

Third Party Trackers on Law School Library Websites

Tom Boone

Georgetown Law Library

tom@tomboone.com

github.com/tomboone



GEORGETOWN LAW LIBRARY

[LIBRARY CATALOGS](#) • [MY ACCOUNT](#)

[CHAT WITH A LIBRARIAN](#)

[About the Library](#)

[Research](#)

[Collections](#)

[Student Resources](#)

[Faculty Resources](#)

[Alumni & Visitor Resources](#)

[Today's Hours](#)

Williams 7am - 2am
Circulation 8am - 10:45pm
Reference 9am - 8pm

[View full calendar](#)

[Catalog](#)

[OneSearch](#)

[Databases](#)

[Website](#)

Search GULLiver for books, journals, media, & more by



What am I searching

[ADVANCED SEARCH](#)

[Request a research consultation](#)

[Book a group study room](#)

[Library FAQs](#)

DUE PROCESS BLOG: LATEST ENTRIES

[Meet your Reference Librarians](#) | January 17, 2018

[Library Training for Faculty Research Assistants](#) | January 11, 2018

[Extended Library Hours and Access Restrictions During Exams](#) | December 6, 2017

[Visit Due Process blog](#)

Research Resources

Georgetown Law Library provides online research guides, treatise finder lists, and on-demand research tutorials, all created specifically for Georgetown Law students by our librarians.

[Explore Research Resources](#)

[Databases](#)

[Research Guides](#)

[Tutorials](#)

[Treatise Finders](#)

[CURRENT STUDENTS](#)

[ALUMNI](#)

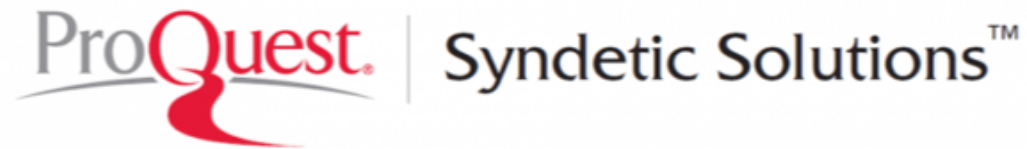
[FACULTY](#)

[STAFF](#)

[SEARCH STAFF & FACULTY DIRECTORY](#)



Google Analytics Solutions



Online Tracking: A 1-million-site Measurement and Analysis

Steven Englehardt
Princeton University
ste@cs.princeton.edu

Arvind Narayanan
Princeton University
arvindn@cs.princeton.edu

This is an extended version of our paper that appeared at ACM CCS 2016.

ABSTRACT

We present the largest and most detailed measurement of online tracking conducted to date, based on a crawl of the top 1 million websites. We make 15 types of measurements on each site, including stateful (cookie-based) and stateless (fingerprinting-based) tracking, the effect of browser privacy tools, and the exchange of tracking data between different sites (“cookie syncing”). Our findings include multiple sophisticated fingerprinting techniques never before measured in the wild.

This measurement is made possible by our open-source web privacy measurement tool, OpenWPM¹, which uses an automated version of a full-fledged consumer browser. It supports parallelism for speed and scale, automatic recovery

to resort to a stripped-down browser [31] (a limitation we explore in detail in Section 3.3). (2) We provide comprehensive instrumentation by expanding on the rich browser extension instrumentation of FourthParty [33], without requiring the researcher to write their own automation code. (3) We reduce duplication of work by providing a modular architecture to enable code re-use between studies.

Solving these problems is hard because the web is not designed for automation or instrumentation. Selenium², the main tool for automated browsing through a full-fledged browser, is intended for developers to test their *own* websites. As a result it performs poorly on websites not controlled by the user and breaks frequently if used for large-scale measurements. Browsers themselves tend to suffer memory leaks over long sessions. In addition, instrument



