

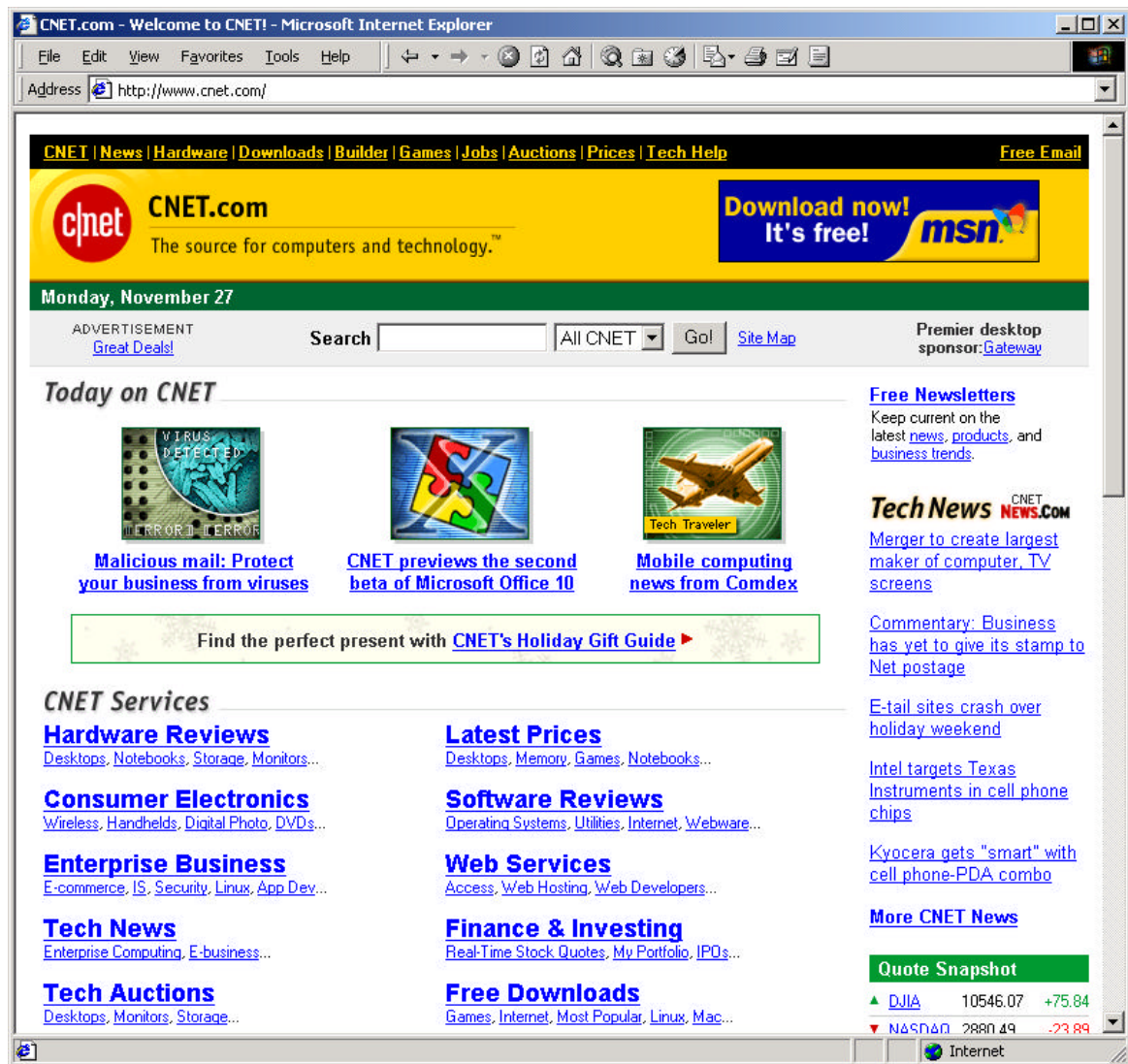
# Improving Web Interactions with Machine Learning

---

Corin Anderson • Pedro Domingos • Daniel Weld  
UW CSE Industrial Affiliates 2002

# The Web for the Masses

- **One-size-fits-all**
  - Same content for everyone
- **Not true!**
  - Different devices
  - Different info needs
  - Different navigation styles
  - Different browsing contexts



