

Security and Performance in Modern Web Browsers

Ben Livshits

Microsoft Research
Redmond, Washington

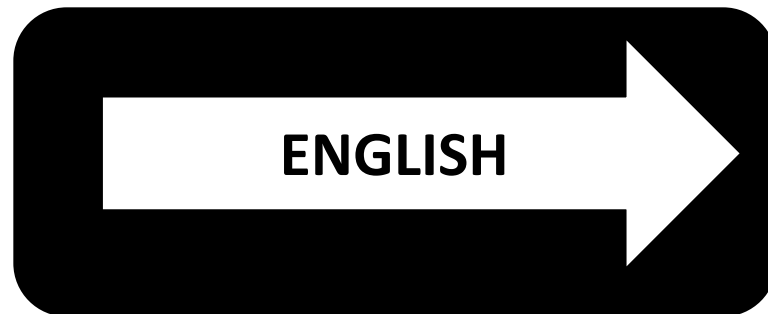
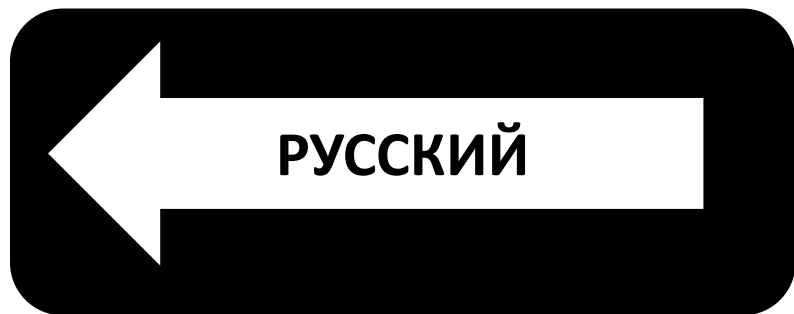


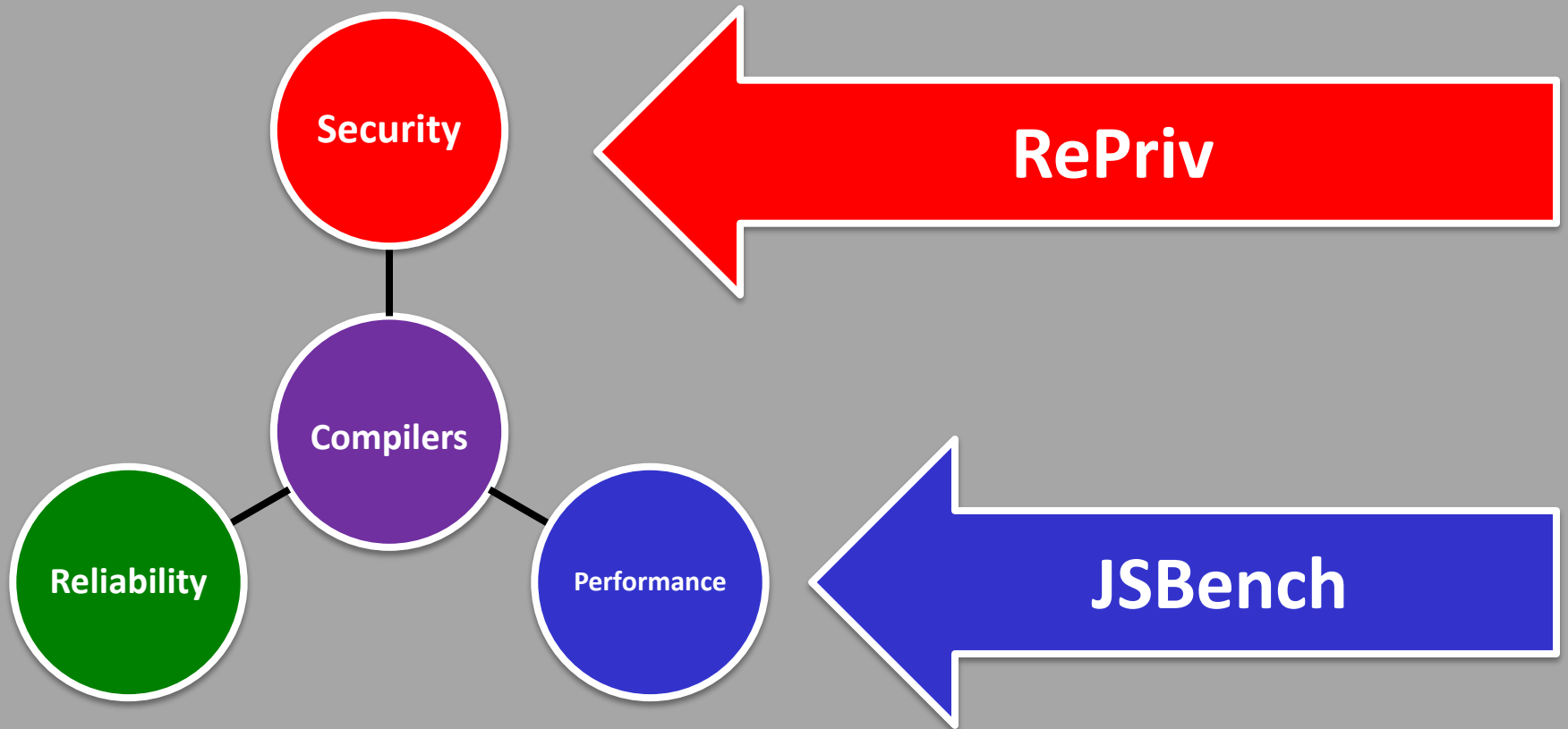
Microsoft Research
Redmond, WA

Организация Microsoft Research

- 300-400 исследователей и ученых
- Исследования в computer science
- ...с двумя основными целями:
 - академическое влияние и публикации
 - исследования, отражающиеся в продуктах Microsoft'а в следующие 5-10 лет







Type systems

Interest mining

Personalization

Provable privacy

RePriv

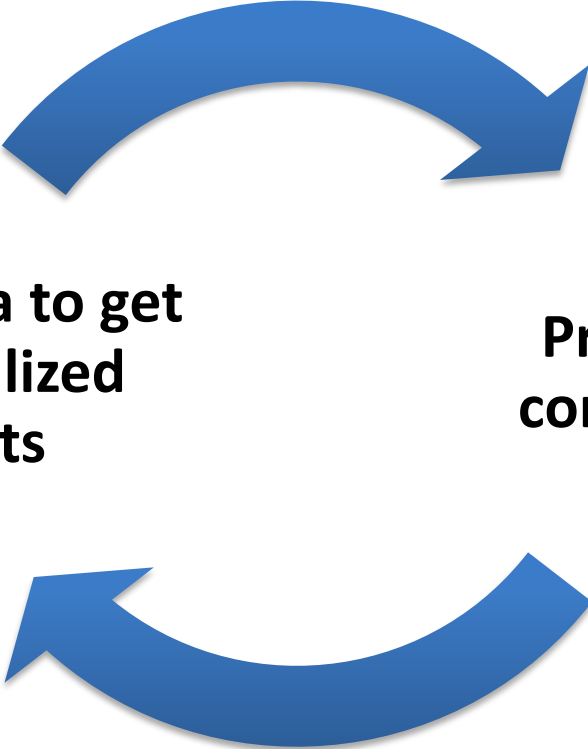
Verification

Re-envisioning in-browser privacy

Browser

Network protocols

Languages



**Share data to get
personalized
results**

**Privacy
concerns**

Approach, Opportunity & Privacy



Browsing history

Top: Computers: Security
Top: Arts: Movies
Top: Sports: Hockey
Top: Science: Math
Top: Recreation: Outdoors

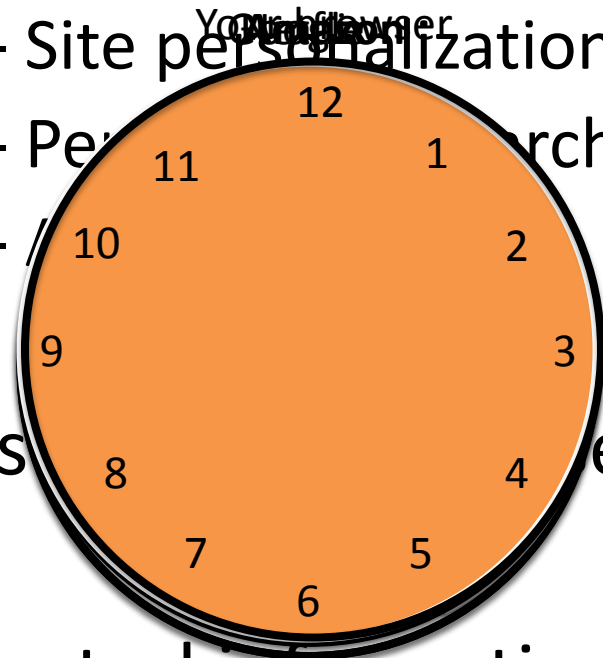


Distill



User interest profile

- Broad applications:
 - Site personalization
 - Personalized search
 - /
- User interest profile
- Control information release