This repository

Search

Pull requests

Issues

Marketplace

Explore



citp / OpenWPM

Watch

60

Star

632

Fork

110

<> Code

Issues 44

Pull requests 2

Projects 0

Wiki

Insights

A web privacy measurement framework <https://webtap.princeton.edu/>

774 commits

3 branches

14 releases

16 contributors

GPL-3.0

Branch: master

New pull request

Create new file

Upload files

Find file

Clone or download

randomwalker typo

Latest commit 63610ca 11 hours ago

| | | |
|------------------|---|--------------|
| automation | Change sqlite timeout to 60 seconds to avoid database locking problems. | 3 months ago |
| test | Adding documentation and tests for content saving | 4 months ago |
| .gitignore | Merge branch 'master' of github.com:citp/OpenWPM | 2 years ago |
| .travis.yml | Bugfix: Adding pillow->PIL map to | 4 months ago |
| CHANGELOG | Merge remote-tracking branch 'origin' into ff52 | 4 months ago |
| LICENSE | Store POST request data as JSON in the database. | a year ago |
| README.md | typo | 11 hours ago |
| VERSION | Version bump to checkpoint final release to support FF45. | 4 months ago |
| __init__.py | Create data_directory if it doesn't exist + drop an init file in main | 3 years ago |
| demo.py | PEP8 Fixes | 4 months ago |
| install-dev.sh | Use apt-get -y so there is no prompt for the user anymore. | 6 months ago |
| install.sh | Bump Firefox to 52.4.1esr | 4 months ago |
| requirements.txt | Restrict Selenium to 3.6.0 or below | 3 months ago |

| | Site | Daily Time on Site ? | Daily Pageviews per Visitor ? | % of Traffic From Search ? | Total Sites Linking In ? |
|---|--|----------------------|-------------------------------|----------------------------|--------------------------|
| 1 | Google.com Enables users to search the world's information, including webpages, images, and videos. Offers... More | 7:23 | 8.26 | 4.70% | 4,006,550 |
| 2 | Youtube.com User-submitted videos with rating, comments, and contests. | 8:17 | | | 3,045,350 |
| 3 | Facebook.com | 10:45 | | | 7,985,144 |
| 4 | Baidu.com The leading Chinese language search engine, provides "simple and reliable" search exp... More | 7:18 | | | 197,243 |
| 5 | Wikipedia.org A free encyclopedia built collaboratively using wiki software. (Creative Commons Attribution-Sh... More | 4:17 | | | 2,171,871 |
| 6 | Reddit.com User-generated news links. Votes promote stories to the front page. | 15:49 | | | 537,191 |
| 7 | Yahoo.com A major internet portal and service provider offering search results, customizable content, cha... More | 3:59 | | | 765,148 |
| 8 | Google.co.in Indian version of this popular search engine. Search the whole web or only webpages from India.... More | 6:52 | | | 33,588 |
| 9 | Qq.com China's largest and most used Internet service portal owned by Tencent, Inc founded in Nov... More | 4:37 | 3.81 | 4.40% | 363,772 |