# The Web for You

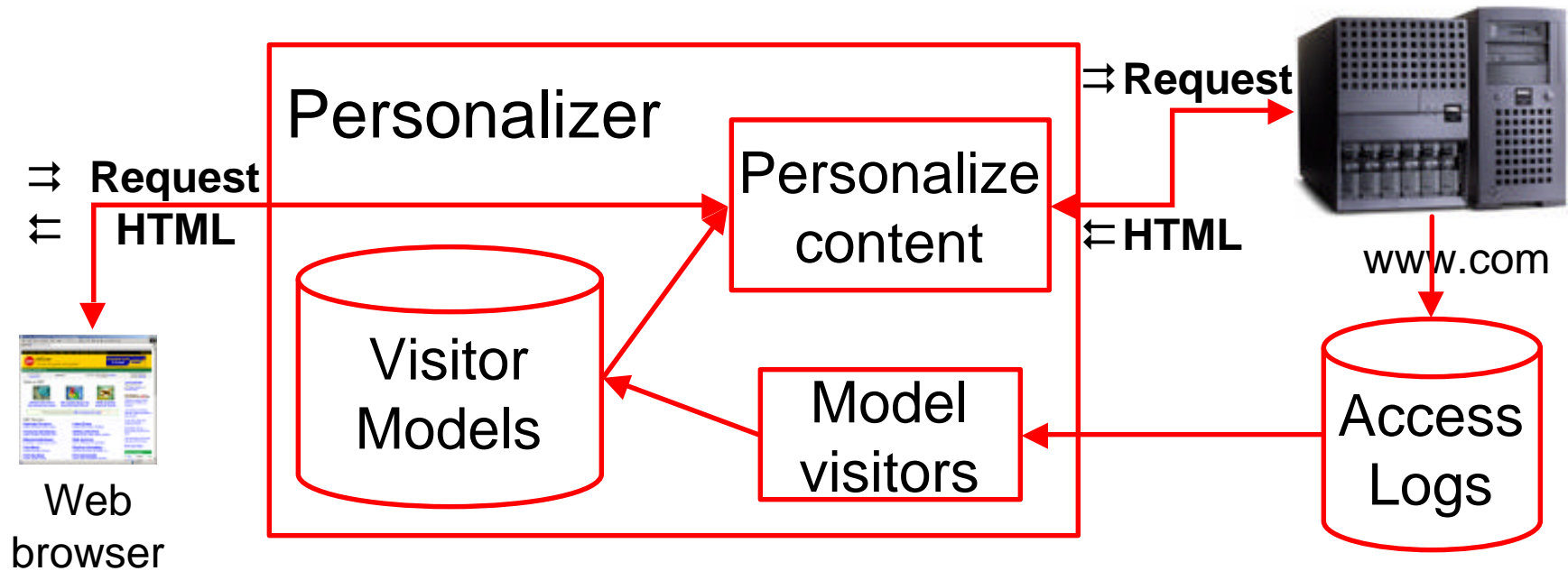
---

- The web interface should be
  - **Personalized** to your browsing style
  - **Dynamic** to your current browsing context
- Our approach: apply machine learning
  - **Proteus**: adapting sites for wireless devices
  - **MinPath**: finding shortcuts by clustering users
  - **Montage**: building dynamic, personalized start pages

# Web personalization process

- Step 1: **model users**
  - Use machine learning to build user models
  - Mine web logs for navigational patterns
  - Analyze text of requested web pages
- Step 2: **adapt interface**
  - Consider a modification
  - Evaluate expected utility based on model
  - Repeat many times; commit the best ones

# The Proteus architecture



- Server provides access logs
- Personalizer builds visitor models, modifies returned content

# Proteus: adapting wireless web

- Small-screen, low-bandwidth devices
- Adapt by **eliding content, adding shortcuts**
- Evaluate personalized site by:
  - Summing **value of each page** in site
    - Pages viewed often have high value
  - But **discounting by difficulty** in reaching it
    - Scrolling to find a link takes time, visitor effort
    - Following a link take a lot of time, may cost money

# Proteus in action

---



CNET.com - Welcome to CNET! - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.cnet.com/