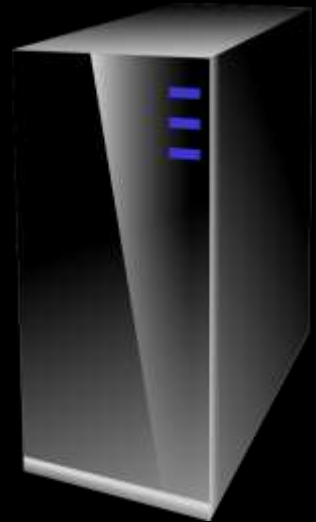
bn.com would like to learn your top interests.
We will let them know you are interested in:

- Science
- Technology
- Outdoors

Accept

Decline

RePriv Protocol



GET /index.html HTTP 1.1
Host: www.example.com

Accept: repriv ...

HTTP/1.1 **300 Multiple Choices**
index.html
index.html?top-n&level=m

POST /index.html HTTP 1.1
Host: www.example.com
Content-Length: x

category1=c1&...

HTTP/1.1 200 OK

Personalized page content

Would you like to install an extension called “Bing Personalizer” that will:

- Watch mouse clicks on bing.com
- Modify appearance of bing.com
- Store personal data in browser

Accept

Decline

Contributions of RePriv

RePriv

- An in-browser framework for collecting & managing personal data to facilitate personalization.

Core Behavior Mining

- Efficient in-browser behavior mining & controlled dissemination of personal data.

RePriv miners

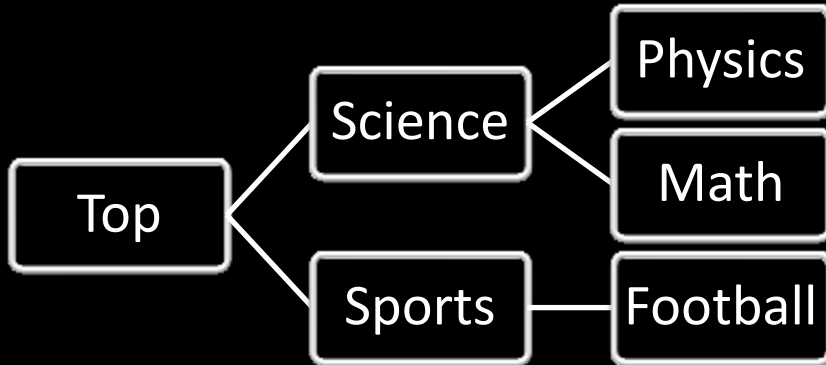
- A framework for integrating verified third-party code into the behavior mining & dissemination of RePriv.

Real-world Evaluation

- Evaluation of above mechanisms on real browsing histories & two in-depth case studies.

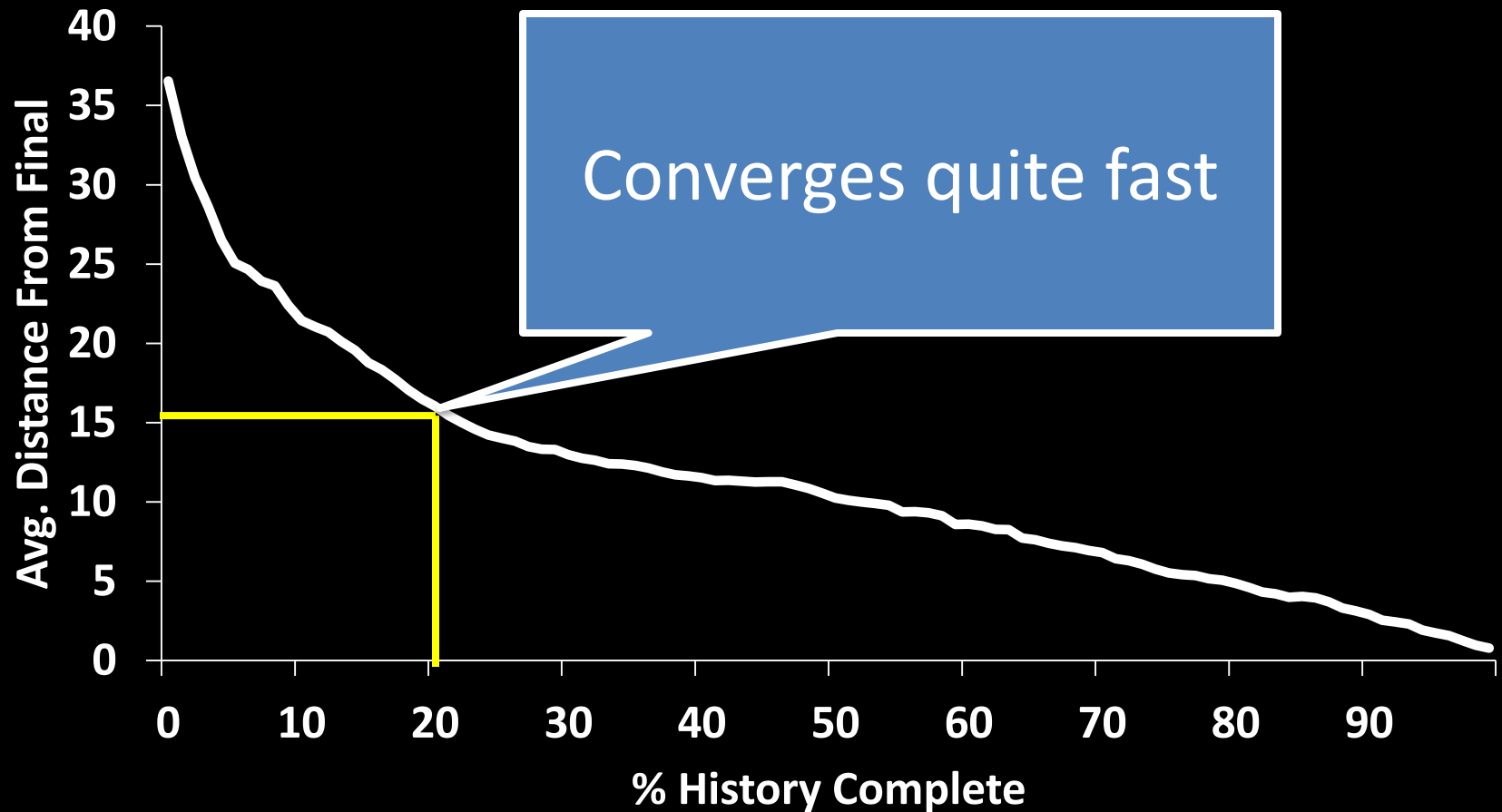
Core Mining

- Taxonomy from first two levels of ODP taxonomy
 - ~450 categories total
 - 20 top-level categories
 - Overlap exists

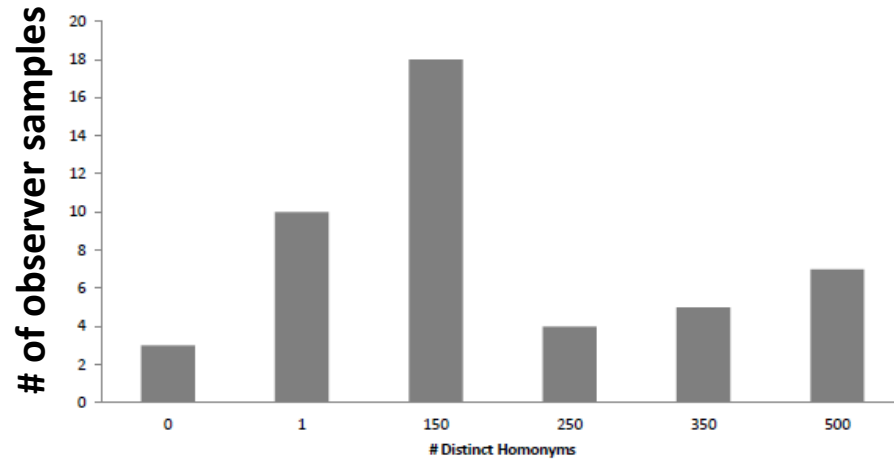


- Naïve Bayes
 - All categories equally likely
 - Training: min(3000, # pages) sites per category
 - Attribute words occur in at least 15% of docs for ≥ 1 category
- Classification is fast enough: $O(c \cdot n)$
 - n is # words in document
 - c is # document categories

