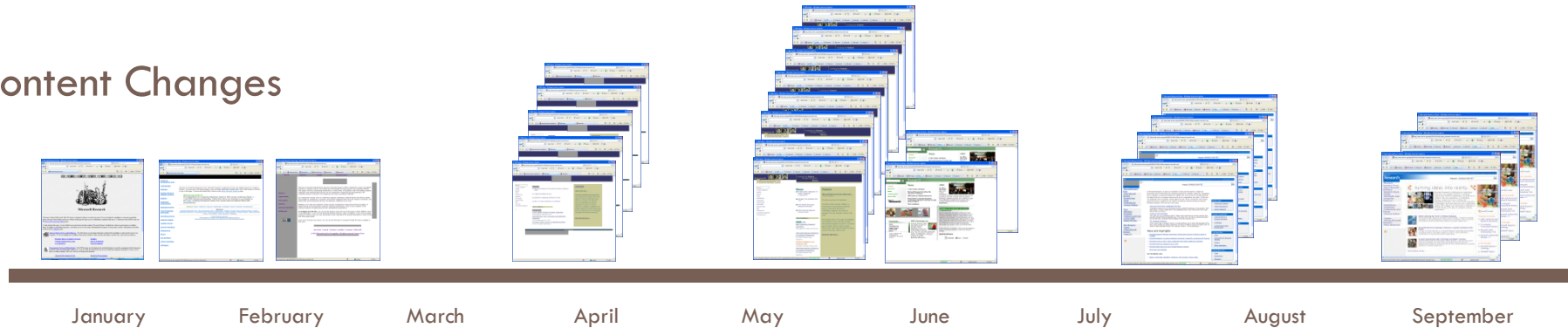
Today's tools focus on the present

But there's so much more information available!



The Web Changes Everything

Content Changes

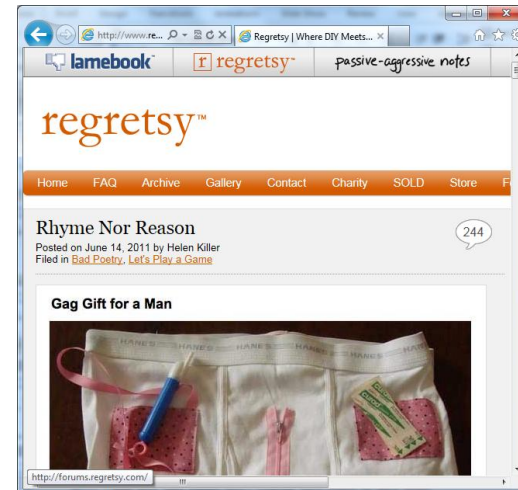


- Large scale Web crawl over time
 - ▣ Revisited pages
 - 55,000 pages crawled hourly for 18+ months
 - ▣ Judged pages (relevance to a query)
 - 6 million pages crawled every two days for 6 months

Measuring Web Page Change

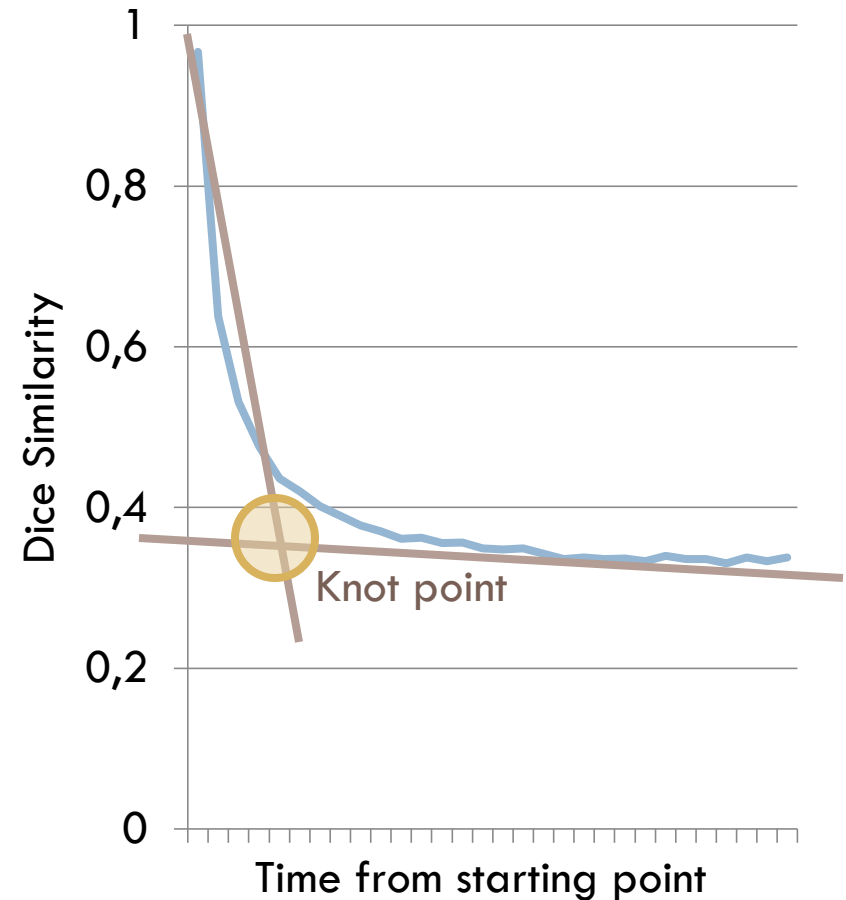
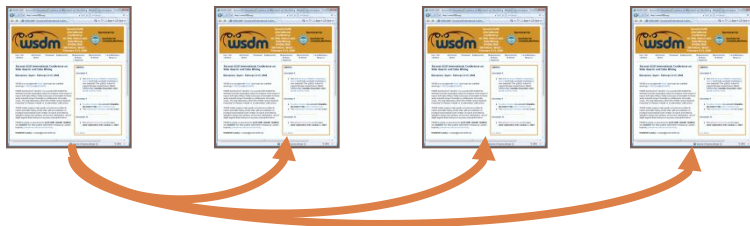
- Summary metrics
 - ▣ Number of changes
 - ▣ Time between changes
 - ▣ Amount of change

*Top level pages change by more and faster than pages with long URLs.
.edu and .gov pages do not change by very much or very often
News pages change quickly, but not as drastically as other types of pages*



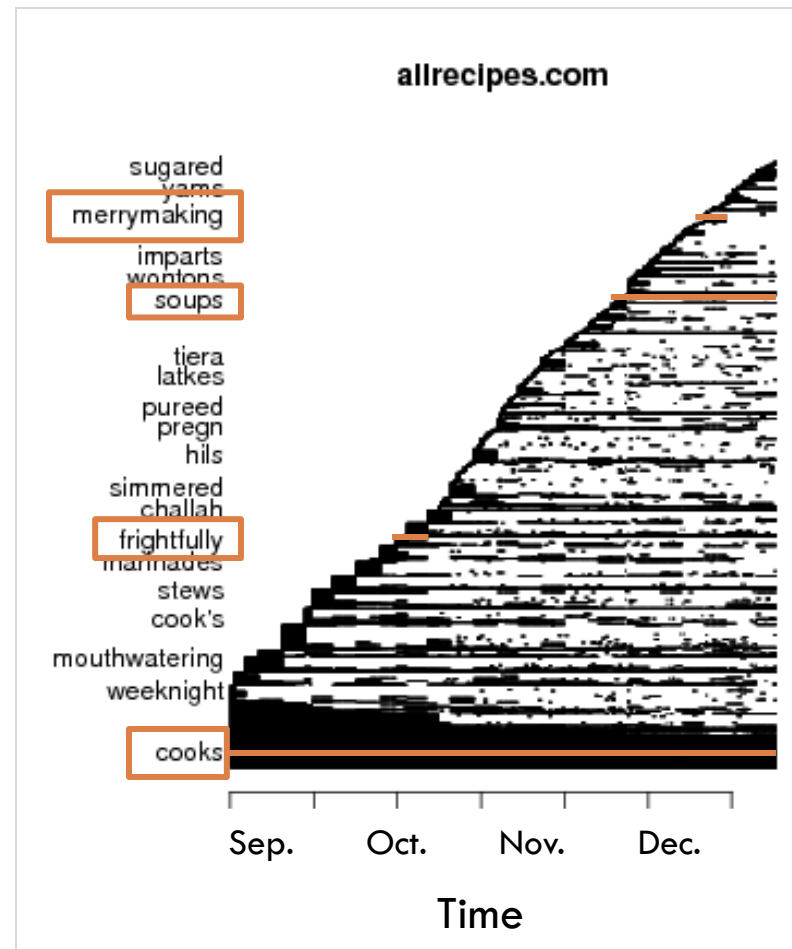
Measuring Web Page Change

- 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
 - frightfully
 - merrymaking
 - ingredient
 - latkes
 - ▣ Staying power in page



Accounting for Web Dynamics

- Avoid problems caused by change
 - ▣ Caching, archiving, crawling
- Use change to our advantage
 - ▣ Ranking
 - Match term's staying power to query intent
 - ▣ Snippet generation