CNET | News | Hardware | Downloads | Builder | Games | Jobs | Auctions | Prices | Tech Help Free Email

**cnet** CNET.com  
The source for computers and technology.™

**Download now!**  
It's free!

Monday, November 27

ADVERTISEMENT [Great Deals!](#) Search  All CNET Go! [Site Map](#) Premier desktop sponsor: [Gateway](#)

**Today on CNET**



**Malicious mail: Protect your business from viruses**



**CNET previews the second beta of Microsoft Office 10**



**Mobile computing news from Comdex**

Find the perfect present with [CNET's Holiday Gift Guide](#) ▶

**CNET Services**

**Hardware Reviews**  
[Desktops, Notebooks, Storage, Monitors...](#)

**Consumer Electronics**  
[Wireless, Handhelds, Digital Photo, DVDs...](#)

**Enterprise Business**  
[E-commerce, IS, Security, Linux, App Dev...](#)

**Tech News**  
[Enterprise Computing, E-business...](#)

**Tech Auctions**  
[Desktops, Monitors, Storage...](#)

**Latest Prices**  
[Desktops, Memory, Games, Notebooks...](#)

**Software Reviews**  
[Operating Systems, Utilities, Internet, Webware...](#)

**Web Services**  
[Access, Web Hosting, Web Developers...](#)

**Finance & Investing**  
[Real-Time Stock Quotes, My Portfolio, IPOs...](#)

**Free Downloads**  
[Games, Internet, Most Popular, Linux, Mac...](#)

**Free Newsletters**  
Keep current on the latest [news](#), [products](#), and [business trends](#).

**Tech News** CNET NEWS.COM  
[Merger to create largest maker of computer, TV screens](#)

[Commentary: Business has yet to give its stamp to Net postage](#)

[E-tail sites crash over holiday weekend](#)

[Intel targets Texas Instruments in cell phone chips](#)

[Kyocera gets "smart" with cell phone-PDA combo](#)

**More CNET News**

**Quote Snapshot**

▲ DJIA	10546.07	+75.84
▼ NASDAQ	2880.49	-23.89

Internet





Elide  
content

**EudoraWeb**

[&] CNET tackles 4...  
[CNET Services]

**Hardware Reviews**  
Desktops, Notebooks [ CNET.com -  
Hardware - Notebooks - Top 5 - Top  
5 Ultralight Notebooks ], Storage,  
Monitors...

← [Details] [Bookmarks]