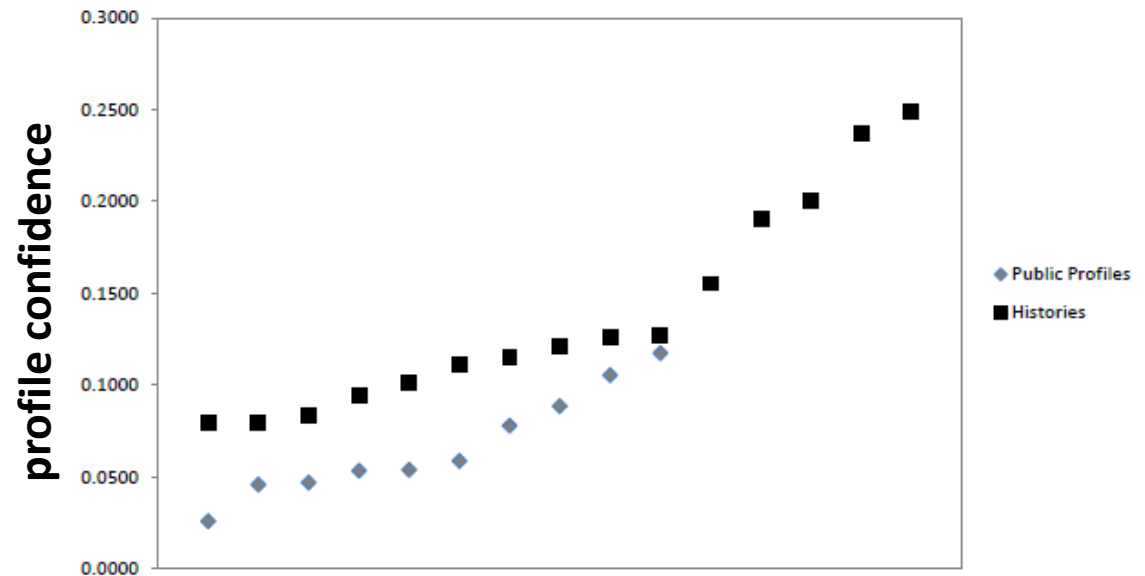
Global Mining Convergence



RePriv vs. the White Pages



Source:
WebMii.com



RePriv

- An in-browser framework for collecting & managing personal data to facilitate personalization.

Core Behavior Mining

- Efficient in-browser behavior mining & controlled dissemination of personal data.

RePriv miners

- A framework for integrating verified third-party code into the behavior mining & dissemination of RePriv.

Real-world Evaluation

- Evaluation of above mechanisms on real browsing histories & two in-depth case studies.

Verifying Miners

```
~/src/f9/browser-sec/SecxIE/FinIE
s --genIL ../lib/Preamble.f9 ../lib/DOM.f9 ../lib/ChromeCompat.f9 ../FacePalm.f9
t-arguha@A2734695 ~/src/f9/browser-sec/SecxIE/FinIE
$ make bin/FacePalm.dll
cd bin && \
    fine --noss --keyfile ../../keys.snk --dotnet4 --partiallytrusted-caller
s --genIL ../lib/Preamble.f9 ../lib/DOM.f9 ../lib/ChromeCompat.f9 ../FacePalm.f9
ERROR: (../FacePalm.f9(72.2)-(72.55)): Expected an expression of type:
{x:elt | DOM.CanReadValue ((x) : elt )}
but got:
{x:elt | DOM.EltParent ((gensym_72_1) : elt ) ((x) : elt )}
Type checking failed: ../FacePalm.f9
Proof obls for module FacePalm = 5
Proof obls for module DOM = 4
t-arguha@A2734695 ~/src/f9/browser-sec/SecxIE/FinIE
$
```


assume ExtensionId "twitterminer"

assume CanCommunicateXHR "twitter.com" Nil

assume CanUpdateStore("twitter.com" "twitterminer")

Netflix Example

- `let doGetMovies genre cdom =`

`...`

`let flixEnts = GetStoreEntriesByTopic
myprov "movie" in`

`let genreFlix = bind myprov flixEnts
(filterByGenre genre) in`

`ExtensionReturn cdom myprov genreFlix`

ass
ass
ass
ass
ass
ass
assu

`assume CanReadDOMClass "netflix.com" "rv5"`

`assume CanCaptureEvents "onclick" (P "netflix.com" "netflixminer")`

`assume CanServeInformation "fandango.com" (P "netflix.com" "netflixminer")`

`assume CanServeInformation "amazon.com" (P "netflix.com" "netflixminer")`

`assume CanServeInformation "metacritic.com" (P "netflix.com" "netflixminer")`

`assume CanHandleSites "netflix.com"`

`assume CanReadStore (P "netflix.com" "netflixminer")`

`assume CanReadLocalFile "moviegenres.txt"`

14 lines of Fine code



RePriv

- An in-browser framework for collecting & managing personal data to facilitate personalization.

Core Behavior Mining

- Efficient in-browser behavior mining & controlled dissemination of personal data.

RePriv miners

- A framework for integrating verified third-party code into the behavior mining & dissemination of RePriv.

Real-world Evaluation

- Evaluation of above mechanisms on real browsing histories & two in-depth case studies.

Privacy-Aware News Personalization

Map RePriv intereststo del.icio.us topics



```
graph TD; A[Map RePriv intereststo del.icio.us topics] --> B[Query personal store for top interests]; B --> C[Ask del.icio.us API for “hot” stories in appropriate topic areas from nytimes.com]; C --> D[Replace nytimes.com front page with del.icio.us stories];
```

The diagram consists of four rectangular boxes arranged in a descending staircase pattern from top-left to bottom-right. Each box contains a step in the process. The boxes are colored in a gradient: the top box is purple, the second is dark blue, the third is medium blue, and the bottom box is light blue. White arrows point downwards from the bottom-right corner of one box to the top-right corner of the next box, indicating the flow of the process.

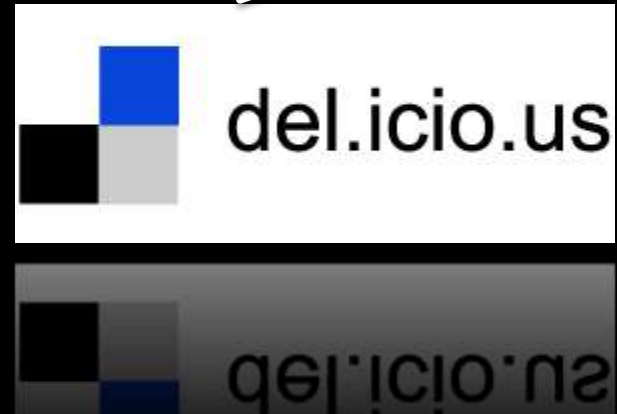
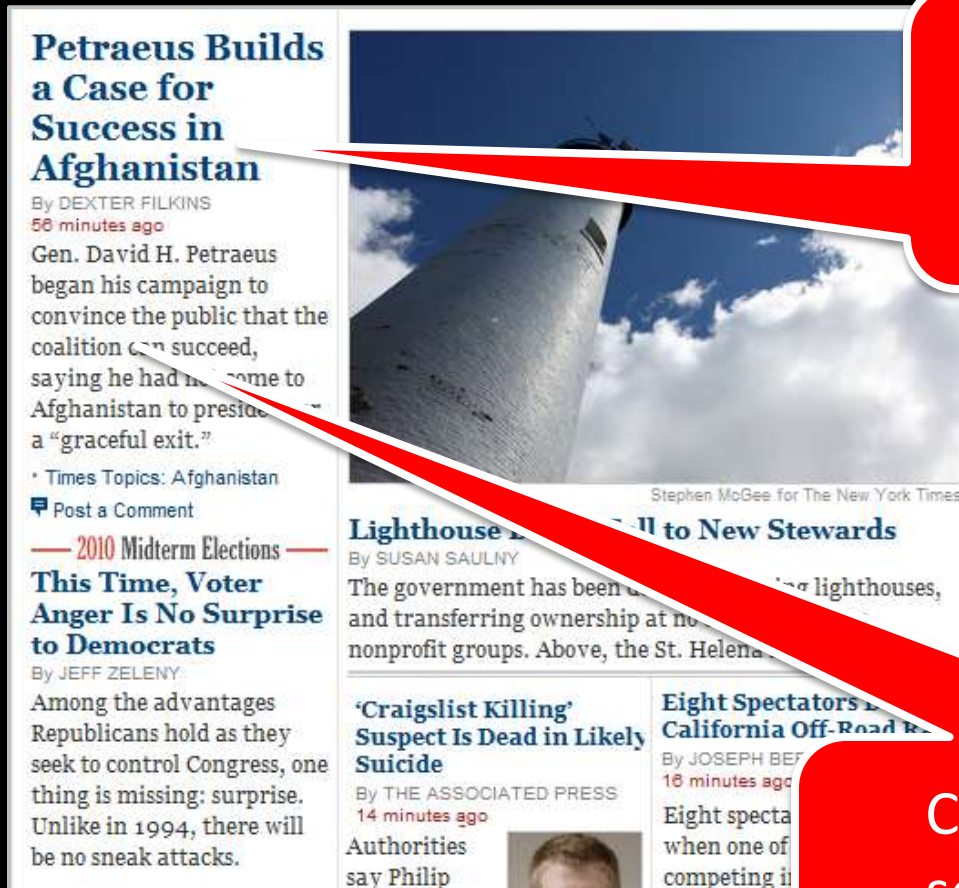
Query personal store for top interests