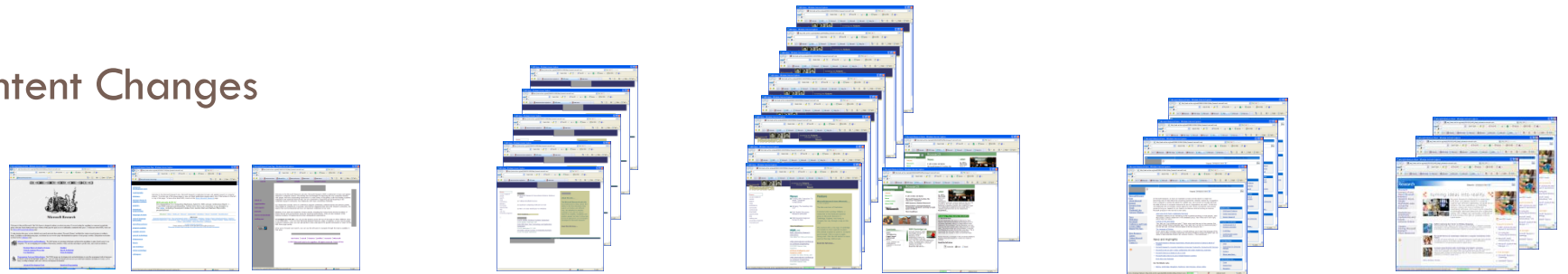
[Tom Bosley - Wikipedia, the free encyclopedia](#)

Bosley died at 4:00 a.m. of heart failure on October 19, 2010, at a hospital near his home in Palm Springs, California. ... His agent, Sheryl Abrams, said **Bosley** had been battling lung cancer.

en.wikipedia.org/wiki/tom_bosley

Revisitation on the Web

Content Changes



January

February

March

April

May

June

July

August

September

January

February

March

April

May

June

July

August

September

People Revisit

What's the last Web page you visited?

Revisitation on the Web

□ Revisitation patterns

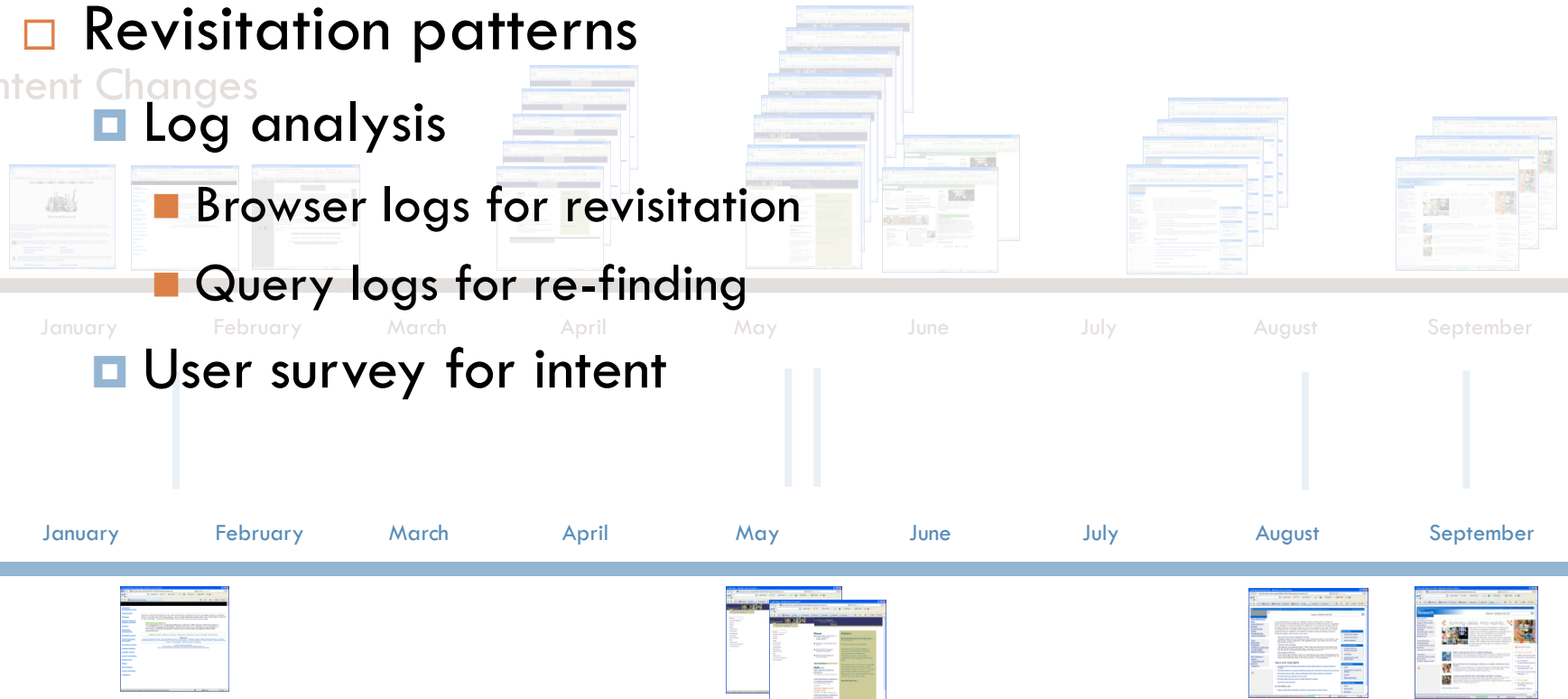
Content Changes

□ Log analysis

■ Browser logs for revisitation

■ Query logs for re-finding

□ User survey for intent

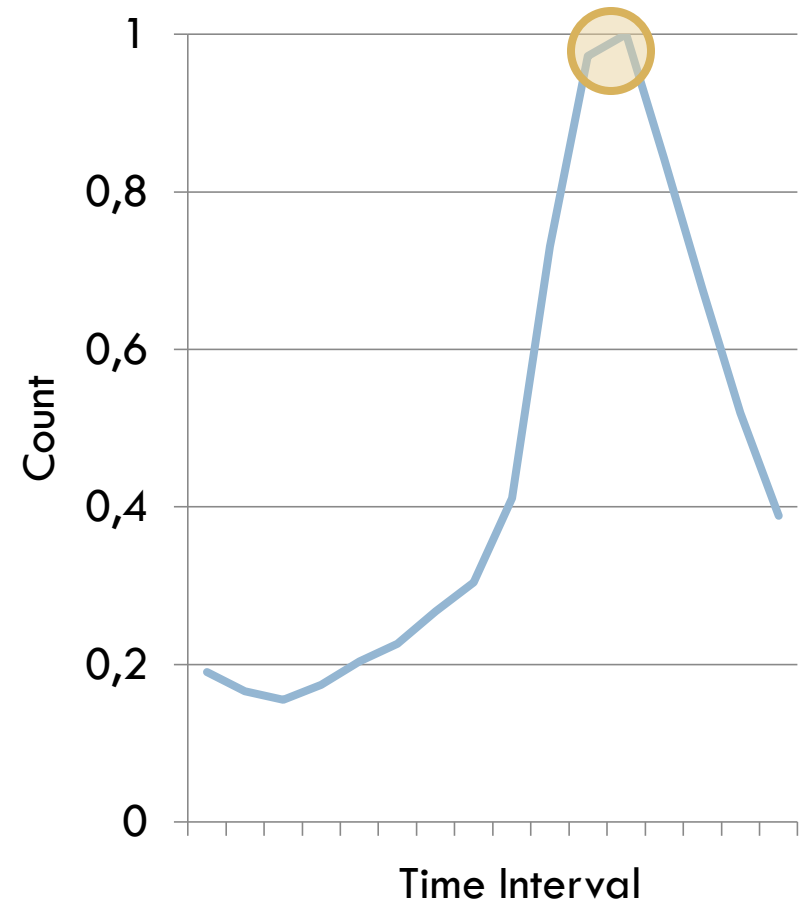


People Revisit

What's the last Web page you visited?