Add shortcut link

# Proteus results

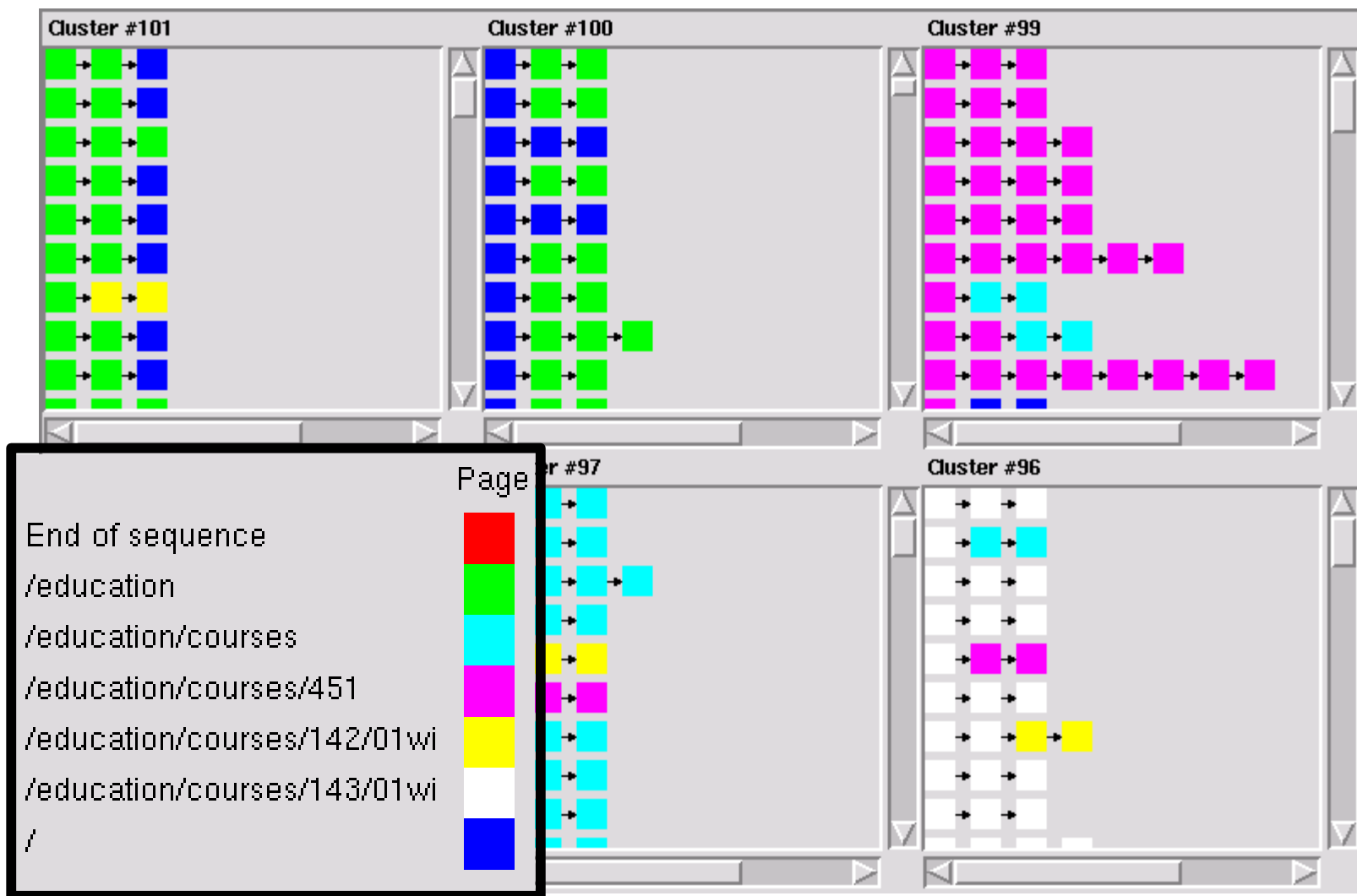
---

- Evaluated Proteus with user study
- Measure user effort required to answer questions using mobile device
  - Unmodified vs. Proteus-personalized sites
- Results: **Proteus wins**
  - Required less time
  - Required fewer navigation actions

