



Consumer privacy I

Privacy Technology in Context

David Sidi (dsidi@email.arizona.edu)



Nest
Google Wallet Gmail
Slides
Google Analytics Books
Chromecast
Google+ Maps AdSense
Drive
Sheets Android
Google News Pixel
ChromeOS
DNS Calendar Chrome Allo
PRISM Home Duo
Google Wifi
Docs DeepMind
YouTube Search Photos
Translate
Google Fiber Keep

it's fiiiiiiiiine



Google has “vertical” data on customers

- Chromebook hardware / Nexus Phone
- Chrome OS / Android OS
- Google App store
- Gmail/Inbox, YouTube, Docs, Hangouts, Google Pay,



Google has “horizontal” data too

- Visiting non-Google sites:
 - DNS resolution via 8.8.8.8
 - Google Analytics trackers, Google Fonts, Google Ajax site search
- Not at a computer: maps lookups, google home searches, chromecast viewing

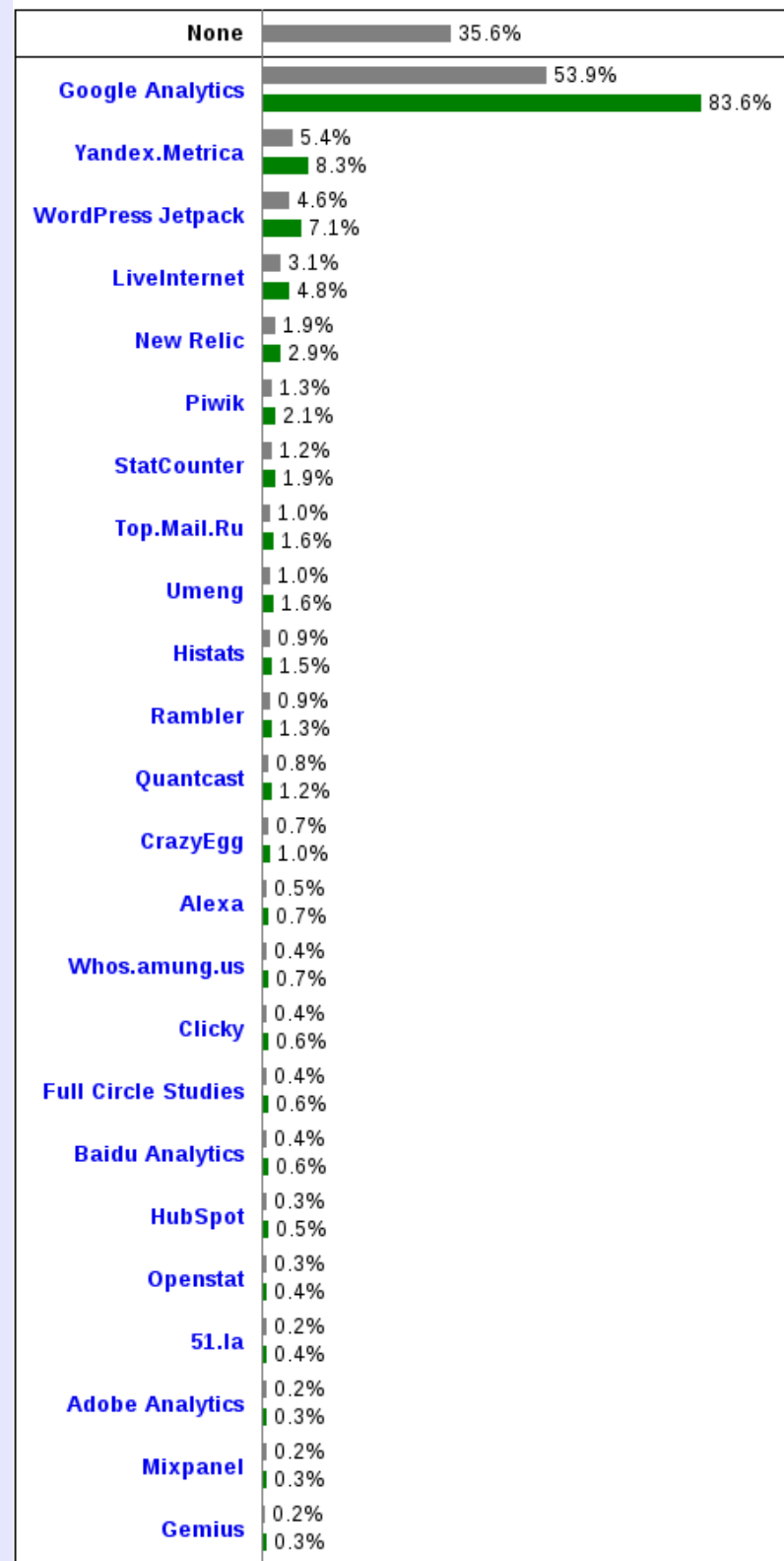


Third-party analytics

- 53.9% of all sites use Google Analytics

https://w3techs.com/technologies/overview/traffic_analysis/all

CC-SA License by David Sidi





Third-party analytics

- 53.9% of all sites use Google Analytics

Creating a complete picture

Begin by centralizing your data. Analytics 360 pulls in data across:



Online Sources



Offline Sources



Google Sources

Data from your site, app, internet-connected devices, and even offline sources will be connected in one place. If you're using Google and DoubleClick advertising products, seamless, out-of-the-box integrations mean you can pull in that information to create a single, complete data source across all customer touchpoints.

