

**Privacy and Security for Browser Extensions:
A Language-Based Approach**

Ben Livshits

Microsoft Research

Redmond, Washington

Microsoft[®]
Research

- Provide missing functionality
- Faster evolution than browsers
- Embed themselves into browser
- ...which has security implications

Language-based foundations

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:

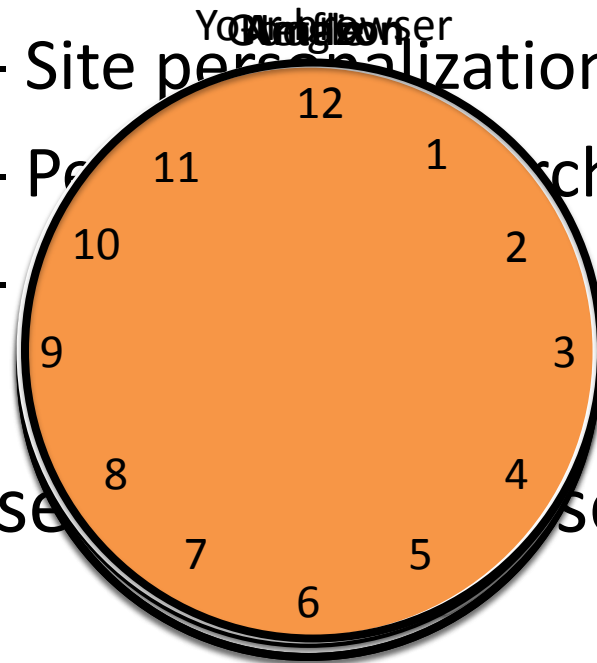
– Site personalization

– Product recommendation

– Search

- User

- Control data release



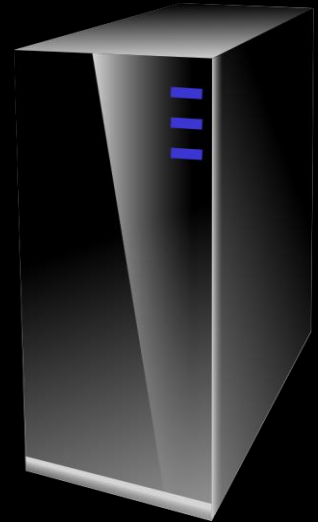
bn.com would like to learn your top interests.
We will let them know you are interested in:

- 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

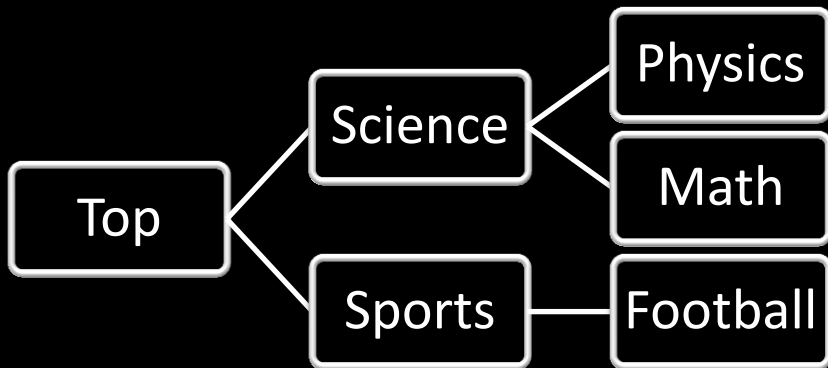
- A framework for integrating verified third-party code into the behavior mining & dissemination of RePriv.