Ask del.icio.us API for “hot” stories in appropriate topic areas from nytimes.com

Replace nytimes.com front page with del.icio.us stories

Privacy Policy

Change “href” attribute of anchor elements on nytimes.com



Change TextContent of selected anchor and div elements on nytimes.com

User profile:

- Games/Card_Games
- Games/Conventions
- Games/Video_Games

Do Video Games Equal Less Crime?

That's one theory for the continuing fall in crime, despite the recession.

Gamers Finally Get Their Wheaties Box ...sort of

Dr Pepper is featuring the Halo 3 player Tom Taylor, who goes by Tsquared, on the labels, which will appear on about 175 million 20-ounce bottles from January to April.

Sony To Shut its SF Metreon PlayStation store

Sony is closing down its one-and-only U.S. PlayStation store at the Metreon mall in San Francisco. The recession is clearly to blame, but it's happening at time when Microsoft - which opened and shut its own Microsoft store at the Metreon - is going to open a chain of its own stores.

Microsoft Takes on Cable With Xbox Streaming Video

If talks with Disney work out, the game console could stream ESPN content, making it that much easier to watch TV without cable.

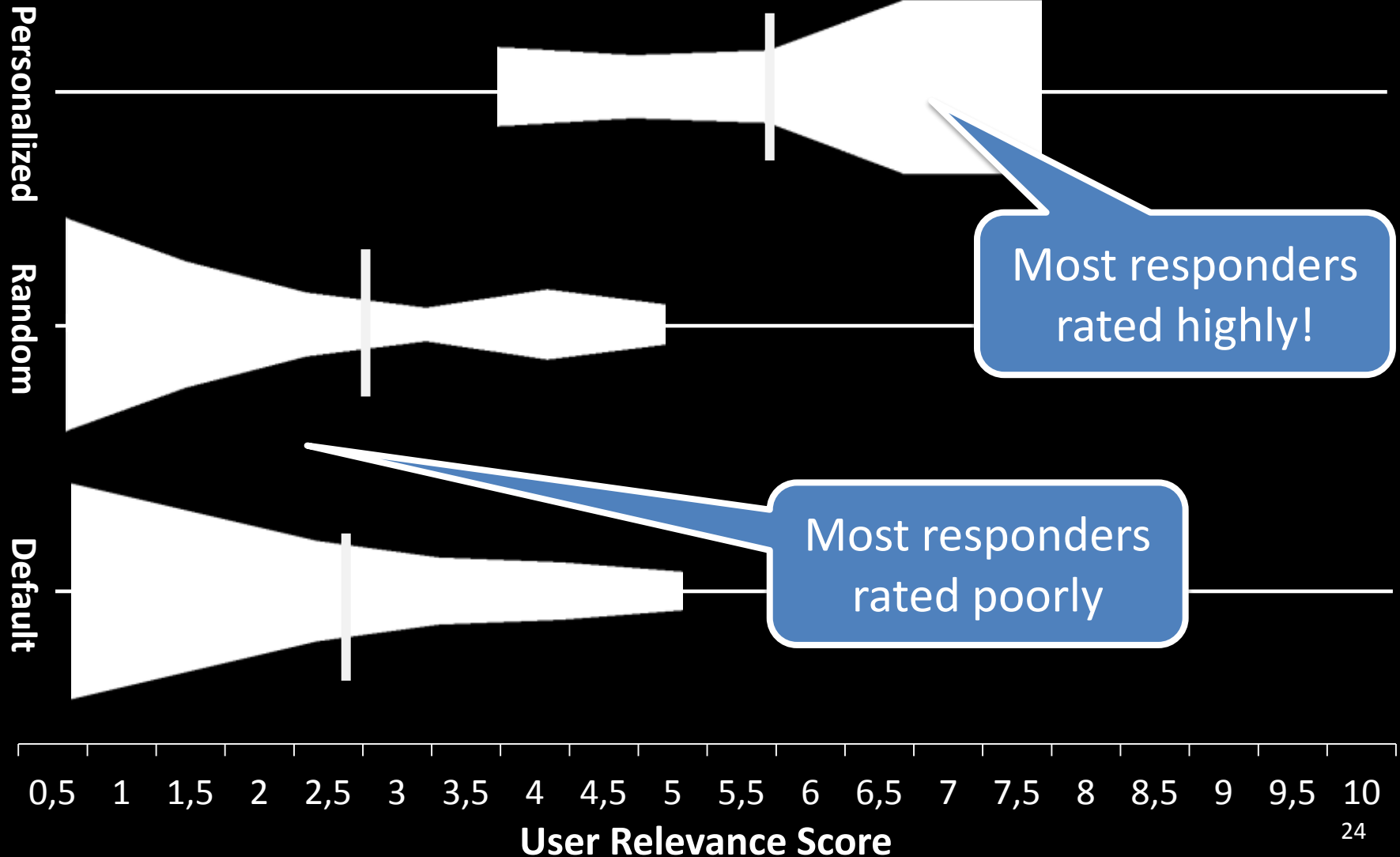
Some Video Gamers Leery of Obama's Views

Gamers are worried that the president-elect's positions on video games may signal new regulations or restrictions on the industry.

Relevance: (required)



News Personalization: Effectiveness



Non-determinism

JavaScript

Record-replay systems

JSBench:

Next-generation JavaScript Benchmarks

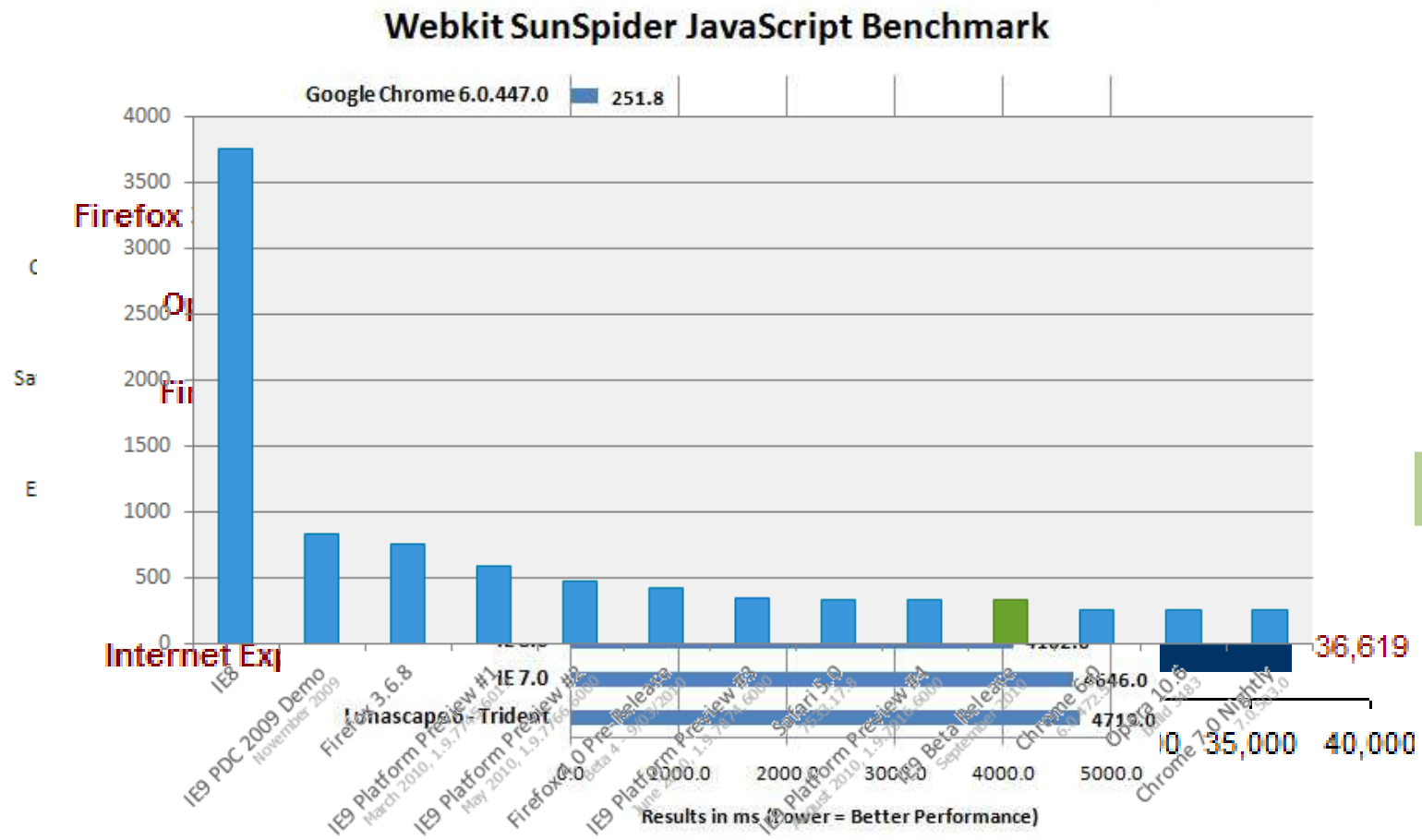
Browsers

Performance metrics

JavaScript Performance. Why?

- Why JavaScript:
 - Language of the web
 - Support in every browser
 - Much existing code
- Browser and JavaScript performance is important
 - Limited understanding of JavaScript behavior in real sites
 - Are current JavaScript benchmarks representative?

SunSpider... Your Results May Vary



12/2007

6/2008

2/2008

6/2010

9/2010

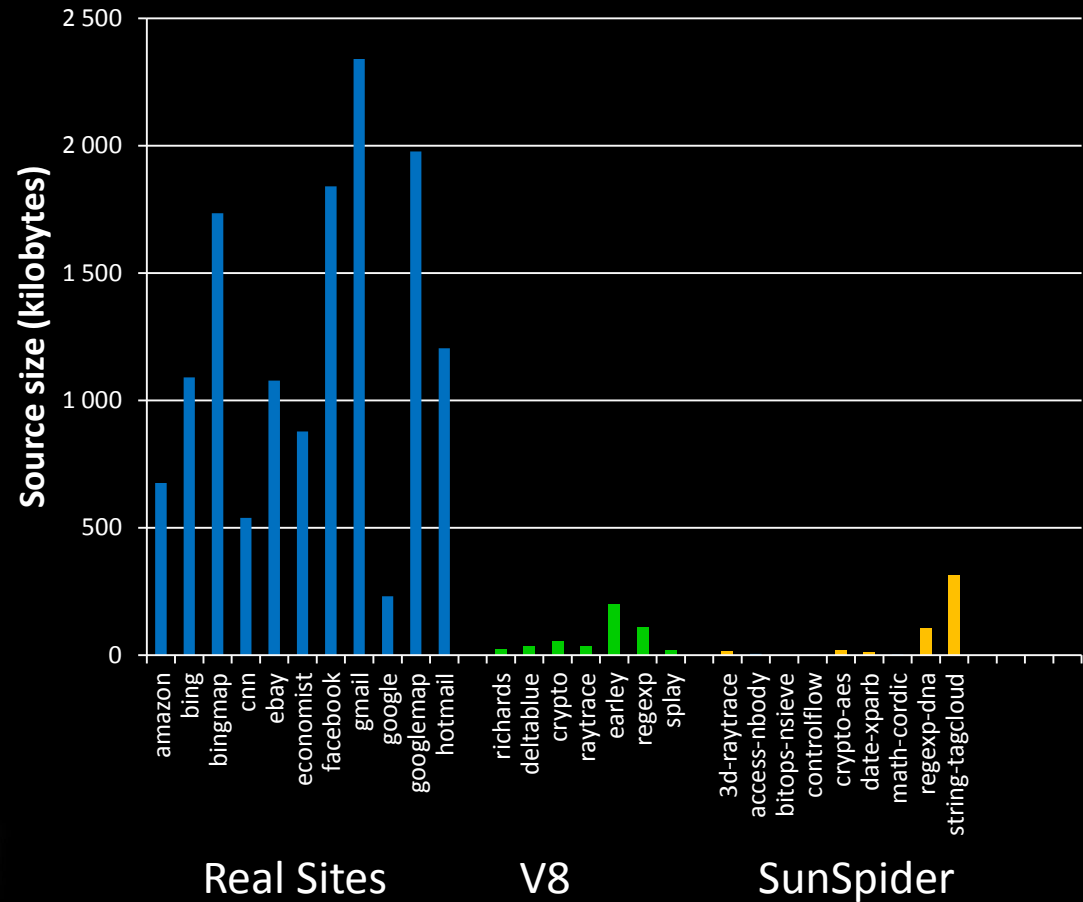
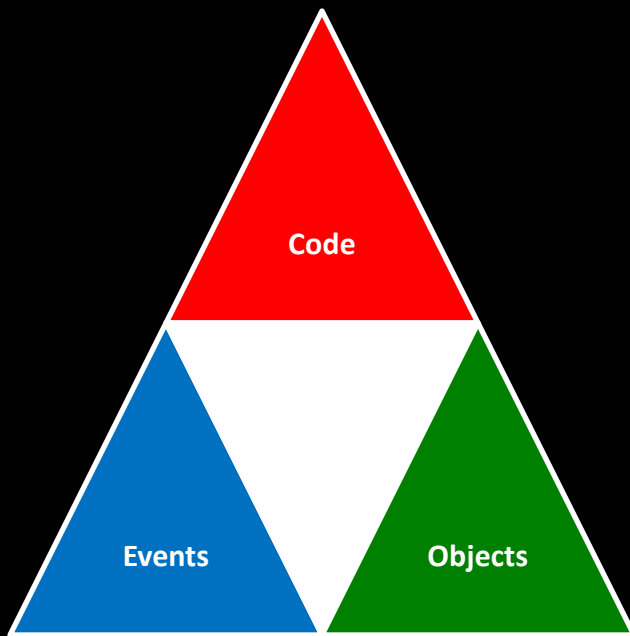
Landscape of JavaScript Benchmarks

SunSpider (Apple)