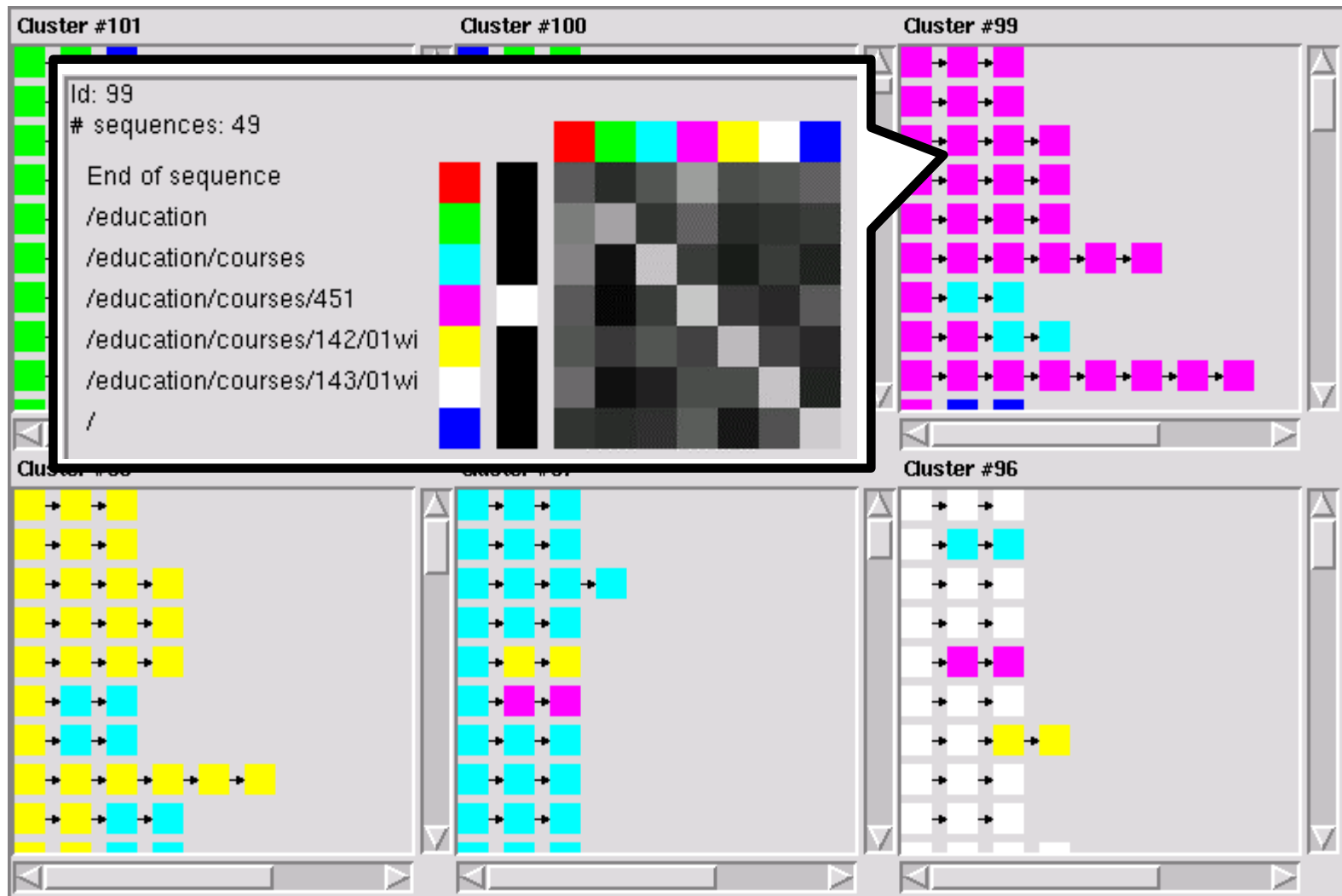
# MinPath: finding shortcuts

- Concentrate exclusively on **shortcut links**
  - Replace  $A \rightarrow B \rightarrow \dots \rightarrow X$  with  $A \rightarrow X$
  - Can add only 3 – 5 shortcuts per page
- Choose links with most **expected savings**
  - Calculate **probability** of each link destination
  - Scale by how many links shortcut would **save**
- Model **clusters of visitors**, not individuals

# Visitor clusters at UW/CSE



# Clusters models



# MinPath results

---

- Evaluated MinPath with UW/CSE log data
- Compared non-clustered with clustered
- Compared various sequence models
- Results: **MinPath saves 40% nav. effort**
  - 10 clusters of users better than only one
  - MinPath runs very quickly

# Montage: improved start page

- Dynamic, personalized start page
  - Browser “Home” page
- Display links and content user wants **now**
- Montage is **dynamic**, based on:
  - Time of day
  - Topic of recent browsing

*Work conducted while on  
internship at Microsoft Research*

# Main Montage

- **Embedded lenses** on most interesting content
- Links to other pages of value
- Links to **topic-specific** montages





# Montage clipping

The Seattle Times Home Page - Microsoft Internet Explorer

A Service of The Seattle Times Company seattletimes.com NWclassifieds NWsource

THURSDAY November 1, 2001

seattletimes.com The Seattle Times

NWsource 51°F

Page updated at 11:30 a.m.

## Microsoft, U.S. shape tentative settlement

MARK HARRISON / THE SEATTLE TIMES

### A bite out on the town

Flashing enhanced incisors, Jen Taylor arrives at the Troll under the Aurora Bridge for a Halloween celebration. Fremont's Cirque du Flamé performance group played host to a gathering under the bridge that culminated with a costume parade of 200 people through the neighborhood.