Real-world Evaluation

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

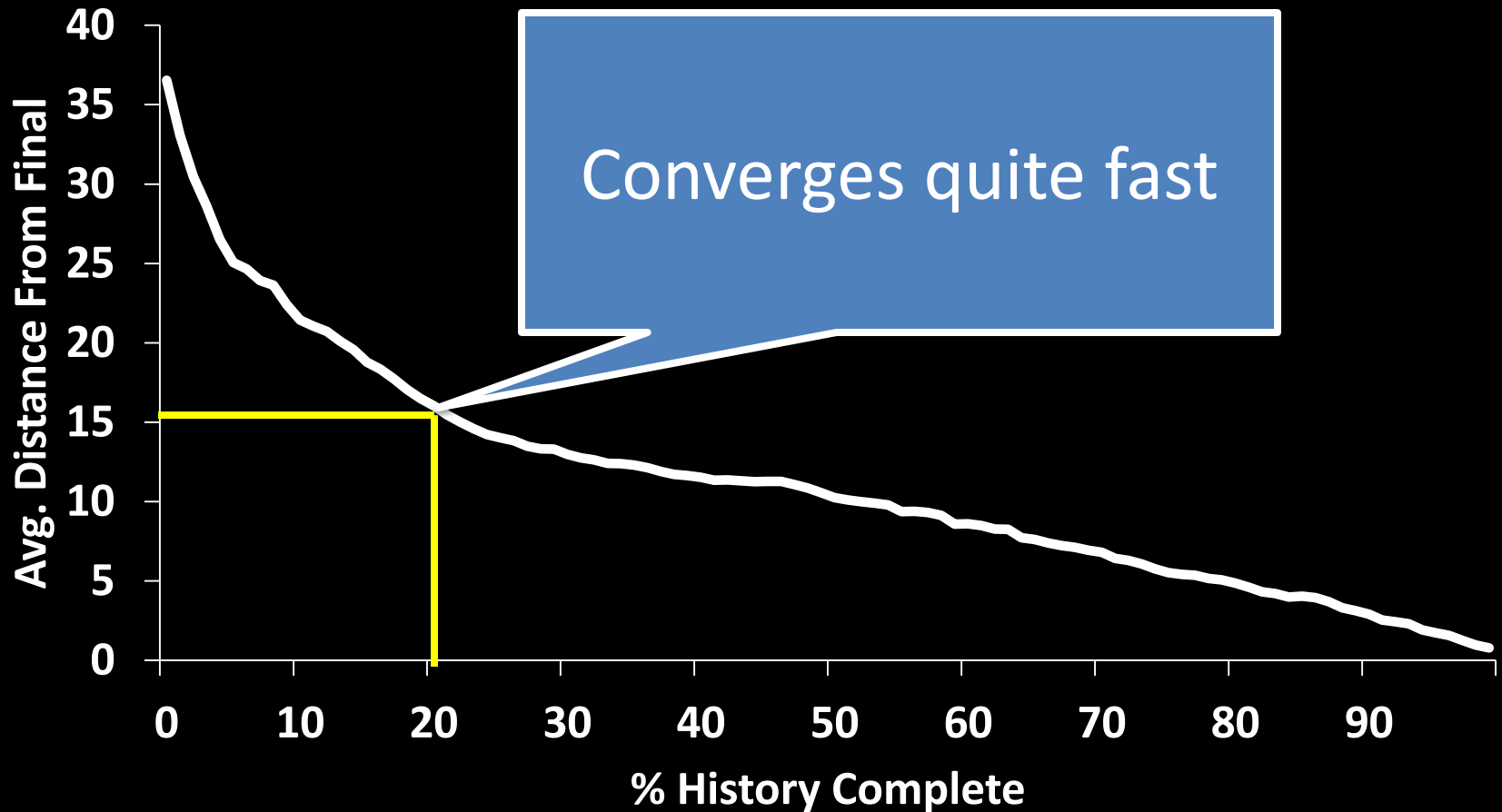
Core Mining

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

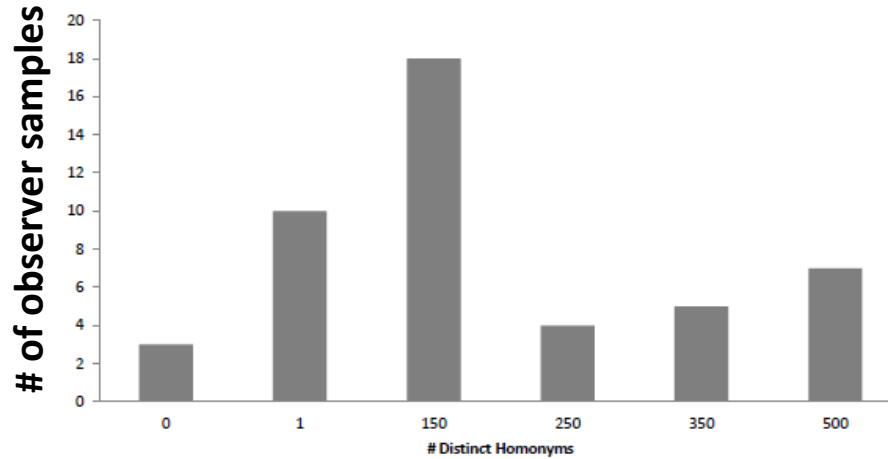


- Naïve Bayes
 - All categories equally likely
 - Training: min(3000, # pages) sites per category
 - Attribute words occur in at least 15% of docs for ≥ 1 category
- Classification is fast enough: $O(c \cdot n)$
 - n is # words in document
 - c is # document categories

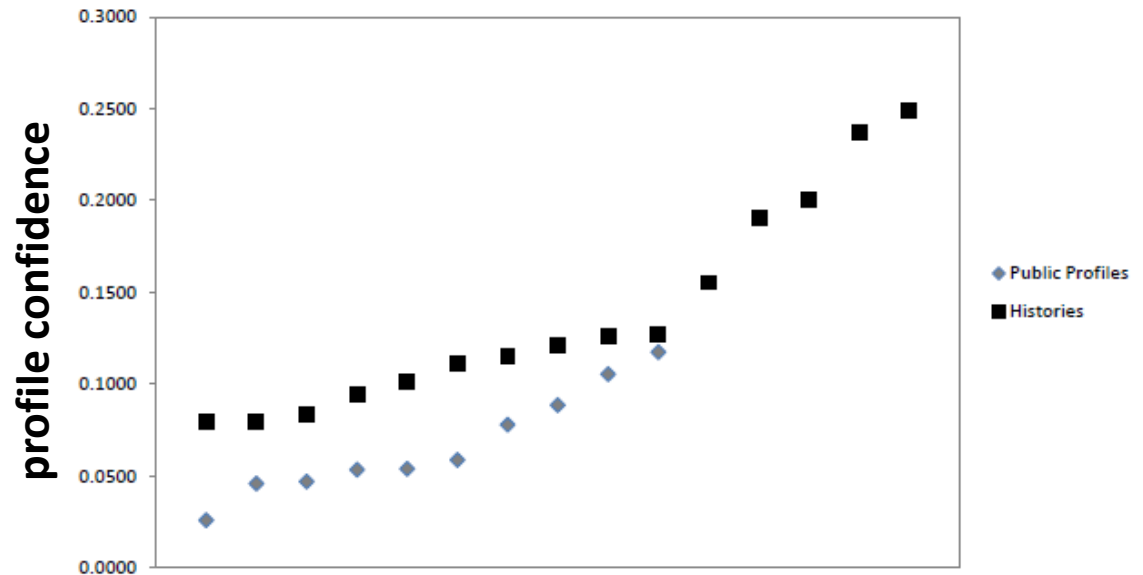
Global Mining Convergence



RePriv vs. the White Pages



Source:
WebMii.com



RePriv Miners **VERIFIED**
^

Miner Verification Strategy

```
~/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
$
```