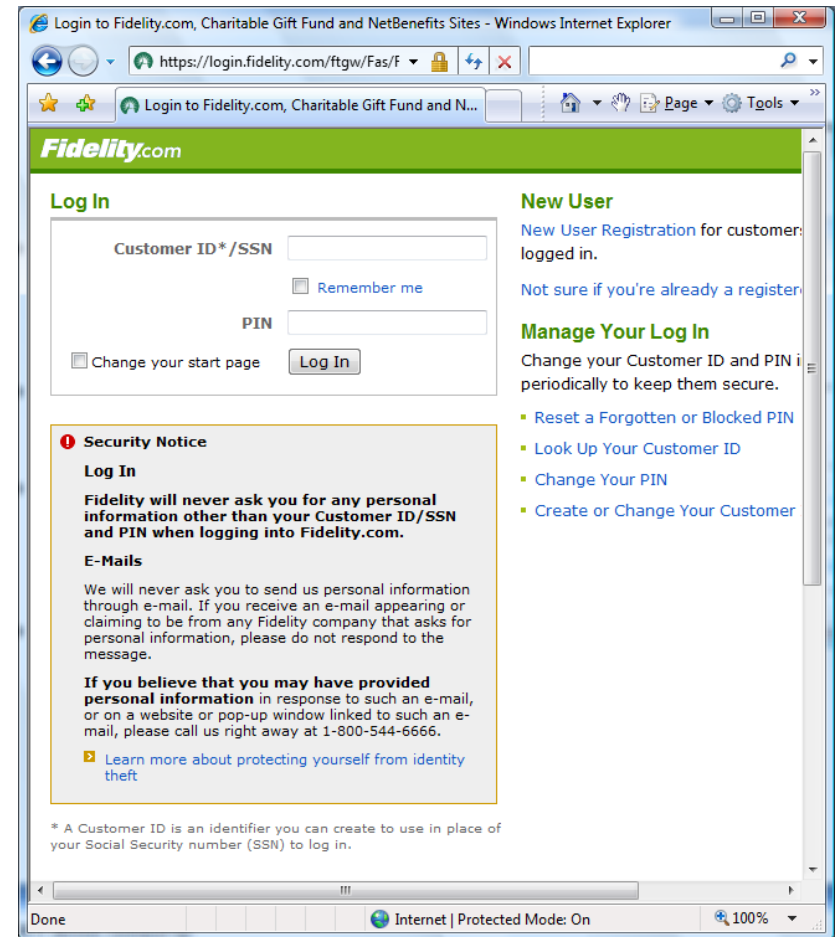
Measuring Revisitation

- Summary metrics
 - ▣ Unique visitors
 - ▣ Visits/user
 - ▣ Time between visits
- Revisitation curves
 - ▣ Revisit interval histogram
 - ▣ 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%)
 - *web science conference*
- Repeat click (39%)
 - <http://websci11.org>
 - Query → *websci 11*
- Lots of repeats (43%)
 - Many navigational

Repeat Query	33%
New Query	67%

Web Images Videos Shopping News Maps More | MSN Hotmail Jaime f Sign out 783 Bellevue, Washington Pr

bing

web science

Web News More

RELATED SEARCHES

- Life Science
- Science Kids
- Science News
- Science
- Popular Science Web Site
- Science Web Search
- Web of Science GA Tech
- ISI Web Science

SEARCH HISTORY

Turn on search history to start remembering your searches.

Turn history on

4 NARROW BY DATE


- All results
- Past 24 hours
- Past week
- Past month

ALL RESULTS 1-10 of 379,000,000 results - [Advanced](#)

Web Science - Web Science Trust

Web Science. The largest human information construct in history. The Web is transforming society. We have a unique opportunity and obligation to understand what ...

webscience.org

 Max 'Electronic' Van Kleek likes Web Science Trust

ISI Web of Knowledge - IP Error

All other users - If you are having difficulty accessing ISI Web of Knowledge, please log-in via your institution's remote access proxy (usually located on the library ...

www.isiknowledge.com

Web of Science - Science - Thomson Reuters

Web of Science © provides researchers, administrators, faculty, and students with quick, powerful access to the world's leading citation databases.

thomsonreuters.com/products_services/science/science_products/a-z/web_of_science

Web Science Trust - Wikipedia, the free encyclopedia

The Web Science Trust is a joint effort originally started between MIT and University of Southampton to bridge and formalize the social and technical aspects of the ...

en.wikipedia.org/wiki/Web_science

Welcome | Web Science @ Georgia Tech

Web Science @ Georgia Tech is an interdisciplinary research initiative that seeks to understand the impact of the Internet and participate in inventing its future.

webscience.cc.gatech.edu

http://www.bing.com/serp web science - Bing

Web Images Videos Shopping News Maps More | MSN Hotmail Jaime Sign out 783 Bellevue, Washington Pr

bing

web science

Web News More

RELATED SEARCHES

- Life Science
- Science Kids
- Science News
- Science
- Popular Science Web Site
- Science Web Search
- Web of Science GA Tech
- ISI Web Science

SEARCH HISTORY

Turn on search history to start remembering your searches.

Turn history on

NARROW BY DATE

- All results
- Past 24 hours
- Past week
- Past month

ALL RESULTS 1-10 of 379,000,000 results - Advanced

Web Science - Web Science Trust

Web Science. The largest human information construct in history. The Web is transforming society. We have a unique opportunity and obligation to understand what ...

webscience.org

Max 'Electronic' Van Kleek likes Web Science Trust

ISI Web of Knowledge - IP Error

All other users - If you are having difficulty accessing ISI Web of Knowledge, please log-in via your institution's remote access proxy (usually located on the library ...

www.isiknowledge.com

Web of Science - Science - Thomson

Web of Science © provides researchers, administrators, and librarians with quick, powerful access to the world's leading citation databases.

thomsonreuters.com/products_services/science/z/web_of_science

ACM Web Science Conference 2011: Home

ACM WebSci '11 - 3rd International Conference on Web Science ... Web Science is concerned with the full scope of socio-technical relationships that are engaged in the 7th

www.websci11.org

Web Science Trust - Wikipedia, the free encyclopedia

The Web Science Trust is a joint effort originally started between MIT and University of Southampton to bridge and formalize the social and technical aspects of the ...

en.wikipedia.org/wiki/Web_science

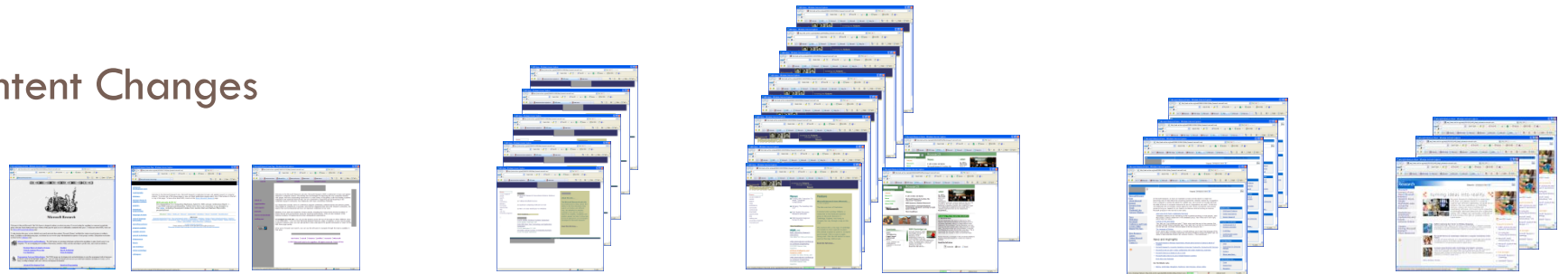
Welcome | Web Science @ Georgia Tech

Web Science @ Georgia Tech is an interdisciplinary research initiative that seeks to understand the impact of the Internet and participate in inventing its future.