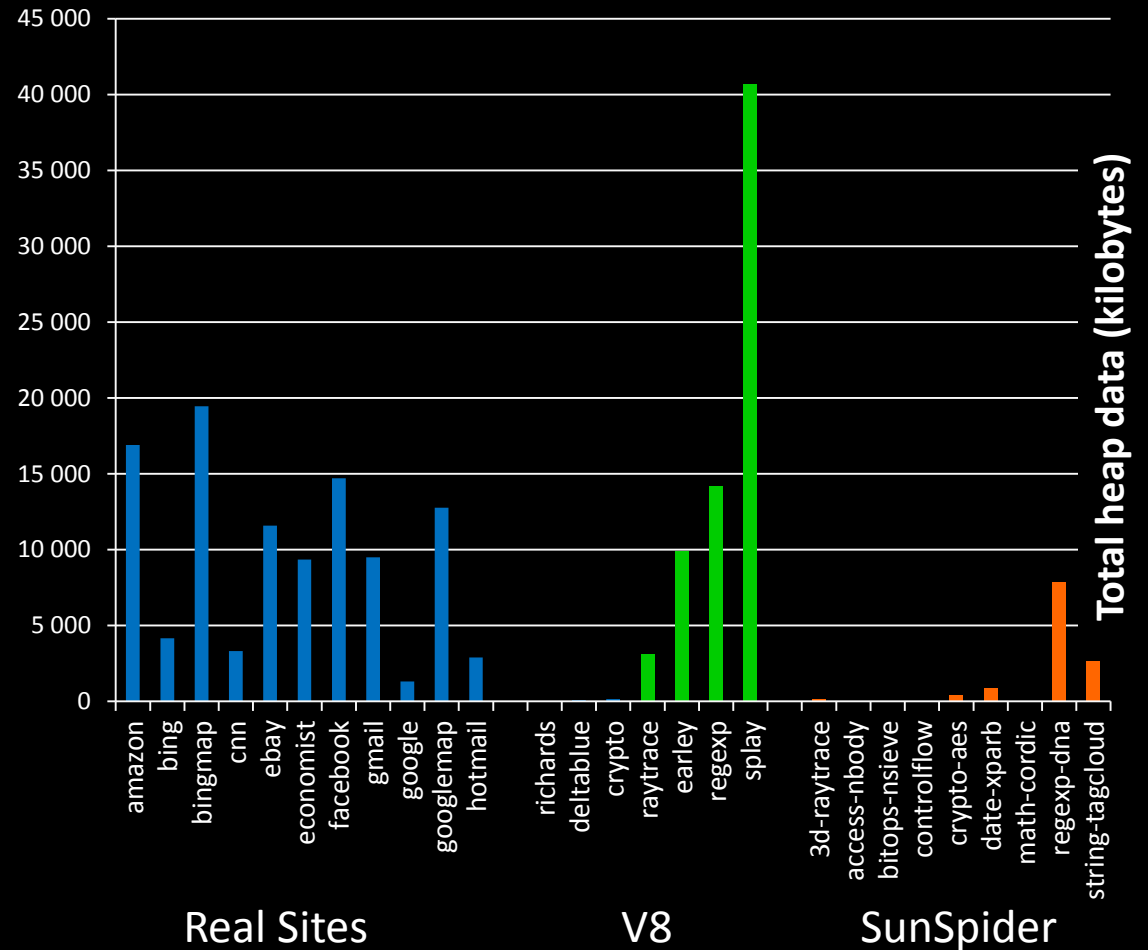
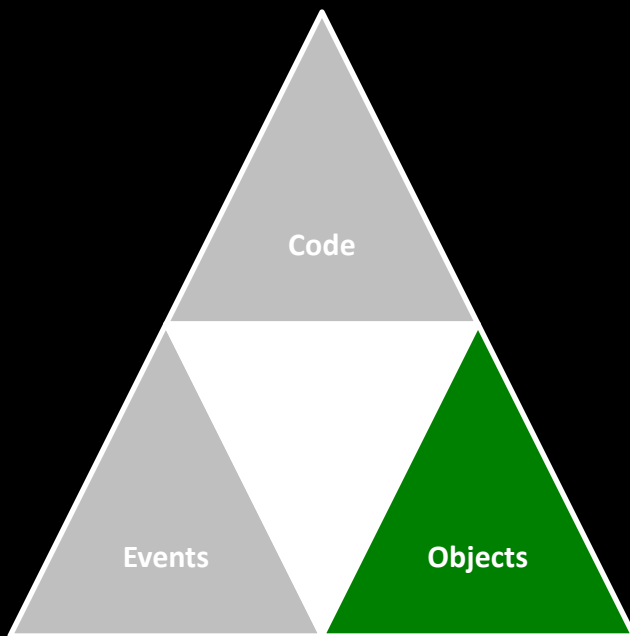
- 3-draytrace
- access-nbody
- bitops-nsieve
- controlflow
- crypto-aes
- date-xparb
- math-cordic
- string-tagcloud