REAL-WORLD

EXPERIMENTAL EVALUATION

Privacy-Aware News Personalization

Map RePriv intereststo del.icio.us topics



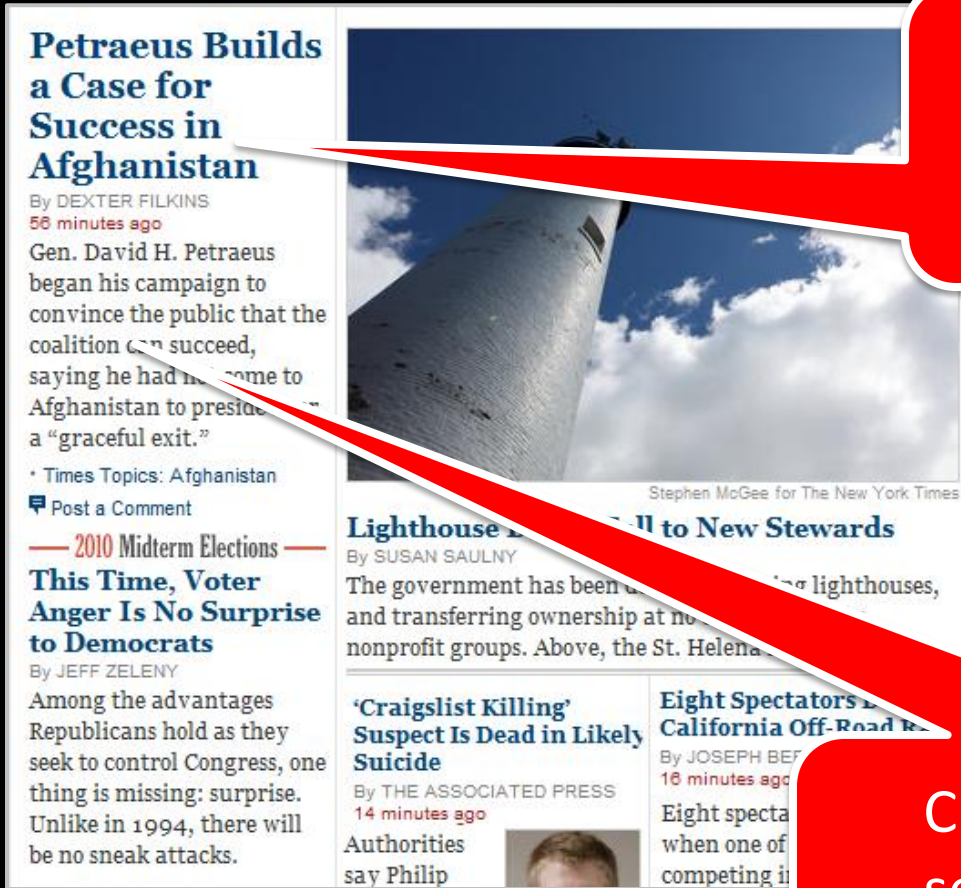
```
graph TD; A[Map RePriv intereststo del.icio.us topics] --> B[Query personal store for top interests]; B --> C[Ask del.icio.us API for "hot" stories in appropriate topic areas from nytimes.com]; C --> D[Replace nytimes.com front page with del.icio.us stories];
```

Query personal store for top interests

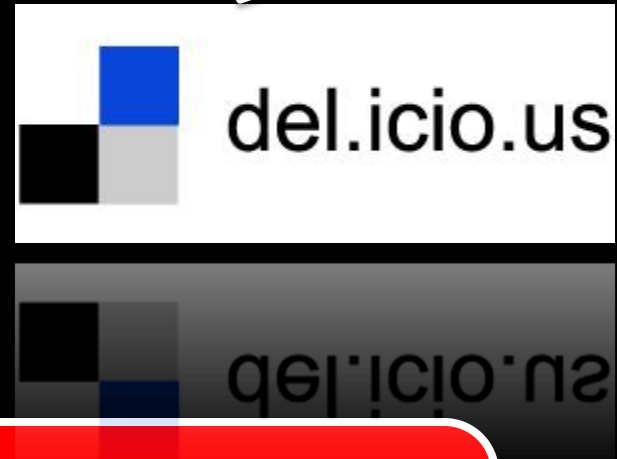
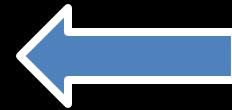
Ask del.icio.us API for “hot” stories in appropriate topic areas from nytimes.com

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.