webscience.cc.gatech.edu

How Revisitation and Change Relate

Content Changes



January

February

March

April

May

June

July

August

September

January

February

March

April

May

June

July

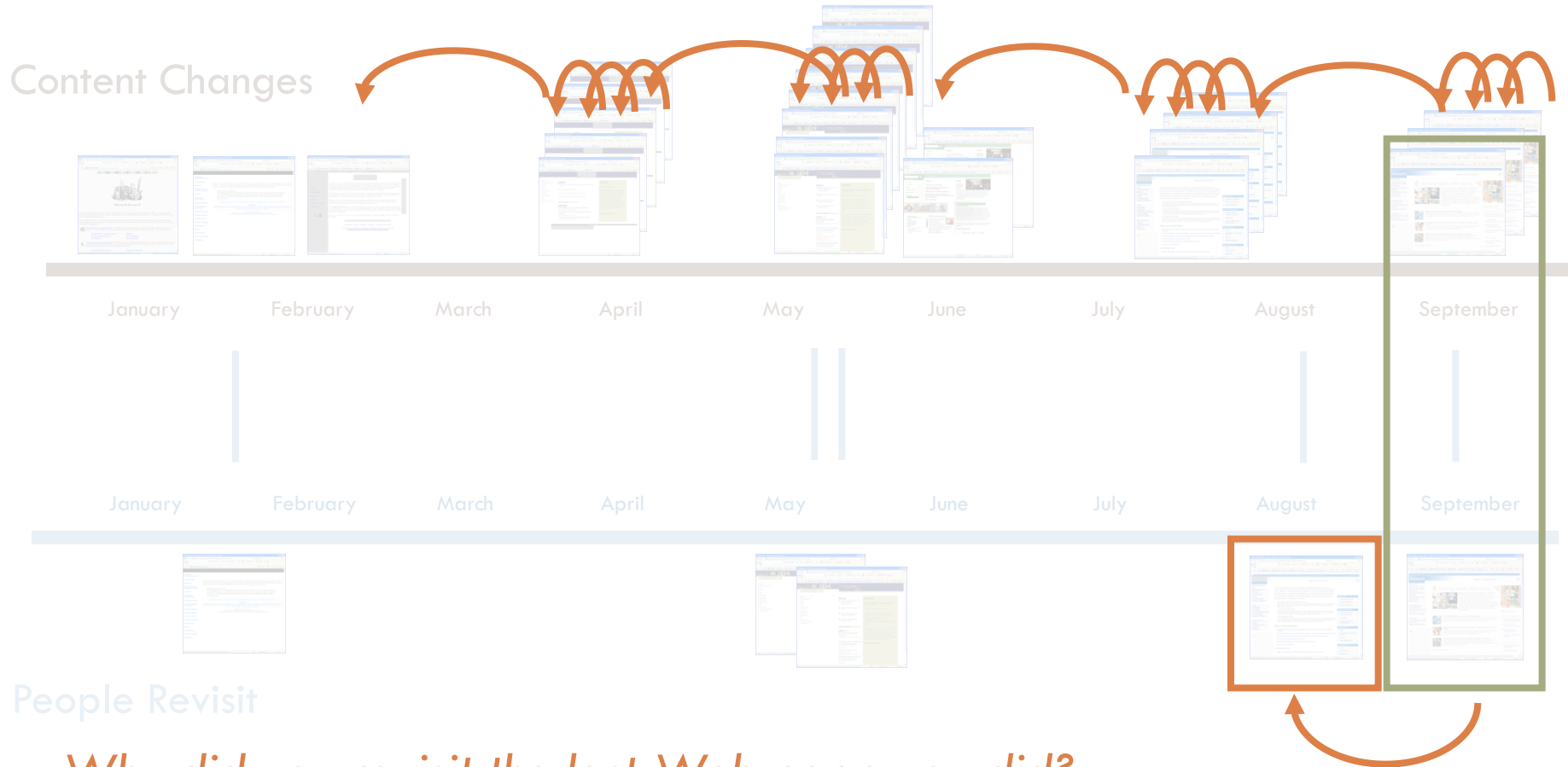
August

September

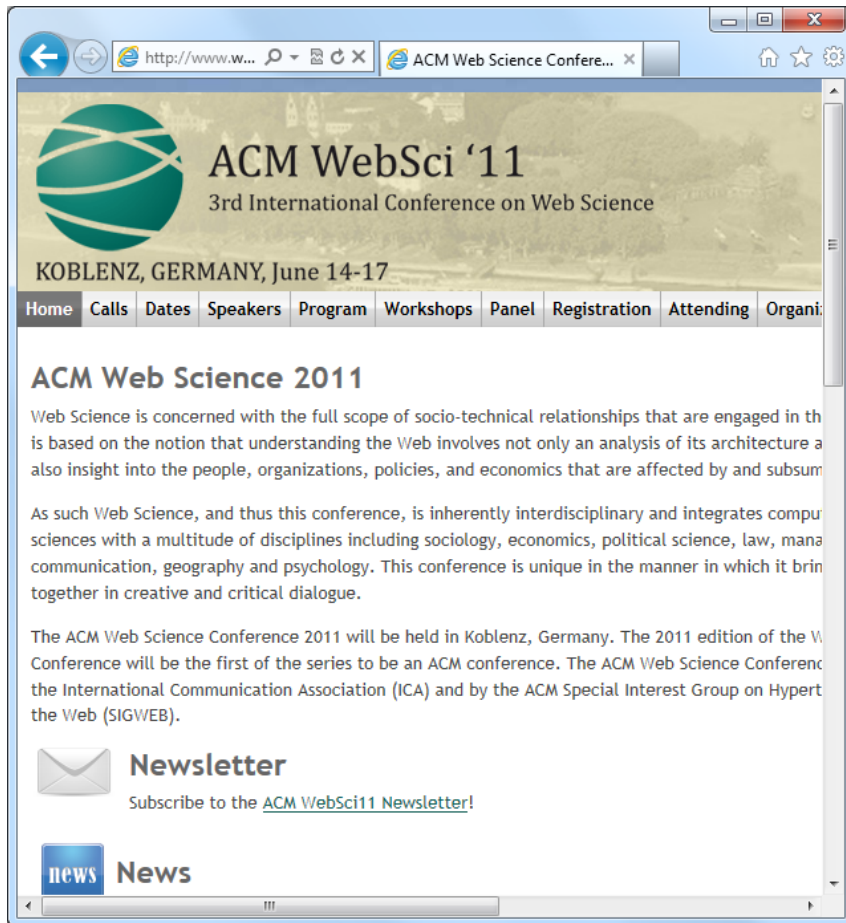
People Revisit

Why did you revisit the last Web page you did?

How Revisitation and Change Relate



Possible Relationships



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

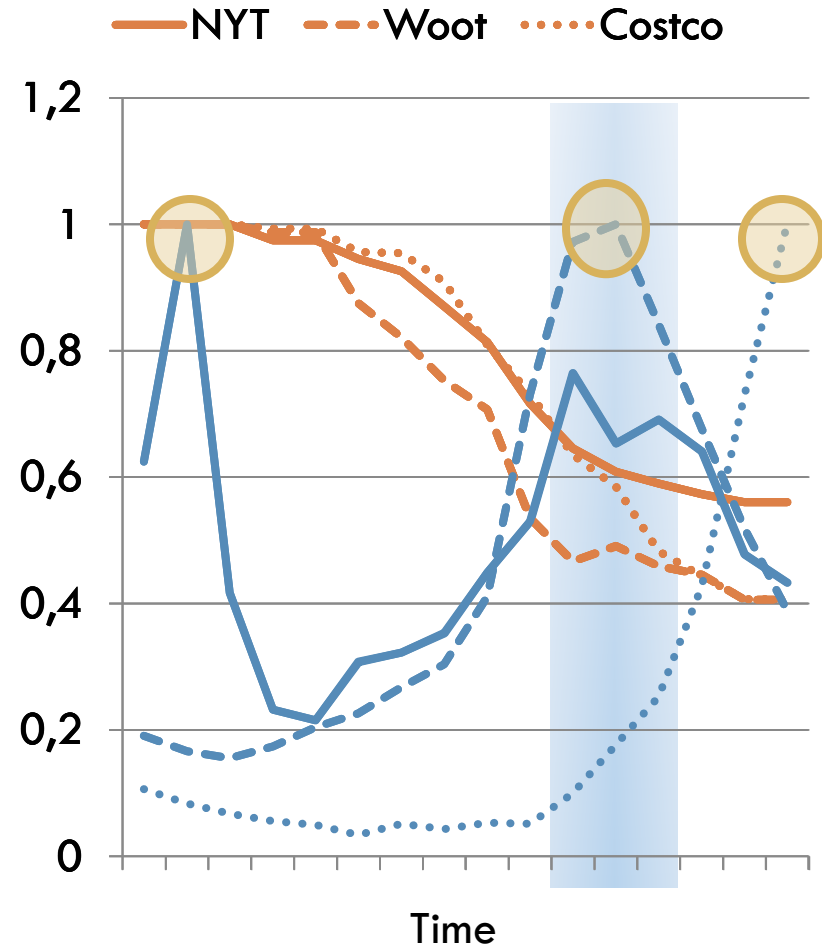
Understanding the Relationship

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

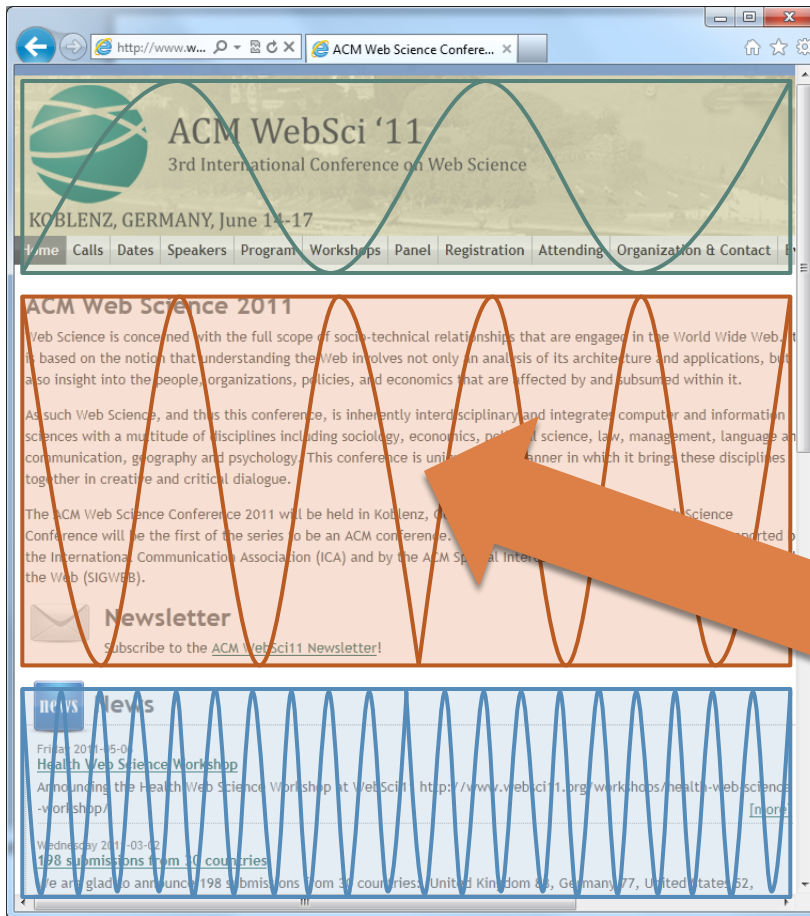
	Number of changes	Time between changes	Similarity
2 visits/user	172.91	133.26	0.82
3 visits/user	200.51	119.24	0.82
4 visits/user	234.32	109.59	0.81
5 or 6 visits/user	269.63	94.54	0.82
7+ visits/user	341.43	81.80	0.81

Comparing Change and Revisit Curves

- Three pages
 - ▣ New York Times
 - ▣ Woot.com
 - ▣ Costco
- Similar change patterns
- Different revisitation
 - ▣ NYT: *Fast*
 - ▣ Woot: *Medium*
 - ▣ Costco: *Slow*



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

OPINION »

An Olive Branch From the Dabai 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.
 No guesswork. [GO](#)

MARKETS »

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.](#)

[GO!](#)
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, pee 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.