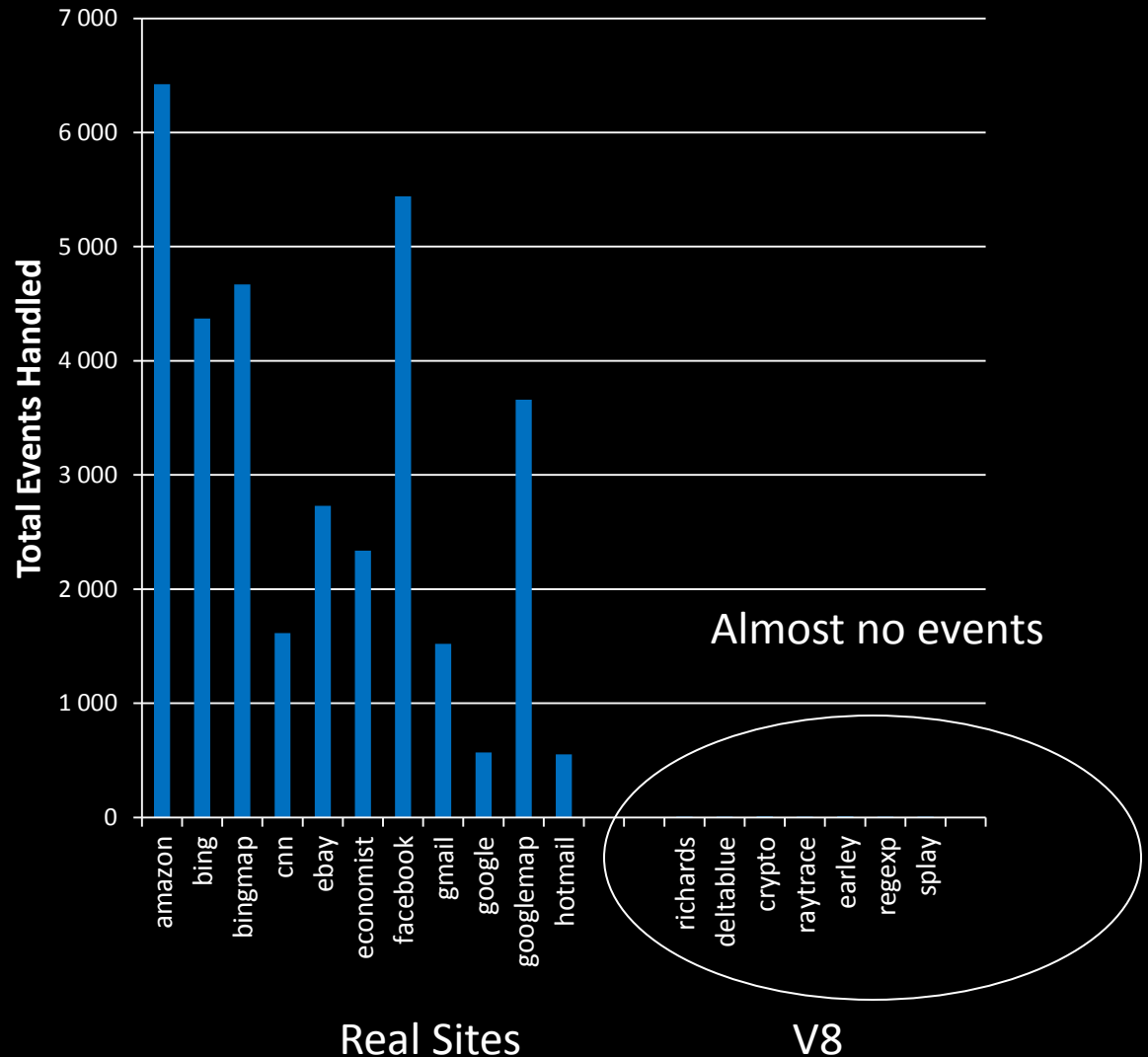
Total Bytes of JavaScript Source



Total Bytes Allocated



Total Events Handled

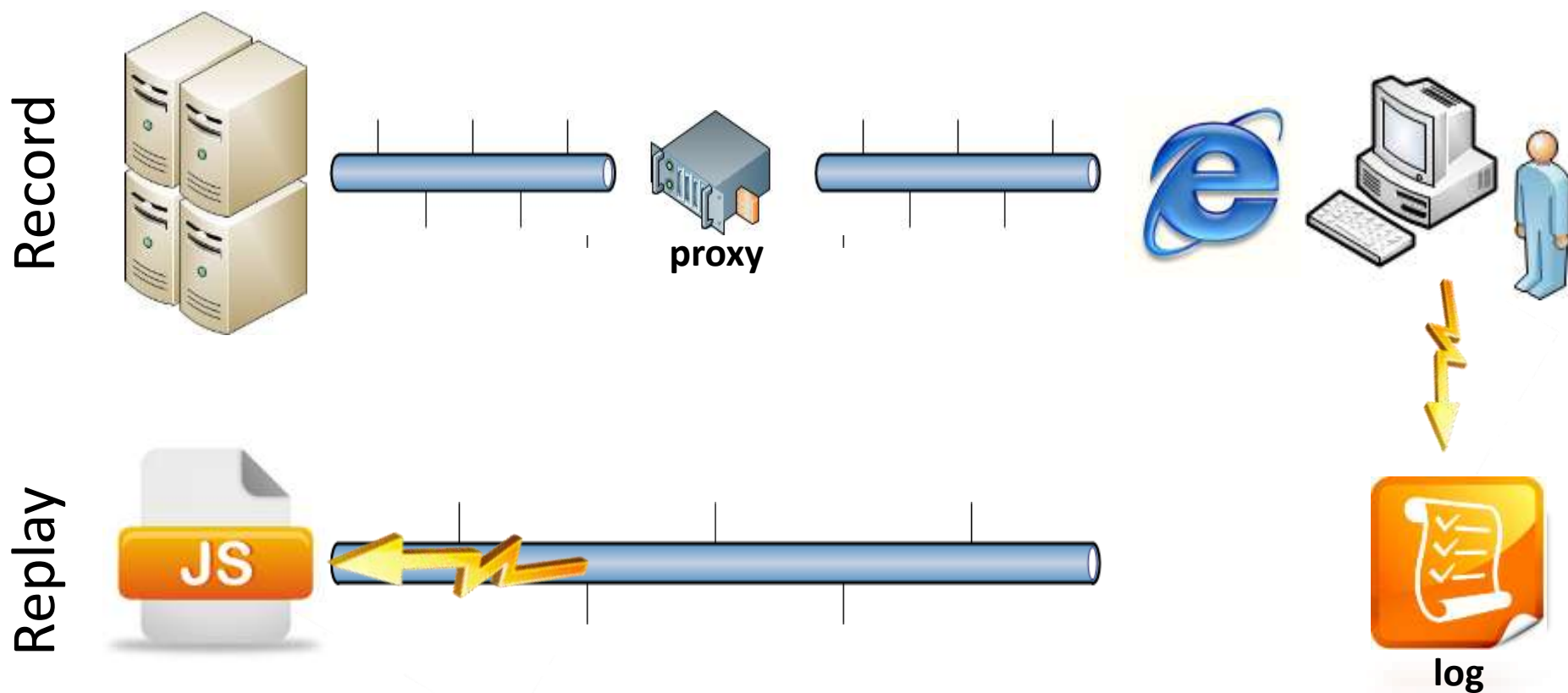


Events

Code

Objects

JSBench Record and Replay in a Nutshell



2. Create and replay a different scenario in a new stage

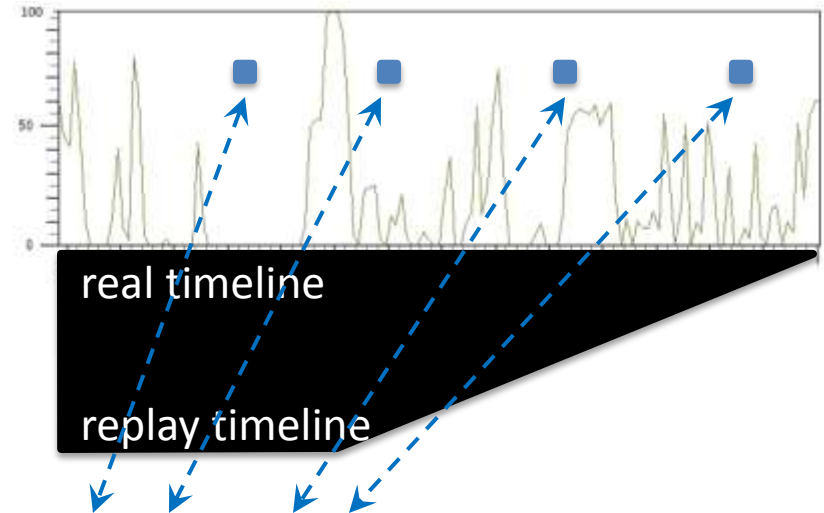
Code -> Record -> Replay

```
function onloadHandler() {
    myvalue = document.getElementById("input").value;
    obj[1] = {}; // 1
}
window.document.addEventListener("load", onloadHandler, false);
func[1] = function() { ... } // 3
obj[1].getElementById = func[1]; // 3   JavaScript
```

obj[2] = {}; // 4	onload event	Log handler	onloadHandler()
func[1].case[obj[1], "input"] = obj[2]; // 4			↓
obj[2].value = "Hello, JavaScript!"; // 5	Inspect document	Log arguments	document.getElementById("input")
onloadHandler.call(window); // 6			
}	↓		
MEMBER window.document	Document object	Log & wrap object	DOM Object
EVENT onloadHandler(window)			obj[1]
MEMBER obj[1].getElementById	Inspect element		↓
RETURN func[1](obj[1], "input")			func[1].value
MEMBER obj[2].value	Get value	Log value	obj[2]
			"Hello, JavaScript!"

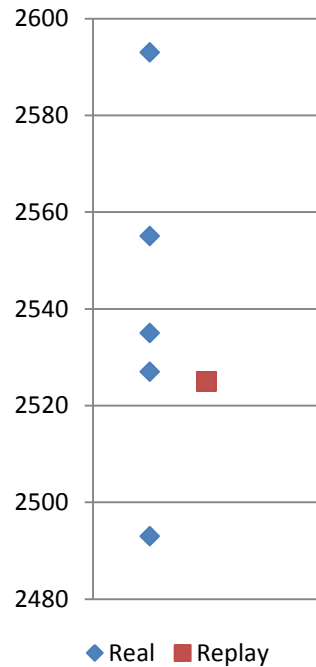
How Accurate Are Our Benchmarks?

- Instrumented browsers for event collection
 - IE8, IE9 have ETW instrumentation
 - Firefox team added instrumentation for us as well
- Add instrumentation to JS code to report events of interest



JavaScript Fidelity

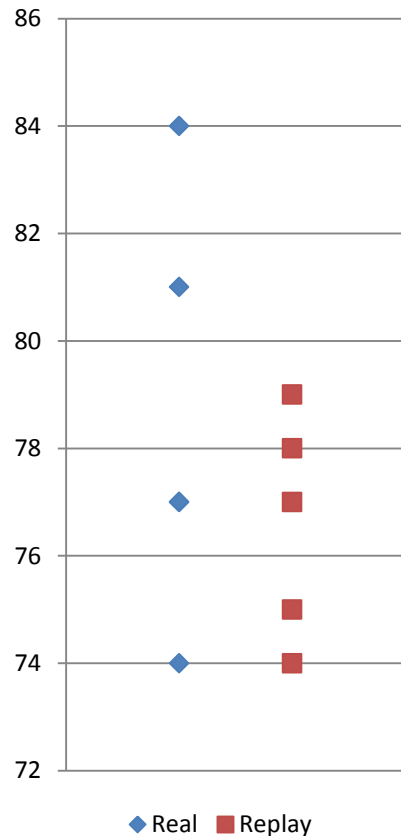
- Run Amazon real and replay, see how JS behaves
- Multiple runs to get range of possible behaviors



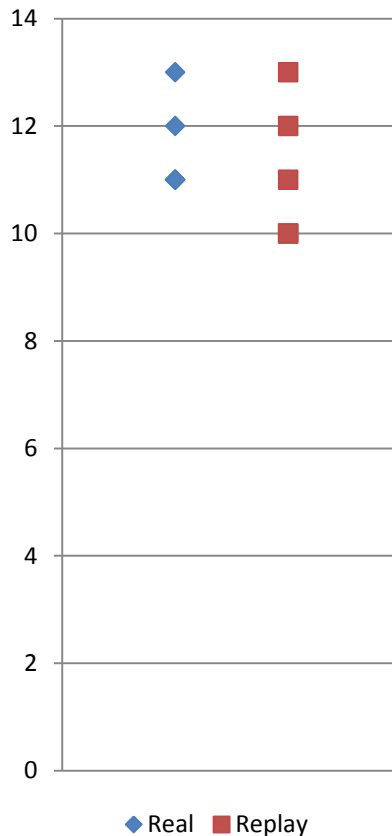
Calls in

Browser Fidelity

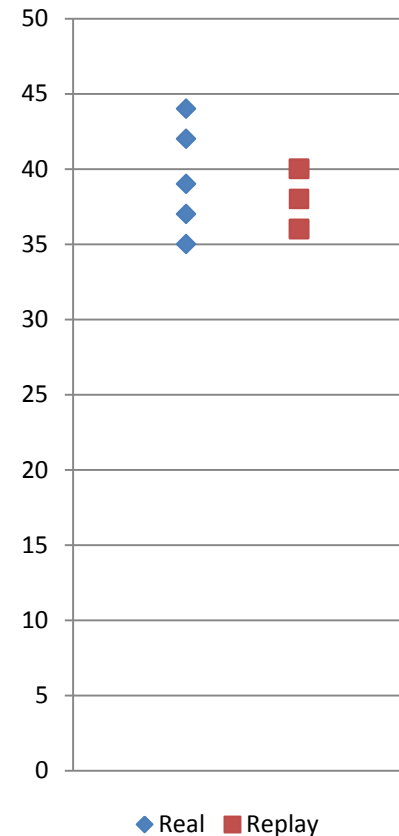
- Run Microsoft real and replay, see how browser behaves