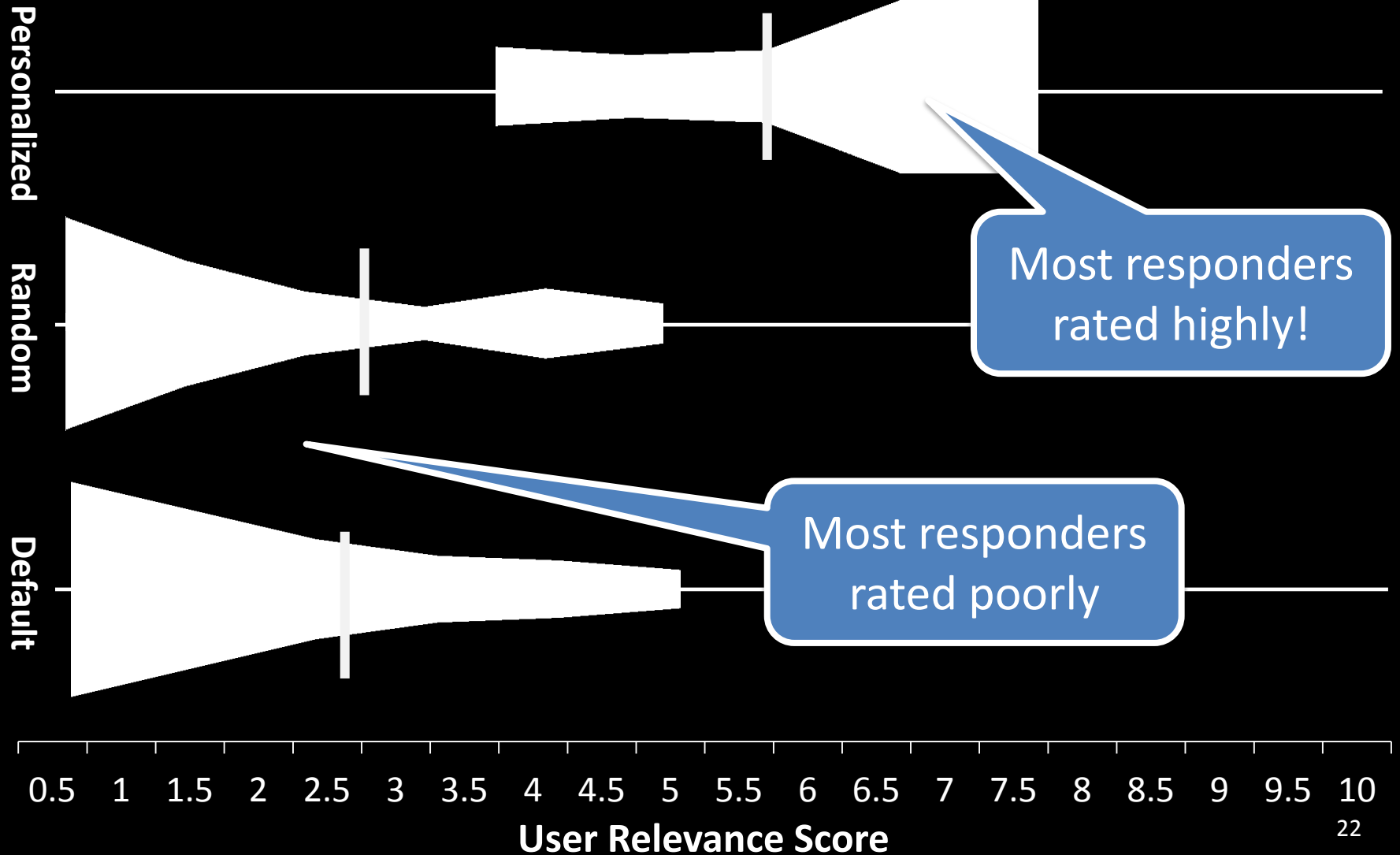
Relevance: (required)

Technology
Technology
Science/C
Science/P

- 2,2
- Ov
- Typ



News Personalization: Effectiveness



Fine

JavaScript

ML

Verified Security for Browser Extensions

Type systems

Verification and analysis

```
"update_url":"http://clients2.google.com/service/...",
"name": "Twitter Extender", "version": "2.0.3",
"description": "Adds new Features on Twitter.com ",
"page_action": { ... }, "icons": { ... }, \\
"content_scripts": [ {
"matches": [
"http://twitter.com/*", "https://twitter.com/*"],
"js": ["jquery-1.4.2.min.js","code.js"]
} ],
"background_page": "background.html",
"permissions": [ "tabs", "http://api.bit.ly/" ]
```

1,139 popular
Chrome extensions



Permission	#	%
all https	143	12%
all http	199	17%
wildcard *	536	47%
history (tabs)	694	60%

60% of all extensions are grossly over-privileged (access to complete history)



Ghostery by Ghostery

★★★★☆ (70) - 18,822 users - Weel


Protect you

Mel Jun 26, 2010
Yo dawg, I heard yo
tracking you so you

DETECT:

Ghostery sees the "invisible" web, detectin
beacons placed on web pages by ad netwo
publishers, and other companies interested in your activity

Confirm Installation



Install Ghostery?

This extension needs access to:

- Your data on all websites
- Your browsing history

Install Cancel

InPrivate Filtering
available on other

ck who's

PROTECT YOUR PRIVACY:

Ghostery is built and maintained for users that care about their online privacy, and is

engine **James** Jul 19, 2010 [Mark as spam](#)

regist Is this a scam? Shi*, I so nearly downloaded it. Why does it have so many good reviews then?

block imag

COLLABOR

Ghostery a
Ghostery s
where you
detectable
ecosystem

Jul 8, 2010 [Mark as spam](#)

DO NOT DOWNLOAD!!! SCAM & PHISHING EXTENSION!!!

anonymous Jul 5, 2010 [Mark as spam](#)

this is stoopid

PROTECT

Ghostery is
is engineer
registration
into your br
any data fro

asmp Jul 5, 2010 [Mark as spam](#)

This extension is dangerous... blog.betteradvertising.com/2010/01/19/better-advertising-acquires-ghostery/ man! To think I was almost installing this piece of sh.....

GhostRank data is anonymous, it is NEVER used for advertising targeting

Moral: security manifests rendered useless by permissively over-privileged extensions

Contributions

Study of Chrome extensions

- Large-scale study of >1,000 Chrome extensions
- Analyze their manifests for security privileges
- Conclude that many or most are over-privileged

Datalog-based policies

- Policy language based on Datalog for specifying fine-grained authorization and data flow policies
- Visualization tools to “apply” authorization policies to web pages

Semantics of policies and extensions

- Formalize the semantics of security policies and extensions in an execution model with arbitrary interleavings
- Security property, (L;P)-safety, suitable for use with extensions that interact with other, untrusted code