A1 fx id

| | A | B | C | D | E | F | G | H | I | J |
|----|---------------------|---|--|---------------------------------------|---|---|---|---|---|---|
| 1 | id | school | library | aba_name | website | | | | | |
| 2 | akron | University of Akron School of Law | University of Akron School of Law Library | University of Akron | http://uakron.edu/law/about-us/law-library.dot | | | | | |
| 3 | alabama | University of Alabama School of Law | Bounds Law Library | University of Alabama | http://library.law.ua.edu | | | | | |
| 4 | albany | Albany Law School, Union University | Schaffer Law Library | Albany Law School | http://www.albanylaw.edu/academic-life/schaffer-law-library | | | | | |
| 5 | american | Washington College of Law, American University | Pence Law Library | American University | https://library.wcl.american.edu | | | | | |
| 6 | appalachian | Appalachian School of Law | Appalachian School of Law Library | Appalachian School of Law | http://www.asl.edu/library/ | | | | | |
| 7 | arizona | James E. Rogers College of Law, University of Arizona | Daniel F. Cracchiolo Law Library | University of Arizona | http://lawlibrary.arizona.edu | | | | | |
| 8 | arizonastate | Sandra Day O'Connor College of Law, Arizona State University | Ross-Blakley Law Library | Arizona State University | https://web.law.asu.edu/library/RossBlakleyLawLibrary.aspx | | | | | |
| 9 | arizonasummit | Arizona Summit Law School | Arizona Summit Law School Library | Arizona Summit Law School | http://www.azsummitlaw.edu/student-resources-law-library.html | | | | | |
| 10 | arkansas | University of Arkansas School of Law | Young Law Library | University of Arkansas - Fayetteville | http://law.uark.edu/library/ | | | | | |
| 11 | arkansas-littlerock | William H. Bowen School of Law, University of Arkansas at Little Rock | UALR Bowen School of Law/Pulaski County Law Library | University of Arkansas - Little Rock | http://ualr.edu/lawlibrary/ | | | | | |
| 12 | avemaria | Ave Maria School of Law | Ave Maria School of Law Library | Ave Maria School of Law | http://www.avemarialaw.edu/library/ | | | | | |
| 13 | baltimore | University of Baltimore School of Law | University of Baltimore School of Law Library | University of Baltimore | http://law.ubalt.edu/library/ | | | | | |
| 14 | barry | Dwayne O. Andreas School of Law, Barry University | Euliano Law Library | Barry University | https://www.barry.edu/law/future-students/library/ | | | | | |
| 15 | baylor | Baylor Law School, Baylor University | Sheridan and John Eddie Williams Legal Research Center | Baylor University | http://www.baylor.edu/law/library/ | | | | | |
| 16 | belmont | Belmont University College of Law | Belmont University College of Law Library | Belmont University | http://www.belmont.edu/law/library/index.html | | | | | |
| 17 | bostoncollege | Boston College Law School | Boston College Law Library | Boston College | https://www.bc.edu/bc-web/schools/law/sites/current-students/library.html | | | | | |
| 18 | bostonuniversity | Boston University School of Law | Samuel M. Fineman Law Library & Pappas Law Library | Boston University | http://www.bu.edu/lawlibrary/ | | | | | |
| 19 | brighamyoung | J. Reuben Clark Law School, Brigham Young University | Howard W. Hunter Law Library | Brigham Young University | http://lawlib.byu.edu | | | | | |
| 20 | brooklyn | Brooklyn Law School | Brooklyn Law School Library | Brooklyn Law School | https://www.brooklaw.edu/intellectuallife/library.aspx | | | | | |
| 21 | buffalo | University at Buffalo Law School, SUNY | Charles B. Sears Law Library | University at Buffalo - SUNY | http://law.lib.buffalo.edu | | | | | |
| 22 | cal-berkeley | University of California, Berkeley School of Law (Boalt Hall) | University of California, Berkeley Law Library | University of California, Berkeley | https://www.law.berkeley.edu/law-library/ | | | | | |
| 23 | cal-davis | University of California, Davis School of Law (King Hall) | UC Davis Mabie Law Library | University of California, Davis | https://law.ucdavis.edu/library/ | | | | | |
| 24 | cal-hastings | University of California, Hastings College of the Law | UC Hastings Law Library | University of California, Hastings | http://library.uchastings.edu/index.php | | | | | |
| 25 | cal-irving | University of California, Irvine School of Law | University of California, Irvine School of Law | University of California, Irvine | http://www.law.uci.edu/library/ | | | | | |
| 26 | cal-losangeles | University of California, Los Angeles School of Law | Hugh & Hazel Darling Law Library | University of California, Los Angeles | http://law.ucla.edu/library/ | | | | | |
| 27 | calwestern | California Western School of Law | California Western School of Law Library | California Western School of Law | https://www.cwsl.edu/student-life/library | | | | | |
| 28 | campbell | Norman Adrian Wiggins School of Law, Campbell University | Campbell Law Library | Campbell University | https://law.campbell.edu/learn/campbell-law-library/ | | | | | |
| 29 | capital | Capital University Law School | Capital University Law School Library | Capital University | http://law.capital.edu/Law_Library/ | | | | | |
| 30 | cardozo | Benjamin N. Cardozo School of Law, Yeshiva University | Dr. Lillian and Dr. Rebecca Chutick Law Library | Yeshiva University (Cardozo) | https://cardozo.yu.edu/library | | | | | |
| 31 | casewestern | Case Western Reserve University School of Law | Judge Ben C. Green Law Library | Case Western Reserve University | http://law.case.edu/Our-School/Library | | | | | |
| 32 | catholic | Columbus School of Law, The Catholic University of America | Judge Kathryn J. DuFour Law Library | Catholic University of America | http://www.law.edu/library/index.cfm | | | | | |
| 33 | chapman | Dale E. Fowler School of Law, Chapman University | Hugh & Hazel Darling Law Library | Chapman University | https://www.chapman.edu/law/student-resources/library/index.aspx | | | | | |
| 34 | charleston | Charleston School of Law | Sol Blatt Jr. Law Library | Charleston School of Law | http://charlestonlaw.edu/sol-blatt-jr-law-library/ | | | | | |
| 35 | chicago | University of Chicago Law School | D'Angelo Law Library | University of Chicago | https://www.lib.uchicago.edu/law/ | | | | | |
| 36 | chicagokent | Chicago-Kent College of Law, Illinois Institute of Technology | IIT Chicago-Kent Law Library | IIT Chicago-Kent College of Law | https://www.kentlaw.iit.edu/library | | | | | |
| 37 | cincinnati | University of Cincinnati College of Law | Robert S. Marx Law Library | University of Cincinnati | http://www.law.uc.edu/library | | | | | |
| 38 | clevelandmarshall | Cleveland-Marshall College of Law, Cleveland State University | Cleveland-Marshall Law Library | Cleveland State University | https://www.law.csuohio.edu/lawlibrary | | | | | |
| 39 | colorado | University of Colorado School of Law | William A. Wise Law Library | University of Colorado | http://lawlibrary.colorado.edu | | | | | |
| 40 | columbia | Columbia University Law School | Arthur W. Diamond Law Library | Columbia University | http://www.law.columbia.edu/library | | | | | |