**OFFERS
ENDING
SUNDAY**

[CLICK HERE](#)

**NEW
OFFERS**

[CLICK HERE](#)

**MORE THAN \$900
IN SAVINGS INSIDE
EVERY \$1,000**



Select Dell Desktops & Laptops
Up to \$400 OFF



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



BONUS \$150 DIRECTV OFFER
Select Branded TV Hot Buys
Up to \$400 OFF



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



ENDS SUNDAY
Saeco Deluxe Espresso
\$699.99 After \$200 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](#)

Online Savings Account

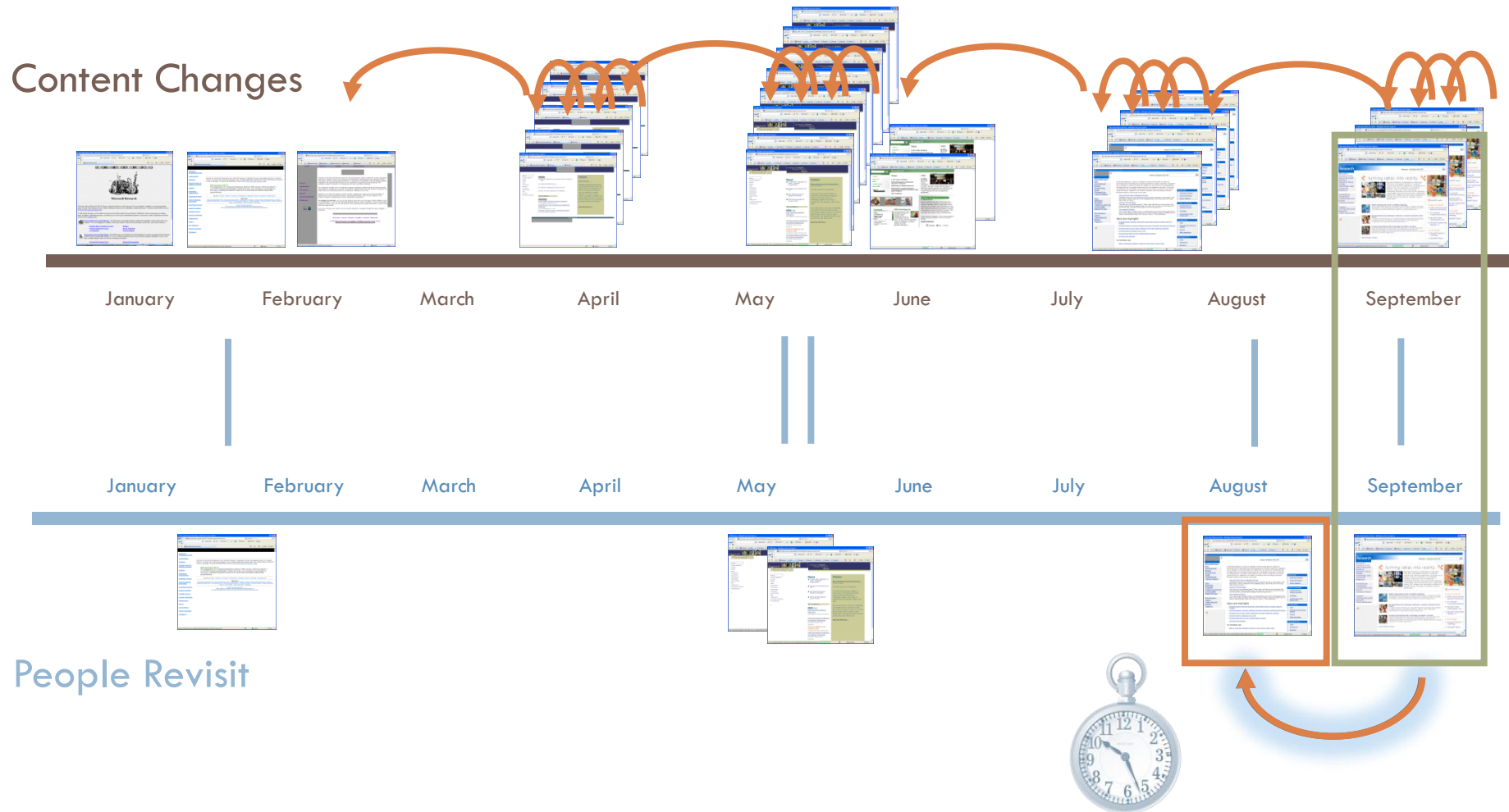
3.75% APY*

on balances of \$10,000 or more

MEMBER FDIC

Give us your Best Shot

Building Support for Web Dynamics



Exposing Change with *Diff-IE*

<http://bit.ly/DiffIE>

Diff-IE
toolbar

Bill Gates - Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Bill_Gates

Page loaded at 3:29 PM

Compare to Today at 10:45 AM

Bill Gates - Wikipedia, the free encyclopedia

interaction

- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link
- Cite this page

languages

- Afrikaans

2009.^[1] During his career at Microsoft, Gates held the positions of CEO and chief software architect, and remains the largest individual shareholder with more than 8 percent of the common stock.^[5] He has also authored or co-authored several books.

Gates is one of the best-known entrepreneurs of the personal computer revolution. Although he is admired by many, a number of industry insiders criticize his business tactics, which they consider anti-competitive, an opinion which has in some cases been upheld by the courts (see Criticism of Microsoft).^{[6][7]} In the later stages of his career, Gates has pursued a number of philanthropic endeavors, donating large amounts of money to various charitable organizations and scientific research programs through the Bill & Melinda Gates Foundation, established in 2000.

Bill Gates stepped down as chief executive officer

Bill Gates at the World Economic Forum in Davos, 2007

Born	October 28, 1955 (age 53) Seattle, Washington, USA
Residence	Seattle, Washington, USA

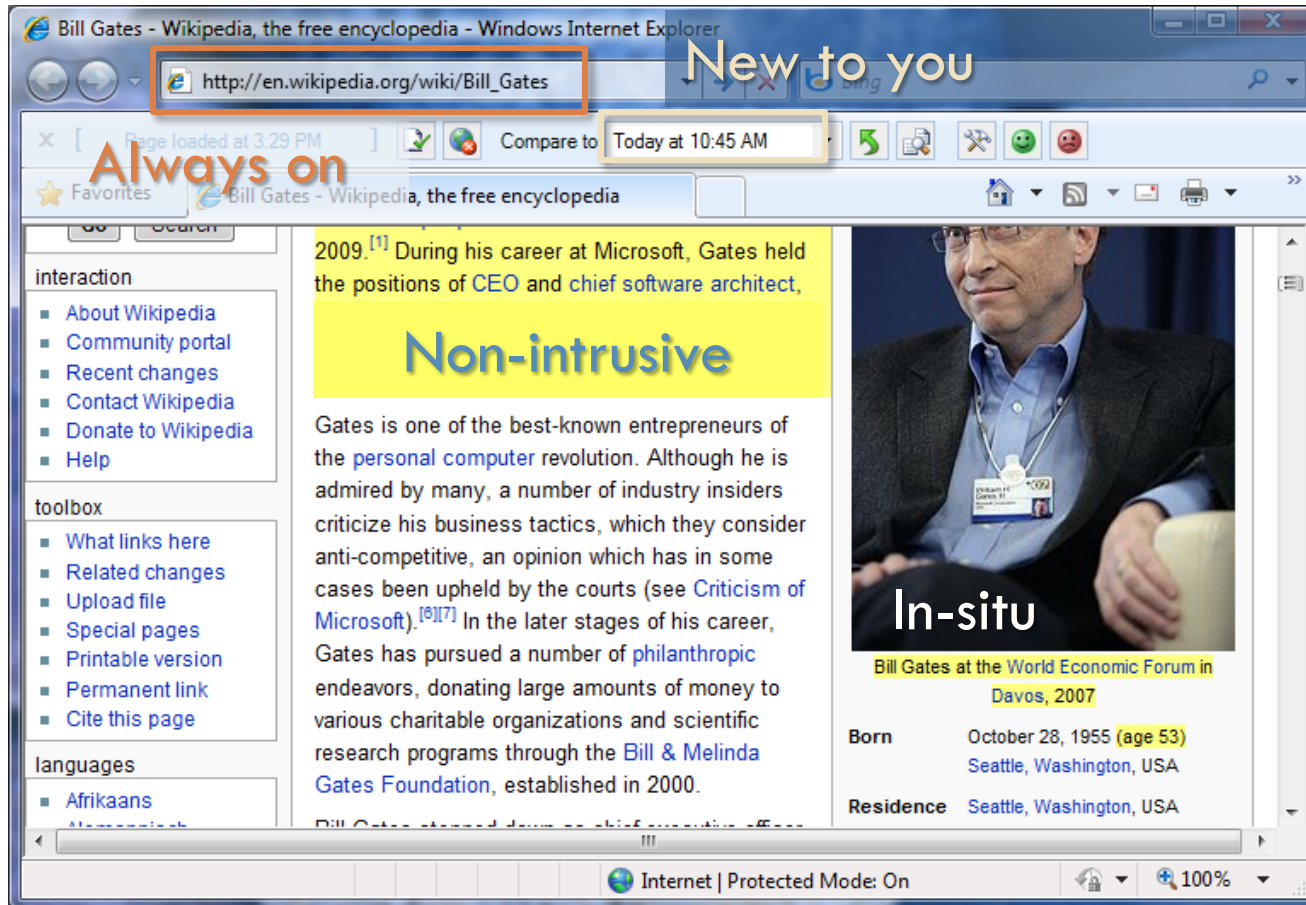
Internet | Protected Mode: On

100%

Changes to page
since your last
visit

Interesting Features of *Diff-IE*

<http://bit.ly/DiffIE>





Examples of *Diff-IE* in Action

<http://bit.ly/DiffIE>

Expected New Content

<http://bit.ly/DiffIE>

msnbc Powered by Bing MSN Home Mail More Sign In **msn**

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 Cha

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 IWM 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 us 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

ally Straightforward.