Extensions implemented

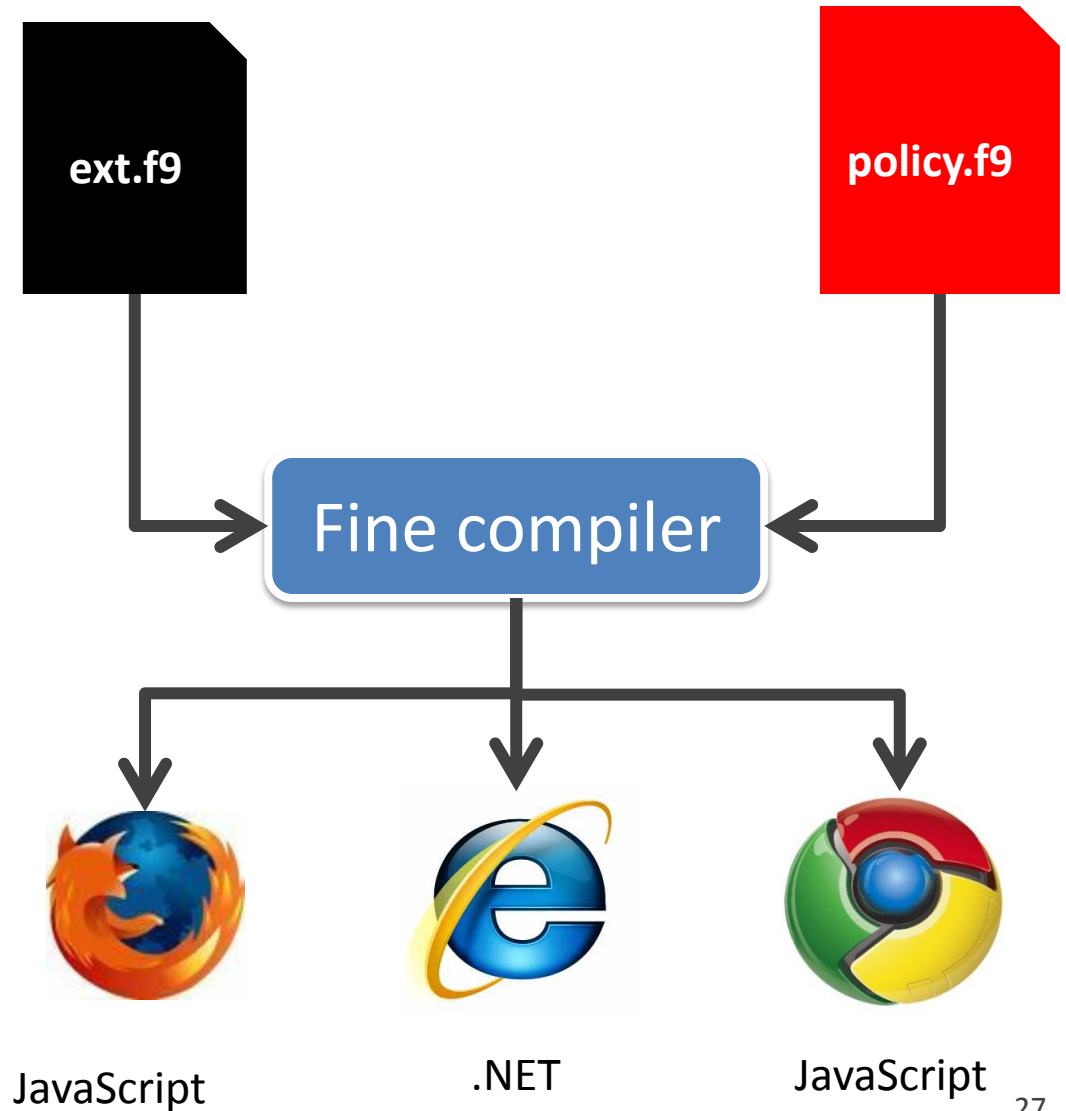
- Programming 17 extensions in Fine covering a range of fine-grained authorization and information flow properties for each, and automatically verifying for policy compliance
- These include several widely-used Chrome extensions, showing that our model brings benefits to legacy extension architectures

Retargeting to multiple browsers

- Extend Fine compiler with a code generator that emits JavaScript (in addition to .NET bytecode)
- Extension development in a platform-independent way, allowing for deployment in IE 8, Chrome, Firefox, and C3