2017-08-20-lawliburls



Most 3rd Party HTTP Requests

(Avg. Page Load – All Contexts)

| | | |
|----|--------------------------------------|--------------|
| 1 | Boston College | 195.8 |
| 2 | Brigham Young University | 128.1 |
| 3 | University of Wyoming | 113.5 |
| 4 | Saint Louis University | 105.5 |
| 5 | University of the Pacific (McGeorge) | 101.0 |
| 6 | Ave Maria School of Law | 96.6 |
| 7 | University of Richmond | 96.4 |
| 8 | Case Western Reserve University | 90.3 |
| 9 | Florida Coastal School of Law | 89.1 |
| 10 | University of Pittsburgh | 79.3 |

Overview

The EasyList filter lists are sets of rules originally designed for [Adblock](#) that automatically remove unwanted content from the internet, including annoying adverts, bothersome banners and troublesome tracking. The filter lists are currently maintained by four authors, Fanboy, MonztA, Famlam and Khrrin, who are ably assisted by an ample forum community.

The links listed below allow you to select filter lists for use in your browser provided that you are using a compatible ad blocker (tested with [Adblock Plus](#), [AdBlock](#) and [uBlock Origin](#)). Furthermore, EasyPrivacy Tracking Protection List is available for [Internet Explorer 9](#) and higher.

EasyList

EasyList is the primary filter list that removes most adverts from international webpages, including unwanted frames, images and objects. It is the most popular list used by many ad blockers and forms the basis of over a dozen combination and supplementary filter lists.

[View EasyList](#) or [add it to your ad blocker](#)

EasyPrivacy

EasyPrivacy is an optional supplementary filter list that completely removes all forms of tracking from the internet, including web bugs, tracking scripts and information collectors, thereby protecting your personal data.

[View EasyPrivacy](#) or [add it to your ad blocker](#)

Fanboy's Annoyance List

Fanboy's Annoyance List blocks Social Media content, in-page pop-ups and other annoyances; thereby substantially decreasing web page loading times and uncluttering them. Fanboy's Social Blocking List is already included, there is no need to subscribe to it if you already have Fanboy's Annoyance List.


[View Fanboy's Annoyance List](#) or [add it to your ad blocker](#)

```
[Adblock Plus 2.0]
! Version: 201802161255
! Title: EasyList
! Last modified: 16 Feb 2018 12:55 UTC
! Expires: 4 days (update frequency)
! Homepage: https://easylist.to/
! Licence: https://easylist.to/pages/licence.html
!
! Please report any unblocked adverts or problems
! in the forums (https://forums.lanik.us/)
! or via e-mail (easylist.subscription@gmail.com).
! GitHub issues: https://github.com/easylist/easylist/issues
! GitHub pull requests: https://github.com/easylist/easylist/pulls
!
! -----General advert blocking filters-----!
! *** easylist:easylist/easylist_general_block.txt ***
&act=ads_
&ad.vid=${xmlhttprequest}
&ad_box_
&ad_channel=
&ad_classid=
&ad_height=
&ad_ids=
&ad_keyword=
&ad_network_
&ad_number=
&ad_revenue=
&ad_slot=
&ad_time=
&ad_type=
&ad_type_
&ad_url=
&ad_zones=
&adbannerid=
&adclient=
&adcount=
&adflag=
&adgroupid=
&adlist=
&admeld_
&admid=
&adname=
&adnet=
&adnum=
&adpageurl=
&Ads_DFP=
&adsafe=
&adserv=
&adserver=
&adsize=
&adslot=
&adslots=
&adsourceid=
```