We do right by you and your money.

Mouse over tabs for details

9-Month No Penalty CD

Great Rates

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

Monitor

<http://bit.ly/DiffIE>

The image shows a screenshot of the Twitter homepage. At the top left is the Twitter logo. To the right are navigation links: Home, Profile, Find People, Settings. Below the logo is a text input field with the placeholder "What are you doing?" and a character count of "140". Below the input field is a "Latest" section showing a tweet from "Coding data and thinking about how people phrase the questions they ask." with a timestamp of "1 month ago" and an "update" button. Below that is a "Home" section with a list of tweets. The first tweet is from "brynn" (@informor) about "Reveal Two Types of Twitter Users, 'Informers' and 'Meformers'", posted "1 minute ago from web". The second tweet is from "emax" about "show everyone's surfing activity in real time", posted "26 minutes ago from web". The third tweet is also from "emax" about "if you have an eyebrowse account, log in + checkout your homepage", posted "29 minutes ago from web". The fourth tweet is from "dtunkelang" about "My blog post today at #thenoisychannel". On the right side, there is a user profile for "jteevan" with 38 following and 77 followers. Below the profile is a search bar and a "Home" section with links for "@jteevan", "Direct Messages", "Favorites", "Trending Topics" (including #haveuever, #stoptheviolence, Google Wave, Kristin, DWTS), and a "Search" input field.

Unexpected Important Content

<http://bit.ly/DiffIE>

The screenshot shows the Evite website interface. At the top left is the Evite logo. To the right is a banner for 'ditech' with the text 'THE DITEC REFI RATE SALE'. Below the banner, the page is divided into a left sidebar and a main content area.

Left Sidebar:

- send free invitations
- view invitations
 - Halloween
 - Halloween for Kids
 - Watch the Game
 - Dinner Party
 - Girls' Night Out
- get party ideas
 - Halloween
 - Halloween - Kids

Main Content Area:

MY UPCOMING EVENTS Welcome Jaime! Not Jaime?

10/17/09	7:30pm	A.J.'s 35th Birthday	19 NEW RSVP(s)	Attending
11/14/09	8pm	Studio 54 for Mark's Big 5-0	6 NEW RSVP(s)	Attending

MY PAST EVENTS

9/07/09	8:30am	Merrie's Bab...	No Albums Upload Photos
9/06/09	10:00am	Merrie's 30t...	No Albums Upload Photos
8/08/09	8:00am	Birthday Bug Bash!	1 Album(s) Upload Photos

[View All Events With Albums](#) | [View My Past Events](#)

Serendipitous Encounters


<http://bit.ly/DiffIE>



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, Danvel Fisher, Steven M. Drucker, Gonzalo Ramos, Paul André and Chang Hu. *Visual Snippets:*

Unexpected Unimportant Content

<http://bit.ly/DiffIE>

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 SLN

Understand Page Dynamics

<http://bit.ly/DiffIE>

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 search bar contains the text "weather 98005" and a magnifying glass icon. Below the search bar, the results are categorized under "ALL RESULTS" with a sub-section for "Weather".

Weather in 98005

Today	Wed	Thu	Fri	Sat
 61° / 50°	 59° / 51°	 61° / 50°	 58° / 39°	 60° / 37°

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

[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 ...

RELATED SEARCHES
[Weather 98052](#)
[Movies 98005](#)

SEARCH HISTORY
[weather 98005](#)
[xoost.com](#)
[fox news](#)
[msn](#)
[See all](#)
[Clear all](#) | [Turn off](#)

Attend to Activity

<http://bit.ly/DiffIE>

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 new topic

Subject	Started by	Replies	Views	Last post ▼
Small Business « 1 2 »	Della Dawn	21	206	February 17, 2009, 09:00 PM by Lori Poythress
Defining Sequencing	Miti Morris	0	93	November 10, 2005, 08:00 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

<http://bit.ly/DiffIE>

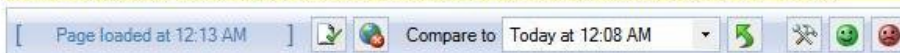


Installing DiffIE

- 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

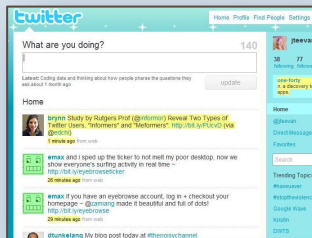
Once installed, you should see the DiffIE toolbar in IE (shown below on Vista):



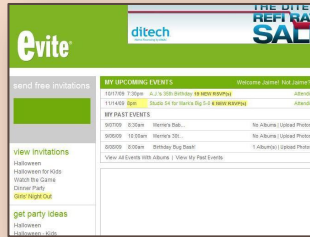
Expected



Expected New Content



Monitor



Unexpected Important Content

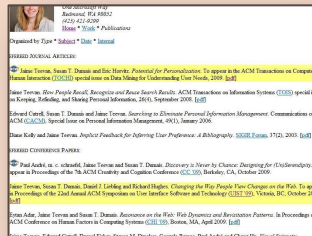
23. Educate Mothers & More Exps. See Chapter Names and Events for more information. Yeah, baby!

Share & More Message Board | Discussion Forum | Scheduling

1 2 3 ... 9 Mark Topics as Read for this Board | Notify of replies | Start new topic | Post

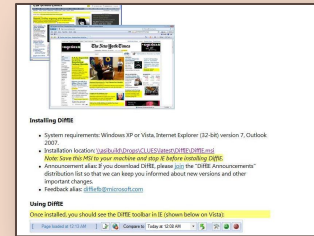
Subject	Started by	Replies	Views	Last post
Small Business > > >	Della Down	21	206	February 27, 2008, 10:04 AM
Defining Inequity	MG Morris	0	93	November 16, 2007, 10:04 AM
Back to work Part Time?	Sue Monahan	0	43	July 10, 2007, 10:04 AM
Almsgate Christian Preschool - HIBING	Heather Ross	0	17	July 10, 2007, 10:04 AM
Do you know anyone in Italy or Sweden?	Virginia Manning	3	70	July 10, 2007, 10:04 AM
Call an Interview and I'm freeling out... > > >	Tina Sooy	94	962	July 10, 2007, 10:04 AM
Accidents or Industry? Finding the Right Fit	Janet Terwan	7	105	July 10, 2007, 10:04 AM
Any members out at Amazon.com?	Kathryn Pizzo	0	108	July 10, 2007, 10:04 AM
Time Magazine on 9/11: getting back into work/life	Kathryn Pizzo	2	75	April 27, 2007, 10:04 AM
Freelance Resurgence	Kathryn Pizzo	1	41	April 26, 2007, 10:04 AM
Graphic designer or Artist?	Heather Ross	0	45	March 16, 2007, 10:04 AM

Attend to Activity

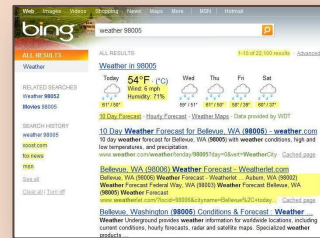


Serendipitous Encounter

Unexpected



Edit



Understand Page Dynamics



Unexpected Unimportant Content

Monitor

<http://bit.ly/DiffIE>

The Weather Channel
weather.com

Local Weather Site Web powered by Google™
98052 Search
Maps | Video | Photos | World | Mobile | Alerts


Home Weather News Travel Driving Health Home & Family **Sports** Outdoor Activities

Sports: [Ski Conditions](#) > **TODAY'S WEATHER for Redmond, WA (98052)**

Yesterday Today Tomorrow Radar Map Hour-by-Hour Weekend

MLB Weather | City Guides | Flu Shot Finder | Big Catch Photos

Right Now for Redmond, WA (98052)
Save Location [English | Metric]

 **Light Rain**


50°F
Feels Like 50°F

Updated Sep 29 07:05 p.m. PT

Weather for your life
Ski Conditions

UV Index: 0 Low
Wind: From E at 4 mph
Humidity: 93%
Pressure: 29.95 in. ↑
Dew Point: 48°F
Visibility: 4.0 miles
Download Now Free Desktop Weather

The Most Fuel-Efficient



Find Expected New Content

<http://bit.ly/DiffIE>

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 IWM 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 us money somehow," one Chicago resident says. [Full story](#)