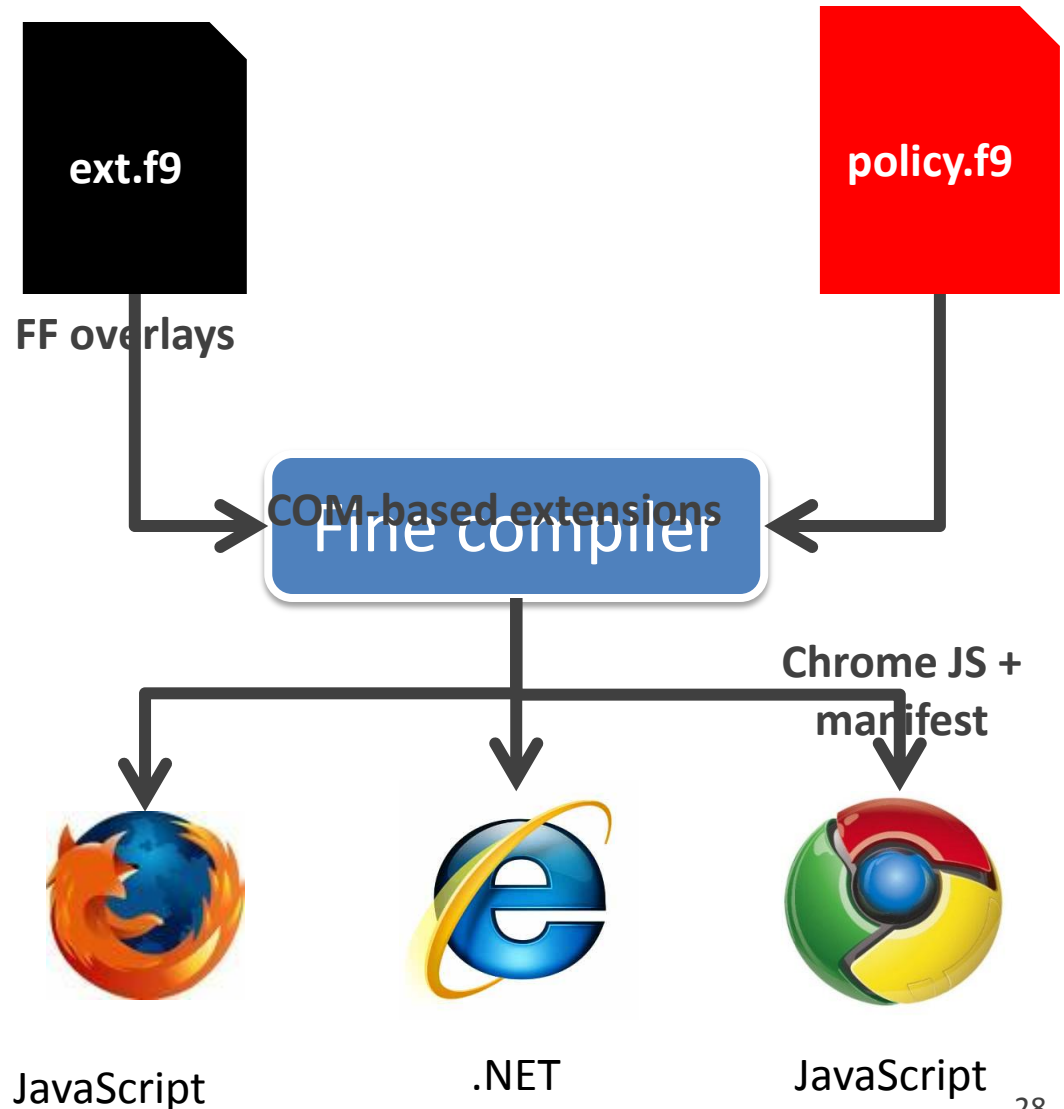
Our Goals

- Explicit and expressive **policy language**
- Automatic **verification**
- **Retargeting** for multiple browsers



Contrast

- Curating process
 - Arbitrary
 - Too permissive
 - Time-consuming
- Difficult to port
- Code in C/C++ or JavaScript is difficult to check



SECURE
EXAMPLE: FACEBOOK EXTENSION

Facebook | Jean Yang - Windows Internet Explorer

http://www.facebook.com/profile.php?id=35107474&ref=ts#!/profile.php?id=11544&v=info&ref=ts

Facebook | Jean Yang

facebook Search Home Profile Account

Jean Yang (杨龄晶) discovered that O.N.E.'s flavored coconut juices taste like what's left at the bottom of bad mixed drinks. Once all the ice has melted... On a more positive note, I'm back in Boston in exactly one week! 18 hours ago

Info Photos Boxes Wiki Video

Contact Information

Contact Info

Email: jean.yang.writeme@gmail.com
jeanyang@mit.edu

Mobile Phone:

Website: <http://people.csail.mit.edu/jeanyang>
<http://jxyzabc.blogspot.com>
<http://gsc.mit.edu/gwamit>

Harvard Alumni
Google

Birthday: October 11
Current City: Seattle, WA

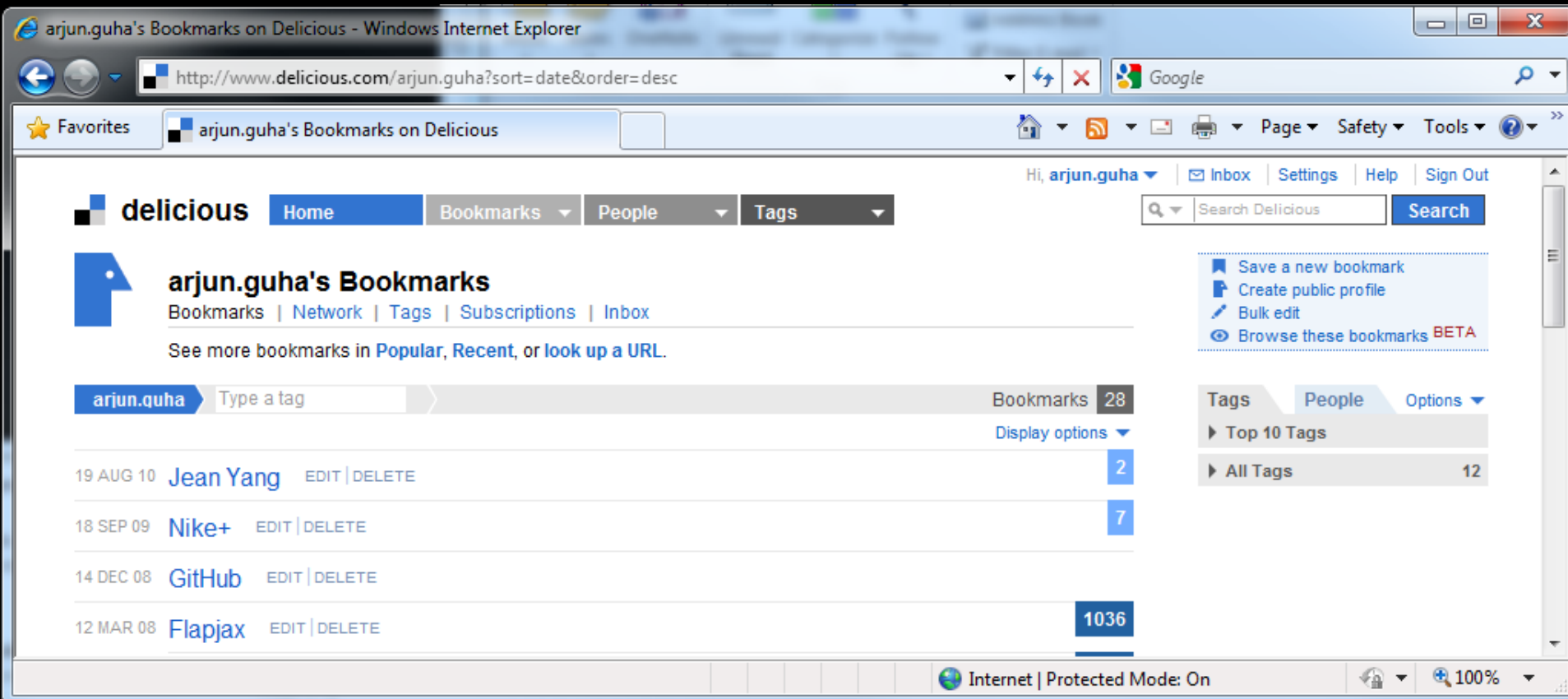
I love vi... at.