Most 3rd Party HTTP Requests

(Avg. Page Load – Tracking Context)

| | | |
|----|--------------------------------------|-------------|
| 1 | Suffolk University | 29.9 |
| 2 | University of the Pacific (McGeorge) | 26.0 |
| 3 | University of Pittsburgh | 25.7 |
| 4 | Arizona State University | 23.7 |
| 5 | Concordia University | 23.1 |
| 6 | University of Wyoming | 20.5 |
| 7 | Pepperdine University | 20.2 |
| 8 | Drexel University | 20.1 |
| 9 | Elon University | 19.2 |
| 10 | Case Western Reserve University | 19.1 |

| | all | cookie | css | image | media | script | XHR | frame | other |
|--|-----|--------|-----|-------|-------|--------|-----|-------|-------|
| 1st-party | | | | | | | | | |
|  google.com | 2 | | | | | | | | |
| adservice.google.com | | | | 1 | | | | | |
| apis.google.com | | | | | | 1 | | | |
| www.google.com | | | 1 | 9 | | 3 | | | |

Most 3rd Party Domains Requested

(Avg. Page Load – Tracking Context)

| | | |
|----|--------------------------------------|----|
| 1 | University of Pittsburgh | 13 |
| 2 | Suffolk University | 12 |
| 3 | University of Wyoming | 11 |
| 3 | University of the Pacific (McGeorge) | 11 |
| 5 | Concordia University | 10 |
| 5 | Elon University | 10 |
| 7 | Saint Louis University | 9 |
| 7 | University of Baltimore | 9 |
| 9 | Creighton University | 8 |
| 10 | Barry University | 7 |

| | all | cookie | css | image | media | script | XHR | frame | other |
|----------------------|-----|--------|-----|-------|-------|--------|-----|-------|-------|
| 1st-party | | | | | | | | | |
| google.com | 2 | | | | | | | | |
| adservice.google.com | | | | 1 | | | | | |
| apis.google.com | | | | | | 1 | | | |
| www.google.com | | | 1 | 9 | | 3 | | | |
| gstatic.com | | | | | | | | | |
| ssl.gstatic.com | | | | 1 | | | | | |
| www.gstatic.com | | | 1 | | | 2 | | | |

Most 3rd Party Orgs Requested

(Avg. Page Load - Tracking Context)

| | | |
|----|--------------------------------------|----|
| 1 | University of Pittsburgh | 10 |
| 1 | University of Wyoming | 10 |
| 3 | Elon University | 9 |
| 3 | Suffolk University | 9 |
| 3 | University of the Pacific (McGeorge) | 9 |
| 6 | Concordia University | 8 |
| 6 | Saint Louis University | 8 |
| 8 | Creighton University | 6 |
| 8 | University of Baltimore | 6 |
| 10 | Barry University | 5 |

| | Mean | Median |
|---------------------------------------|-------------|---------------|
| Third Party HTTP Request | 29.5 | 22 |
| Third Party Tracker Requests | 7.32 | 6 |
| Third Party Tracker Domains Requested | 2.99 | 2 |
| Third Party Tracker Orgs Requested | 2.08 | 2 |

Most Common Domains

(Avg. Page Load - All Contexts)

| | | |
|----|----------------------|-------|
| 1 | google-analytics.com | 85.9% |
| 2 | googleapis.com | 63.4% |
| 3 | doubleclick.net | 52.2% |
| 4 | google.com | 49.8% |
| 5 | gstatic.com | 42.0% |
| 6 | googletagmanager.com | 35.6% |
| 7 | facebook.com | 31.7% |
| 8 | facebook.net | 30.2% |
| 9 | bootstrapcdn.com | 24.4% |
| 10 | googleadservices.com | 22.9% |