Layout



Paint

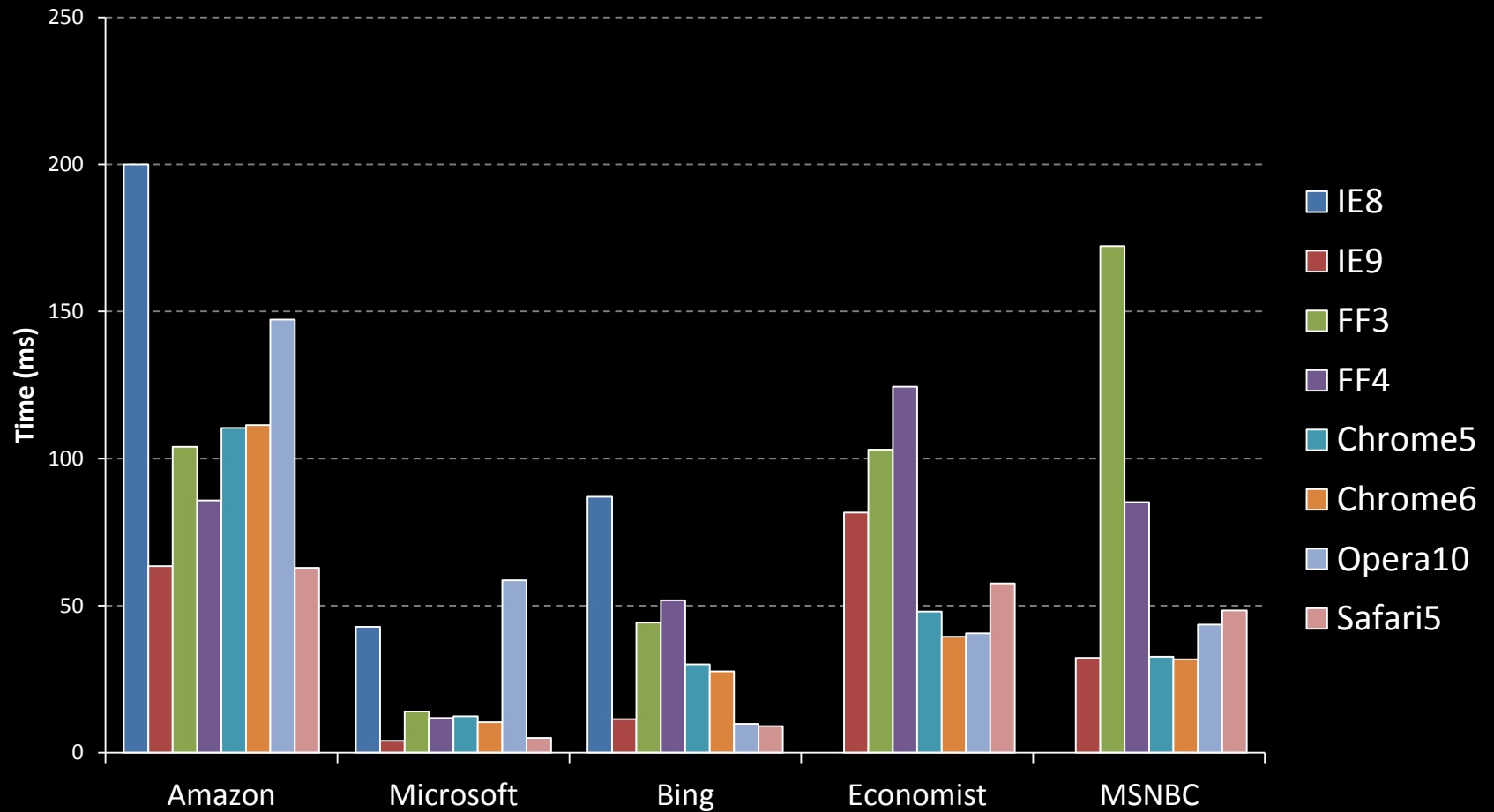


CSS calculations

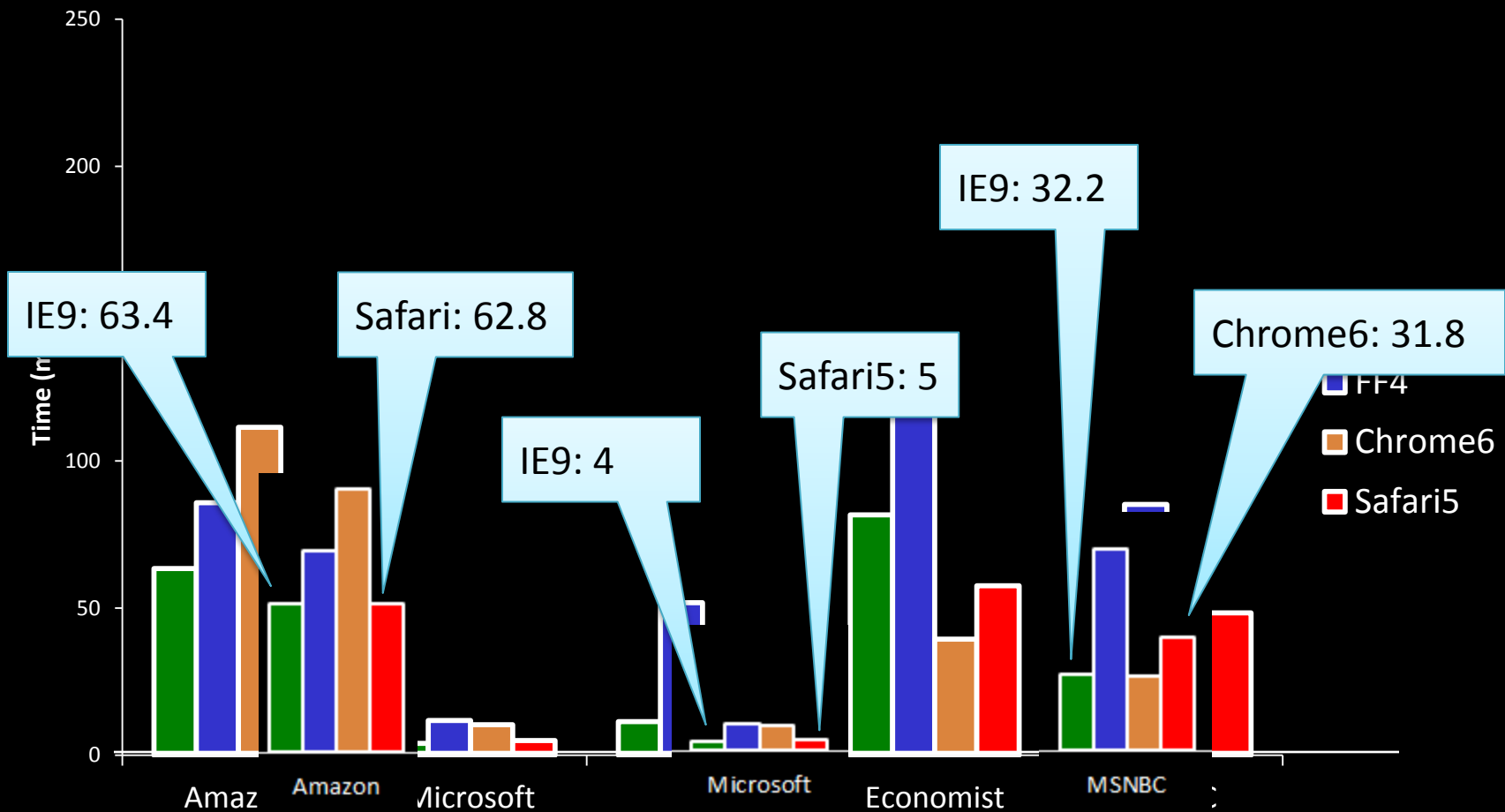
Demo

(removed)

Cross-browser Performance Comparison



Cross-browser Performance Comparison



- Towards better JavaScript benchmarks

- Improve on the current state-of-the art of SunSpider & co
- Do so through a record-replay strategy on real sites

- Show

- How to do this in practice
- Present fidelity results

- A collaboration with Mozilla (Firefox) and academia

- Personalized content & privacy can coexist

Microsoft® Research



Ben Livshits

Microsoft Research

Redmond, Washington

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