"One... think well, love well, sleep well, if one has not dined well. The lamp in the hum... does not light on beef and prunes." -Virginia Woolf, A Room of One's Own

Stone Age, Classical Age, Modern Age... and beyond!
Advance through the ages of time and build a thriving Civilization! Play City of

<https://api.del.icio.us/y.../posts/add?url=http://people.csail.mit.edu/jeanyang&description=Jean+Yang>

http://www.facebook.com/?ref=logo Internet | Protected Mode: On 100%



```
https://api.del.icio.us/v1/posts/add?  
url=http://people.csail.mit.edu/jeanyang&  
description=Jean+Yang
```

[View Photos of Jean \(584\)](#)[View Videos of Jean \(2\)](#)[Send Jean a Message](#)[Poke Jean](#)

My goal is to be more Indian than Rishabh Singh.

Jean Yang (杨龄晶) discovered that O.N.E.'s flavored coconut juices taste like what's left at the bottom of bad mixed drinks after all the ice has melted... On a more positive note, I'm back in Boston for the weekend 18 hours ago

[Wall](#) [Info](#) [Photos](#) [Boxes](#) [Wiki](#) [Videos](#)

About Me

Basic Info	Birthday:	October 1
	Parents:	Jie Yang
	Siblings:	Rishabh Singh
	Current City:	Seattle, Washington
	Political Views:	Very Liberal
	Religious Views:	Morristect

Bio "Would you tell me, please, which way I ought to go from here?"
"That depends a good deal on where you want to get to," said the Cat

getName and
getWebsites
do not exist. ..

:-)

```
let name = document.getName() in  
let website = document.getWebsites()[0] in  
...
```

Mutual Friends

2 friends in common

[See All](#)

Nikhil
Swamy



Leo
Meyerovich

Work and Education

Employers

Microsoft Research May 2010 - Present
Research Intern
Redmond, Washington
Combining programming languages and cryptography to make the world a better place.

[Like](#)

MBA

[Chat \(11\)](#)

How Do We Pattern-match?

```
<th class="label">Contact Info</th>
▼<td class="data">
  ▼<table class="uiInfoTable noBorder">
    ▼<tbody>
      ▼<tr>
        <th class="label">Email:</th>
        <th class="label">Address:</th>
        ▼<td class="data">
          "jean.yang.writeme@gmail.com"
          "jeanyang@mit.edu"
        </td>
      </tr>
    </tbody>
  </table>
  ▼<tbody>
    ▼<tr>
      <th class="label">Mobile Phone:</th>
      <td class="data">4123026391</td>
    </tr>
  </tbody>
  ▼<tbody>
    ▼<tr>
      <th class="label">Website:</th>
      ▼<td class="data">
        <a href="http://people.csail.mit.edu/jeanyang">http://people.csail.mit.edu/jeanyang</a>
        <a href="http://jxyzabc.blogspot.com">http://jxyzabc.blogspot.com</a>
        <a href="http://gsc.mit.edu/gwamit">http://gsc.mit.edu/gwamit</a>
      </td>
    </tr>
  </tbody>
  ...

```

lbls = document.getElementsByClassName("label")
> [<th class="label">Email:</th>;
<th class="label">Address:</th>;
<th class="label">Website:</th>;
...]

websiteLbl = (filter isWebsite lbls)[0]
> <th class="label">Website:</th>

websiteLbl.nextSibling
> <td class="data"> ...

Extension Name	Extension Behavior
PrintNewYorker	Appends “?printable=true” to internal links on newyorker.com
Google Reader client	Sends RSS feed links to Google Reader
Gmail checker	Rewrites “mailto:” links to open Gmail’s compose page
Bookmarking	Sends selected text to delicious.com
Dictionary lookup	Queries online dictionary with selection; displays definition in a floating <div>
Facebook data miner	Sends friends’ web-addresses to delicious.com
JavaScript toolbox	Edits selected text
Password manager	Stores and retrieves passwords on each page
Magnify under mouse	Modifies CSS on the page
Short URL expander	Sends URLs to longurlplease.com
Typography	Modifies values of <input> elements

USING REFINED TYPES

POLICIES IN FINE 

```
type elt
```

Native DOM
elements, abstract to
Fine

```
val getAttr :  
    elt  
    -> string  
    -> string
```

Defined in
F#/JavaScript with
this type

```
assume  $\forall$  (e:elt) . EltTagName e "a"  
   $\Rightarrow$  CanReadAttr e "href"
```

```
type elt
```

```
val getAttr :  
  elt  
  -> { key:string | CanReadAttr e key }  
  -> string
```

```
val getTagName :  
  elt  
  -> string
```