<https://www.google.com/analytics/analytics/features/>



Third-party analytics

- 53.9% of all sites use Google Analytics

If you want to learn more about how your customers behave away from your site, you can share your Analytics 360 customer segments with Audience Center 360 to get additional insights like demographics, interests, and in-market information.

<https://www.google.com/analytics/analytics/features/>



- Remember the course introductions, and their relation to passwords?
- That kind of thing, sharing in one understood context for your own purposes and having what was shared used in some other context that you find troubling, underlies a lot of consumer privacy issues



Everybody lies

- Sensitive survey bias
- National Household Survey on Drug Abuse (NHSDA)
 - a. During the past 30 days, on how many days did you use cocaine?
 - b. When was the most recent time you smoked a cigarette?
 - c. How many times in the past 12 months have you gotten very high or drunk on alcohol?



Sensitivity depends on context

- Sensitive: relative to anyone, to the interviewer, to third parties?
- Who are you talking to, when you email your friends on Gmail, or search on Google's search?



Google as digital truth serum

- Statisticians attempts to solve the problem of bias are varied and interesting (see: Randomized Response Techniques)
- They have one big flaw: they aren't talking to economists



Incentives matter

- “People have no incentive to tell surveys the truth.... [A] huge advantage that Google searches have in getting people to tell the truth [is] incentives. If you enjoy racist jokes, you have zero incentive to share that un-PC fact with a survey. You do, however, have an incentive to search for the best new racist jokes online.”



Perceived Costs matter too

- “Think of Google searches. Remember the conditions that make people more honest. Online? Check. Alone? Check. No person administering a survey? Check.”
- A feeling of safety, seclusion, or intimacy makes a difference



How is all that sensitive data used?

- “Do corporations have the right to judge our fitness for their services based on abstract but statistically predictive criteria not directly related to those services?”
- Price discrimination



An opposed cost to collecting sensitive data?

- Maciej's point: like nuclear energy, large quantities of data collected in the way Google or Facebook collects it can be a liability to the company collecting it
- Disclosures can lead to fines, public loss of trust



- There is a server challenge assignment due Tuesday; we'll discuss it at the end of class
- Keep your ears pricked for clues!



Let's get down to it

- recording what you're doing: script
- getting help: man
- getting out of trouble: ^C, ^D, ^|, kill
 - kill: processes, how to find the PID. Related: pkill, top
- getting around: ls, cd, pwd, find
 - ls: hidden files (including the special ones in every directory), globs
- Open a terminal. Which directory are you in?
- Change to /var/log, and list all files starting with 's'
- List exactly the names of the hidden directories in the current directory, *using only ls*
- List all the files in any subdirectory below the current one
- Start vim, then kill it from a different window



Linux Shell Basics

- sending output to stdout, elsewhere
 - echo, cat
- view a text file: less, head, tail
 - nerd joke: “less is more” (and then some)
 - try pressing ‘v’ in less: what happens?
- finding text in files: grep
- counting words, lines: wc
- Find all lines in search_me that includes the word ‘Othello’
- Echo “hi!” to a file called ‘conversation’, then use tail to watch it. Open a new terminal (keeping the original visible) and echo “hi there” in to conversation.
- Echo “Bye!” into the same file (conversation), and show all lines but the first



Linux shell basics

- chmod and permissions
- Do
 - \$ ls -l
 - what does drwxrwxrwx mean?
- Wrinkle: setuid bit
 - \$ chmod u+s ./foo



Linux Shell Basics

- finding files: `find`, `whereis`, `whence`
- composing commands: `|`, `$()`, `xargs`
 - FIFOs: named, unnamed, and anonymous
- sequences of commands: `&&`, `;`
- storing command outputs: `>` and `>>`
- “anonymous documents” with heredocs, herestrings: `<<` and `<<<`
- find all `.conf` files in `/etc`, pipe to `wc` to get the number of matches
- pipe “hi” from `figlet` to `cowsay`
- Challenge: Can you find all files named `syslog` (not `syslog.blah`), using `find`? You should find exactly three results, and no error messages



Vim

- text editor
- two modes: edit and command.
 - typing 'i' puts you in insert mode, <esc> puts you in command mode
- In command mode,
 - 'h' moves left, 'l' moves right
 - 'j' moves up, and 'k' moves down
 - 'x' is like delete, and 'X' is like backspace
 - 'u' is undo, '^r' is redo
- In insert mode,
 - arrows move you around
 - backspace removes text
- colon (:) lets you put in commands
 - :e <filename> opens an arbitrary filename
 - :w saves
 - :q quits without saving (you will get a prompt) Note: you can use :q! to force a quit without saving.



Assignment 1: Permissions trouble

- Your task is to
 - find and read a file called SECRET that is somewhere on our server
 - create a file called NOT_SECRET in your home directory containing the following, each on a separate line
 - your netid
 - the contents of SECRET
 - A brief explanation of how you did it
- You can do this with only the tools we now know