Most Common Domains

(Avg. Page Load – Tracking Context)

| | | |
|----|----------------------|--------------|
| 1 | google-analytics.com | 85.9% |
| 2 | doubleclick.net | 51.7% |
| 3 | facebook.com | 30.7% |
| 4 | typekit.net | 15.1% |
| 5 | googleadservices.com | 12.7% |
| 6 | google.com | 11.2% |
| 7 | addthis.com | 8.8% |
| 8 | twitter.com | 8.3% |
| 9 | newrelic.com | 7.8% |
| 10 | addthisedge.com | 6.8% |

Most Common Orgs

(Avg. Page Load – Tracking Context)

| | | |
|----|-----------|--------------|
| 1 | Alphabet | 86.8% |
| 2 | Facebook | 30.7% |
| 3 | Adobe | 16.1% |
| 4 | Oracle | 8.8% |
| 5 | Twitter | 8.3% |
| 6 | New Relic | 7.8% |
| 7 | LinkedIn | 6.3% |
| 8 | Amazon | 5.4% |
| 9 | AppNexus | 4.4% |
| 10 | Verizon | 3.4% |

Most 3rd-party cookies set

(Avg. Page Load)

| | | |
|----|--------------------------------------|------|
| 1 | University of Wyoming | 55 |
| 2 | Suffolk University | 53 |
| 3 | Vermont Law School | 46 |
| 4 | University of the Pacific (McGeorge) | 43.5 |
| 5 | Saint Louis University | 37 |
| 6 | IIT Chicago-Kent College of Law | 35 |
| 7 | Loyola University Chicago | 34 |
| 7 | University of Pittsburgh | 34 |
| 9 | Capital University | 33 |
| 10 | Albany Law School | 30.5 |

Most 3rd-party cookie domains

(Avg. Page Load)

| | | |
|----|--------------------------------------|----|
| 1 | University of Wyoming | 20 |
| 2 | Suffolk University | 15 |
| 2 | University of the Pacific (McGeorge) | 15 |
| 4 | IIT Chicago-Kent College of Law | 12 |
| 5 | Saint Louis University | 11 |
| 6 | University of Pittsburgh | 10 |
| 7 | Elon University | 9 |
| 8 | Concordia University | 8 |
| 8 | Loyola University Chicago | 8 |
| 10 | Barry University | 7 |

Most 3rd-party cookie orgs

(Avg. Page Load)

| | | |
|---|--------------------------------------|-----------|
| 1 | University of Wyoming | 17 |
| 2 | Suffolk University | 14 |
| 3 | University of the Pacific (McGeorge) | 12 |
| 4 | IIT Chicago-Kent College of Law | 11 |
| 4 | Saint Louis University | 11 |
| 6 | Concordia University | 8 |
| 6 | Elon University | 8 |
| 6 | University of Pittsburgh | 8 |
| 9 | Barry University | 6 |
| 9 | Loyola University Chicago | 6 |
| 9 | Southwestern University | 6 |
| 9 | Stetson University | 6 |
| 9 | University of Akron | 6 |
| 9 | Vermont Law School | 6 |

Most common 3rd-party cookie domains

(Avg. Page Load)

| | | |
|---|-----------------|-------|
| 1 | doubleclick.net | 34.6% |
| 2 | google.com | 32.7% |
| 3 | facebook.com | 24.4% |
| 4 | adnxs.com | 9.8% |
| 5 | addthis.com | 9.3% |
| 5 | cloudflare.com | 9.3% |
| 5 | twitter.com | 9.3% |
| 8 | newrelic.com | 7.3% |
| 9 | adsrvr.org | 6.8% |
| 9 | linkedin.com | 6.8% |

Most common 3rd-party cookie orgs

(Avg. Page Load)

| | | |
|---|----------------|--------------|
| 1 | Alphabet | 51.2% |
| 2 | Facebook | 24.4% |
| 3 | AppNexus | 9.8% |
| 4 | Cloudflare | 9.3% |
| 4 | Oracle | 9.3% |
| 4 | Twitter | 9.3% |
| 7 | Verizon! | 8.4% |
| 8 | New Relic | 7.3% |
| 9 | LinkedIn | 6.8% |
| 9 | The Trade Desk | 6.8% |