Precondition

```
assume  $\forall$  (e:elt) . EltTagName e "a"  
   $\Rightarrow$  CanReadAttr e "href"
```

```
type elt
```

```
val getAttr :  
  e:elt  
  -> { key:string | CanReadAttr e key }  
  -> string
```

```
val getTagName :  
  e:elt  
  -> { name:string | EltTagName e name }
```



Postcondition

```
assume  $\forall$  (e:elt) . EltTagName e "a"  
   $\Rightarrow$  CanReadAttr e "href"
```

```
type elt
```

```
val getAttr
```

1. No runtime overhead (fast)

2. No runtime security exceptions (robust)

3. Fine + Z3 check pre- and post-conditions

```
  e:elt  
  -> {
```

```
  -> str
```

```
val getTagName
```

```
  e:elt  
  -> { name:String | EltTagName e name }
```

```
// code
```

```
let getLink elt =
```

```
  if getTagName elt = "a" then
```

```
    // true  $\Rightarrow$  EltTagName elt "a"
```

```
    getAttr elt "href" // requires CanReadAttr elt "href"
```

```
  else
```

```
    "not a link"
```

Postcondition

assume \forall (e:elt) . **CanReadAttr** e "class"

assume \forall (label:elt), (labelText:elt) .
EltParent labelText label
&& EltAttr label "class" "label"
 \Rightarrow **CanReadValue** labelText

Can read

assume \forall (data:elt), (label:elt), (labelText:elt),
(website:elt), (parent:elt) .

EltParent data parent
&& EltParent label parent
&& EltParent website data
&& EltParent labelText label
&& EltAttr label "class" "label"
&& EltTextValue labelText "Website:"

\Rightarrow **CanReadAttr** website "href"

*data and
label are
siblings via
parent*

(L;P)-safety: Semantics of policies

- Execution of browser extensions interleaved with JS-code on a web page

Key feature of (L;P)-safety: Security of an extension is independent of effects of JS on the page

⇒ Security of browser does not depend on the page currently being viewed

⇒ Simply programming model: extension author does not have to consider JS interleavings in order to comply with security policy

Safety by typing

Main theorem:

- Given a Datalog policy P , a set of ground facts L , an environment Γ such that $\Gamma \models L$, a program e and a type t .
- $P; \Gamma \vdash e : t \Rightarrow e$ is $(L;P)$ -safe

Reduction relation: $P \vdash (L;e) \rightarrow (L';e')$

- Reduction steps are guarded by policy queries evaluated over a set of accumulated ground facts L
- Theorem says: **well-typed programs never raise security exceptions**

Visualizing Policies

The image shows a screenshot of a Facebook profile page for Jean Yang. The browser address bar displays the URL: <http://www.facebook.com/profile.php?id=11544&v=info&ref=ts>. The profile header includes a cover photo titled "Settling Back into Seattle" and a profile picture. The left sidebar contains sections for "Notes" (2 notes), "Gifts" (4 of 12 gifts), and "Suggest Friends for Jean". The main content area is titled "Likes and Interests" and lists several categories, each with a red box around its label: "Activities" (Running, Yoga, Acrobatics, Blogging), "Interests" (Coconut Juice, Polymorphism, The Sun), "Music" (Amanda Palmer, The Dresden Dolls, Rasputina, Tom Waits, Opera, Show tunes), and "Books" (Mrs Dalloway, A Room of One's Own, The Garden of Eden). Below this is the "Contact Information" section, which includes a "Contact Info" label and several fields with red boxes: "Email" (jean.yang.write@gmail.com, jeanyang@mit.edu), "Mobile Phone" (4123026391), and "Website" (http://people.csail.mit.edu/jeanyang, http://jkyzabc.blogspot.com, http://gsc.mit.edu/gwanit). At the bottom of the page, there is a "Share" button and a "chat" button. The footer contains the text "Facebook © 2010 · English (US)" and a navigation menu: "About · Advertising · Developers · Careers · Privacy · Terms · Help".

Experimental Summary

Name	LOC	# Assumes	Compile (s)	#Z3 q's
Verified for access control properties				
Magnifier	23			
PrintNewYorker	45			
Dictionary lookup	70			
FacePalm	142			
Bib Parser	262			
Verified for access control and data flow properties				
Password Manager	52			
Twitter Miner	36			
Bing Miner	35			
Netflix Miner	110			
Glue Miner	101			
News Personalizer	124			
Search Personalizer	382			
Partially ported Chrome extensions				
Bookmarking	(6K)	19		
Gmail Checker Plus	(7K)	43		
JavaScript Toolbox	(2K)	19		
Short URL Expander	(494)	22		
Typography	(20K)	44		
TOTAL	1,529			

- Variety of extension types
- Over 1,500 LOC total
- Many extensions ported from Chrome
 - Only the content script in Fine
 - Compiled to JavaScript using a new Fine backend
 - Much of the code remains in the extension core

Microsoft® Research

Ben Livshits

Microsoft Research

Redmond, Washington

...with help from Matt Fredrikson, Arjun Guha, Nikhil Swamy, and others

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

Verified Security for Browser

REPRIV: Re-Envisioning In-Browser Privacy

Matthew Fredrikson
University of Wisconsin

Abstract

In this paper, we present REPRIV, a system for managing and controlling the release of private information from the browser. We demonstrate how advanced user interest mining can effectively identify user interests in a real browser. We go on to discuss an extensible framework that allows third-party code to extract and disseminate more detailed information, as well as language-based tools for verifying the absence of privacy leaks in this information. To demonstrate the effectiveness of our model, we present extensions that perform personalization for Microsoft and Google. We evaluated several aspects of REPRIV: (1) We show that REPRIV's default personalization does so with no noticeable coverage overhead to the user. (2) The results it produces are of high quality, similar to those of commercial personalization services. (3) Similar results for each of our extensions, and that the overhead of REPRIV is minimal. (4) REPRIV provides individual privacy control.

1. Introduction

The research data