U.S. news

Guns to be allowed in Arizona bars | Vote

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

advertisement | your ad here

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

Mouse over tabs for details

9-Month No Penalty CD

Great Rates

Categories
U.S. news | World news | Politics | Business | Sports | Entertainment

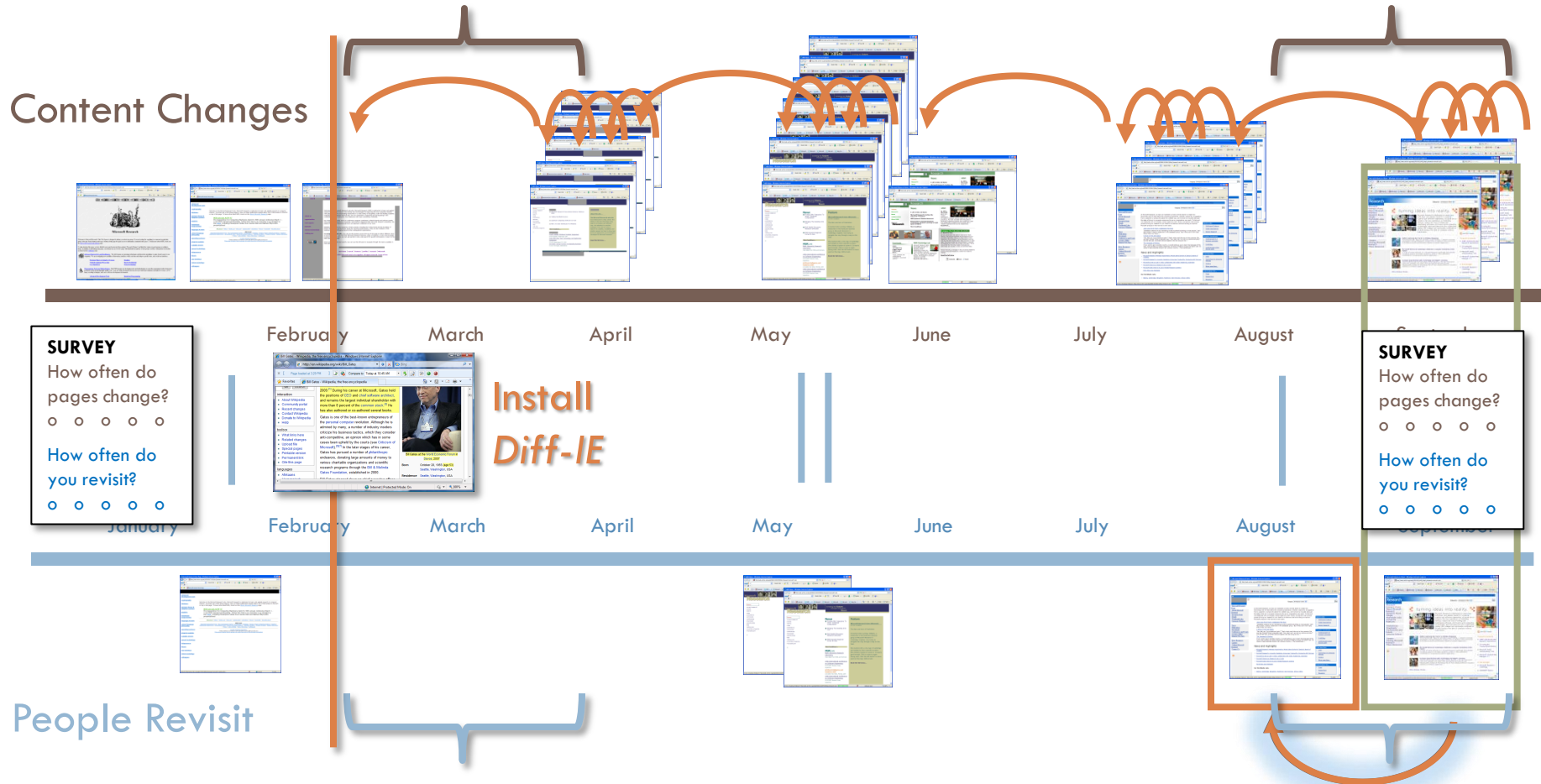
Explore
More news | My news | Pinned

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
[Slideshow: China](#)




Studying *Diff-IE*

<http://bit.ly/DiffIE>

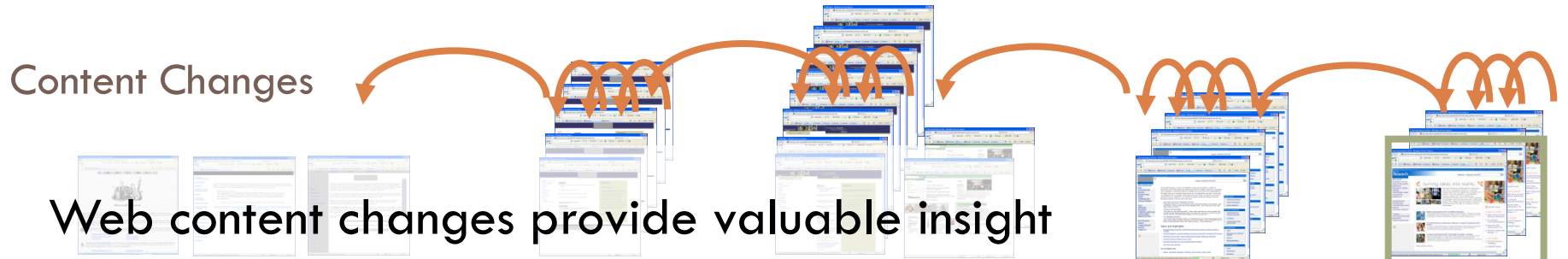


Seeing Change Changes Web Use

<http://bit.ly/DiffIE>

- Changes to perception
 - ▣ *Diff-IE* users become more likely to notice change
 - ▣ Provide better estimates of how often content changes
- Changes to behavior
 - ▣ Diff-IE users start to revisit more  14%
 - ▣ Revisited pages more likely to have changed  51%
 - ▣ Changes viewed are bigger changes  53%
- Content gains value when history is exposed

The Web Changes Everything



Relating revisitation and change enables us to

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

People revisit and re-find Web content

People Revisit



Explicit support for Web dynamics can impact how people use and understand the Web



Thank you.

Web Content Change

Adar, Teevan, Dumais & Elsas. *The Web changes everything: Understanding the dynamics of Web content*. WSDM 2009.

Elsas & Dumais. *Leveraging temporal dynamics of doc. content in relevance ranking*. WSDM 2010.

Kulkarni, Teevan, Svore & Dumais. *Understanding temporal query dynamics*. WSDM 2011.

Web Page Revisitation

Teevan, Adar, Jones & Potts. *Information re-retrieval: Repeat queries in Yahoo's logs*. SIGIR 2007.

Adar, Teevan & Dumais. *Large scale analysis of Web revisitation patterns*. CHI 2008.

Tyler & Teevan. *Large scale query log analysis of re-finding*. WSDM 2010.

Teevan, Liebling & Ravichandran. *Understanding and predicting personal navigation*. WSDM 2011.

Relating Change and Revisitation

Adar, Teevan & Dumais. *Resonance on the Web: Web dynamics and revisitation patterns*. CHI 2009.

Studying Diff-IE

Teevan, Dumais, Liebling & Hughes. *Changing how people view changes on the Web*. UIST 2009.

Teevan, Dumais & Liebling. *A longitudinal study of how highlighting Web content change affects people's web interactions*. CHI 2010.



Example: AOL Search Dataset

- August 4, 2006: Logs released to academic community
 - ▣ 3 months, 650 thousand users, 20 million queries
 - ▣ Logs contain anonymized User IDs
- August 7, 2006: AOL pulled the files, but already mirrored
- August 9, 2006: New York Times identified Thelma Arnold
 - ▣ [“A Face Is Exposed for AOL Searcher No. 4417749”](#)
 - ▣ Queries for businesses, services in Lilburn, GA (pop. 11k)
 - ▣ Queries for Jarrett Arnold (and others of the Arnold clan)
 - ▣ NYT contacted all 14 people in Lilburn with Arnold surname
 - ▣ When contacted, Thelma Arnold acknowledged her queries
- August 21, 2006: 2 AOL employees fired, CTO resigned
- September, 2006: Class action lawsuit filed against AOL



Example: AOL Search Dataset

- August 4, 2006: Logs released to academic community
 - ▣ 3 months, 650 thousand users, 20 million queries
 - ▣ Logs contain anonymized User IDs

□ August 7, 2006: AOL pulled the files but already mirrored

AnonID	Query	QueryTime	ItemRank	ClickURL
1234567	jitp	2006-04-04 18:18:18	1	http://www.jitp.net/
1234567	jipt submission process	2006-04-04 18:18:18	3	http://www.jitp.net/m_mscript.php?p=2
1234567	computational social science	2006-04-24 09:19:32		
1234567	computational social science	2006-04-24 09:20:04	2	
1234567	seattle restaurants	2006-04-24 09:25:50	2	
1234567	perlman montreal	2006-04-24 10:15:14	4	
1234567	jitp 2006 notification	2006-05-20 13:13:13		

- ▣ When contacted, Thelma Arnold acknowledged

□ August 21, 2006: 2 AOL employees

□ September, 2006: Class action lawsuit

A Face Is Exposed for AOL Searcher No. 4417749

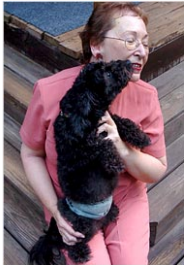
By MICHAEL BARBARO and TOM ZELLER Jr.
Published: August 9, 2006

Buried in a list of 20 million Web search queries collected by AOL and recently released on the Internet is user No. 4417749. The number was assigned by the company to protect the searcher's anonymity, but it was not much of a shield.