**The vaccine may be delayed, but supply should be adequate**

UW study says more children's hospital visits are for mind disorders than for accidents

Terrorism scare delays plane unloading at Sea-Tac

Crash scene clash brings police suit

Official limiting public's opinion on 'credit scoring' by insurers

Seattle city attorney rivals differ most in backgrounds

Sisal server: Fast Flagler and other parks offer unique escapes at off season bargain

ATTACK ON TURKISHES

Understanding the Conflict  
A changed name. A new kind of war.

Photo galleries

WAR

- The trial of bin Laden
- Taliban claim more civilian deaths
- Britain, U.S. begin global PR effort
- Northern Alliance's flaws are tough to camouflage

THE RESPONSE

- House readies competing plans for almost security

ANTRAX

- N.Y. woman fourth to die of mailed letters
- An ordinary life ends swiftly in anthrax mystery

BUSINESS

- Job American's policy won't be allowed
- Boeing ad's offer stuns reassurances

CROSSING AMERICA

- Read our series

BREAKING NEWS ALERTS

Be the first to know! Sign up for breaking news e-mail from seattletimes.com.

150 YEARS

SEATTLE BE AND BY SEATTLE'S HISTORY: A full with a history. See the historical list of 150 most influential people.

NWsource

- Outdoors
- TV Listings
- Vacation Guide
- Shopping Guide

Search

Advanced search

NWclassifieds

- Jobs
- Autom
- Home & Rentals

Highlights

High School sports

Coverage of all the fall sports; plus a new database full of schedules, standings, rosters and statistics for football.

Election 2001

Seattle

Mayor, initiatives, local races.

Suburban news:

Find news about the Eastside and Southside Coast.

Internet

Corin's Main Montage - Microsoft Internet Explorer

Main **Refresh** Your other MyMontage - Business & Finance - Computers & Internet - Entertainment & Media  
Montages: - Home & Family - Society, Politics & News - Sports & Recreation - Unregistered

## Society, Politics, & News

Preferred content:

Customize size & focus [The Seattle Times Home Page](#)

Page updated at 06:54 a.m.

## Microsoft, U.S. shape tentative settlement

MARK HARRISON / THE SEATTLE TIMES

### A bite out on the town

Flashing enhanced incisors, Jen Taylor arrives at the Troll under the Aurora bridge for a Halloween celebration. Fremont's Cirque du Flamé performance group played host to a gathering under the bridge that culminated with a costume parade of 200 people through the neighborhood.

**The vaccine may be delayed, but supply should be adequate**

UW study says more children's hospital visits are for mind disorders than for accidents

Terrorism scare delays plane unloading at Sea-Tac

Crash scene clash brings police suit

Official limiting public's opinion on 'credit scoring' by insurers

Seattle city attorney rivals differ most in backgrounds

Sisal server: Fast Flagler and other parks offer unique escapes at off season bargain

ATTACK ON TURKISHES

Understanding the Conflict  
A changed name. A new kind of war.

Photo galleries

WAR

- The trial of bin Laden
- Taliban claim more civilian deaths
- Britain, U.S. begin global PR effort
- Northern Alliance's flaws are tough to camouflage

THE RESPONSE

- House readies competing plans for almost security

ANTRAX

- N.Y. woman fourth to die of mailed letters
- An ordinary life ends swiftly in anthrax mystery

BUSINESS

- Job American's policy won't be allowed
- Boeing ad's offer stuns reassurances

CROSSING AMERICA

- Read our series

BREAKING NEWS ALERTS

Be the first to know! Sign up for breaking news e-mail from seattletimes.com.

150 YEARS

SEATTLE BE AND BY SEATTLE'S HISTORY: A full with a history. See the historical list of 150 most influential people.

NWsource

- Outdoors
- TV Listings
- Vacation Guide
- Shopping Guide

Search

Advanced search

NWclassifieds

- Jobs
- Autom
- Home & Rentals

Highlights

High School sports

Coverage of all the fall sports; plus a new database full of schedules, standings, rosters and statistics for football.

Election 2001

Seattle

Mayor, initiatives, local races.

Suburban news:

Find news about the Eastside and Southside Coast.