No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog that urinates on everything."

And search by search, click by click, the identity of AOL user No. 4417749 became easier to discern. There are queries for "landscapers in Lilburn, Ga.," several people with the last name Arnold and "homes sold in shadow lake subdivision winnnett county georgia."

It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga., frequently researches her friends' medical ailments and loves her three dogs. "Those are my searches,"



Erik S. Lesser for The New York Times
Thelma Arnold's identity was betrayed by AOL records of her Web searches, like ones for her dog, Dudley, who clearly has a problem.

SIGN IN TO E-MAIL THIS
PRINT
SINGLE PAGE
REPRINTS

127 HOURS



Example: AOL Search Dataset

- August 4, 2006: Logs released to academic community
 - ▣ 3 months, 650 thousand users, 20 million queries
 - ▣ Logs contain anonymized User IDs

□ August 7, 2006: AOL pulled the files but already mirrored

AnonID	Query	QueryTime	ItemRank	ClickURL
1234567	jitp	2006-04-04 18:18:18	1	http://www.jitp.net/
1234567	jipt submission process	2006-04-04 18:18:18	3	http://www.jitp.net/m_mscript.php?p=2
1234567	computational social science	2006-04-24 09:19:32		
1234567	computational social science	2006-04-24 09:20:04	2	http://socialcomplexity.gmu.edu/phd.php
1234567	seattle restaurants	2006-04-24 09:25:50	2	http://seattletimes.nwsourc.com/rests
1234567	perlman montreal	2006-04-24 10:15:14	4	http://oldwww.acm.org/perlman/guide.html
1234567	jitp 2006 notification	2006-05-20 13:13:13		
...				

- ▣ When contacted, Thelma Arnold acknowledged her queries
- August 21, 2006: 2 AOL employees fired, CTO resigned
- September, 2006: Class action lawsuit filed against AOL



Example: AOL Search Dataset

- Other well known AOL users
 - User 927 *how to kill your wife*
 - User 711391 *i love alaska*
 - <http://www.minimovies.org/documentaires/view/ilovealaska>
- Anonymous IDs do not make logs anonymous
 - Contain directly identifiable information
 - Names, phone numbers, credit cards, social security numbers
 - Contain indirectly identifiable information
 - Example: Thelma's queries
 - Birthdate, gender, zip code identifies 87% of Americans



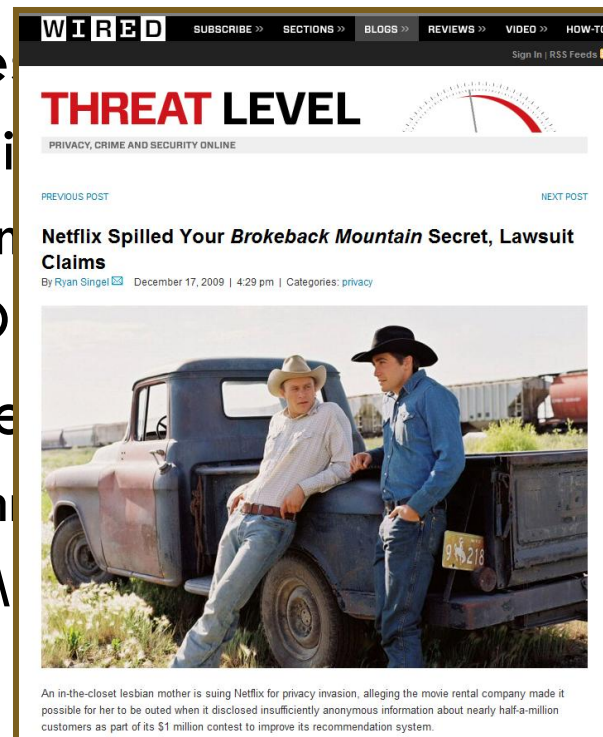
Example: Netflix Challenge

- October 2, 2006: Netflix announces contest
 - Predict people's ratings for a \$1 million dollar prize
 - 100 million ratings, 480k users, 17k movies
 - Very careful with anonymity post-AOL
- May 18, 2008: Data de-anonymized
 - Paper published by Narayanan & Shmatikov
 - Uses background knowledge from IMDB
 - Robust to perturbations in data
- December 17, 2009: Doe v. Netflix
- March 12, 2010: Netflix cancels second competition



Example: Netflix Challenge

- October 2, 2006: Netflix announces
 - Predict people's ratings for a \$1 million prize
 - 100 million ratings, 480k users, 17k movies
 - Very careful with anonymity post-AOL
- May 18, 2008: Data de-anonymized
 - Paper published by Narayanan & Shmatikov
 - Uses background knowledge from IMDb
 - Robust to perturbations in data
- December 17, 2009: Doe v. Netflix
- March 12, 2010: Netflix cancels second competition





Example: Netflix Challenge

- October 2, 2006: Netflix announces
 - Predict people's ratings for a \$1 million prize
 - 100 million ratings, 480k users, 17k new titles
 - Very careful with anonymity post-AOL

Ratings

1: [Movie 1 of 17770]

12, 3, 2006-04-18 [CustomerID, Rating, Date]

1234, 5, 2003-07-08 [CustomerID, Rating, Date]

2468, 1, 2005-11-12 [CustomerID, Rating, Date]

...

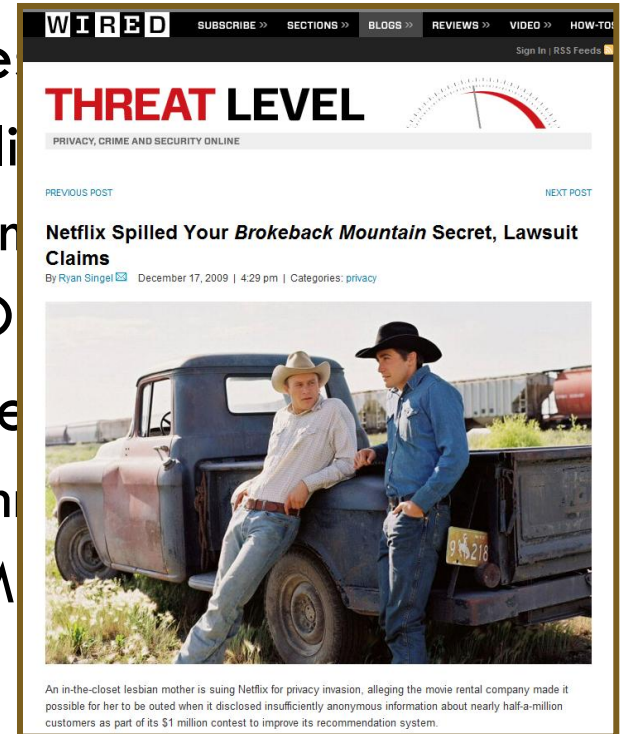
Movie Titles

...

10120, 1982, "Bladerunner"

17690, 2007, "The Queen"

...

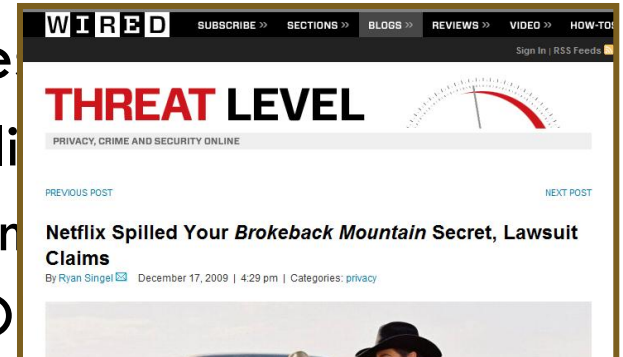


de-anonymize
 arayanan & Sh
 knowledge from IM
 s in data
 Doe v. Netflix
 Netflix cancels second competition



Example: Netflix Challenge

- October 2, 2006: Netflix announces
 - Predict people's ratings for a \$1 million prize
 - 100 million ratings, 480k users, 17k movies
 - Very careful with anonymity post-AOL



Ratings	
1:	[Movie 1 of 17770]
12, 3, 2006-04-18	[CustomerID, Rating, Date]
1234, 5, 2003-07-08	[CustomerID, Rating, Date]
2468, 1, 2005-11-12	[CustomerID, Rating, Date]
...	
Movie Titles	
...	
10120, 1982, "Bladerunner"	
17690, 2007, "The Queen"	
...	

*All customer identifying information has been removed; all that remains are ratings and dates. This follows our privacy policy. . . Even if, for example, you knew all your own ratings and their dates you probably couldn't identify them reliably in the data because **only a small sample was included** (less than one tenth of our complete dataset) and that **data was subject to perturbation.***