Internet

## Computers & Internet

Preferred content:

Customize size & focus [Python 2.1.1 Documentation - July 20, 2001](#)

[Tutorial](#) (start here)

[Global Module Index](#) (for quick access to all documentation)

[Library Reference](#) (keep this under your pillow)

Also of interest in Computers & Internet:

- [Python 2.1.1](#)
- [Python Language Website](#)

## Entertainment & Media

Preferred content:

Customize size & focus [WSDOT Puget Sound Area Traffic Cameras](#)

Also of interest in Entertainment & Media:

- <http://raindotsonmedia.com/mg/home>
- [Ice The Circle: Volume 12, Ice](#)

Done

Unknown Zone (Mixed)

# Montage results

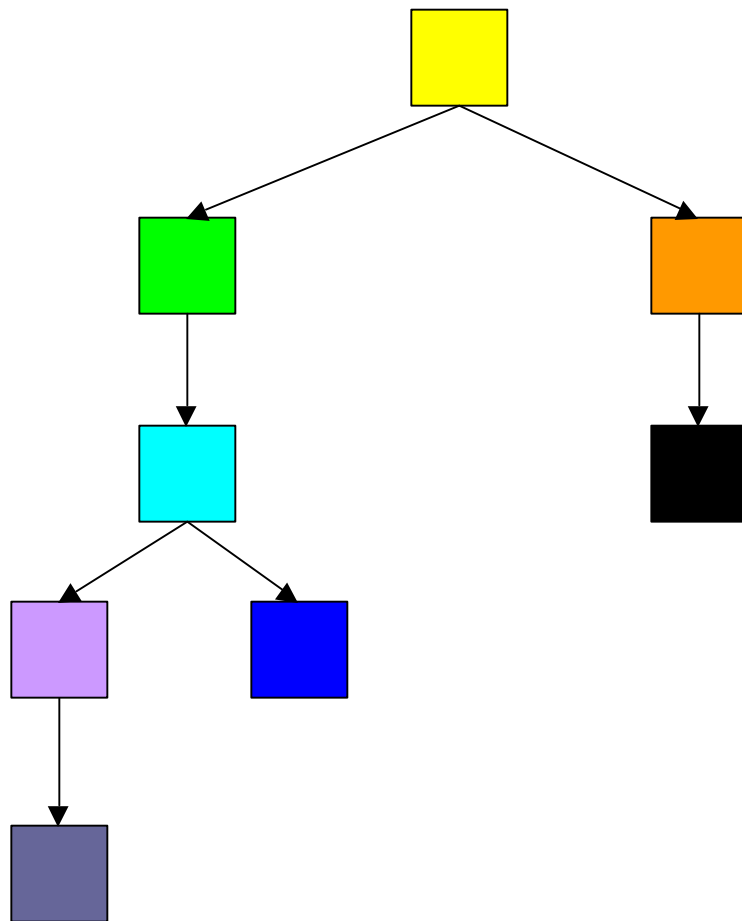
- Evaluated Montage with user study
- Compared predictive user models
- Compared montage visualizations
- Results: **context-sensitive model best**
  - Content changes throughout day
  - Suggestions depend on value of content

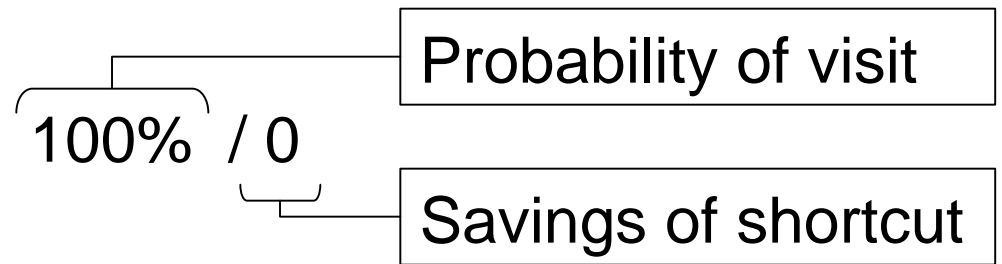
# Summary

---

- Machine learning enables effective improvements to web interactions
- Proteus architecture suitable for variety of systems
- Interactions at many levels (page, site, browser) can be improved substantially

# Calculating expected savings





50% / 0

40% / 0

48% / 1

20% / 2

19% / 3

16% / 2

30% / 1



$$.19 \times 3 = .57$$



$$.48 \times 1 = .48$$



$$.20 \times 2 = .40$$