No boundaries: Exfiltration of personal data by session-replay scripts

Güneş Acar Steven Englehardt Arvind Narayanan

This research is funded by NSF grant CNS 1526353 and by an Amazon AWS Cloud Credits for Research grant.

PRIVACYCON

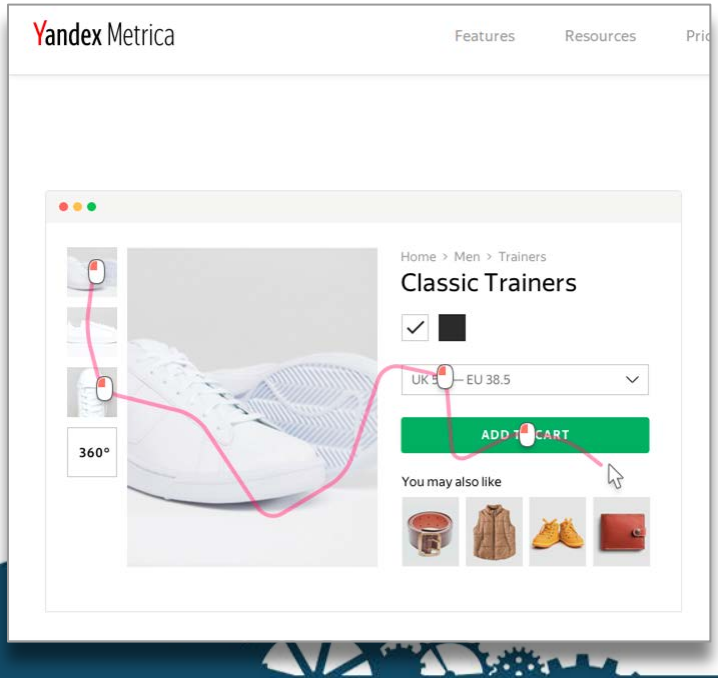
PRIVACYCON

Photo: Dan Saelinger (Prop Stylist: Dominique Baynes)
https://spectrum.ieee.org/computing/software/browser-fingerprinting-and-the-onlinetracking-arms-race

What are session replay scripts?

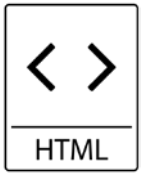
“Watch recordings of your visitors’ sessions. Discover how they browse as if you're looking over their shoulder!” (clicktale.com)

Why use session replay scripts?



- Who are my most valuable customers?
- Who added items to the cart but didn’t convert?
- Where are users frustrated?

The problem: recordings require a ton of data



Full page source and text

Mouse movements & clicks

Keypresses

Automated redaction...

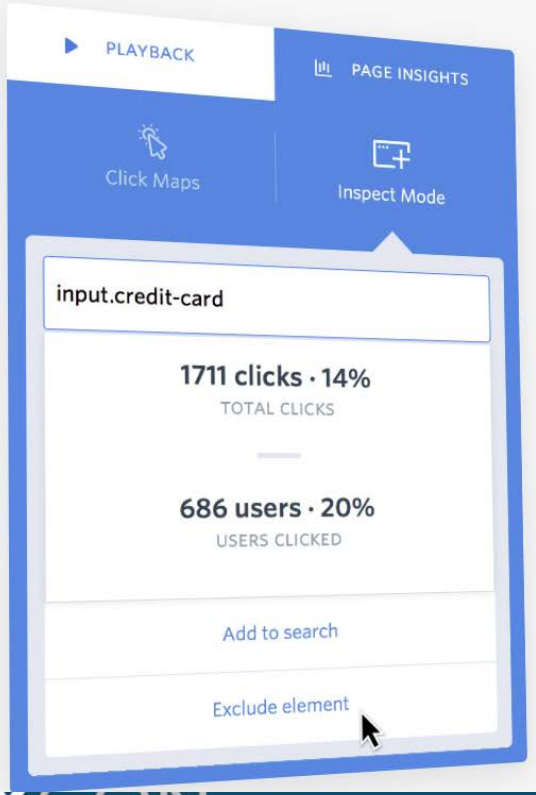
Excluded Elements

Enter CSS selectors on separate lines, describing elements to be excluded from capture. Learn more

```
input[type="password"]
[autocomplete="cc-name"]
[autocomplete="cc-given-name"]
[autocomplete="cc-additional-name"]
[autocomplete="cc-family-name"]
[autocomplete="cc-number"]
[autocomplete="cc-exp"]
[autocomplete="cc-exp-month"]
[autocomplete="cc-exp-year"]
[autocomplete="cc-csc"]
[autocomplete="cc-type"]
```

Redacted Field	FullStory	UserReplay	SessionCam	Hotjar	Yandex	Smartlook
Name	○	◐	◐	○	○	○
Email	○	◐	◐	○	○	○
Phone	○	◐	◐	○	○	○
Address	○	◐	◐	○†	○	○
SSN	○	◐	◐	○	○	○
DOB	○	◐	◐	○	○	○
Password	●	◐	●	●	●	◐
CC Number	●	◐*	◐	◐	○	●
CC CVC	●	◐	◐	○	○	●
CC Expiry	●	◐	◐	○	○	●

Manual redaction



Easily protect your user's privacy.

Exclude sensitive customer data from ever leaving your customer's browser by using our in-app point and click system.

How can things go wrong?

We found session recordings containing:

Passwords
(*Propeller Ads*, …)

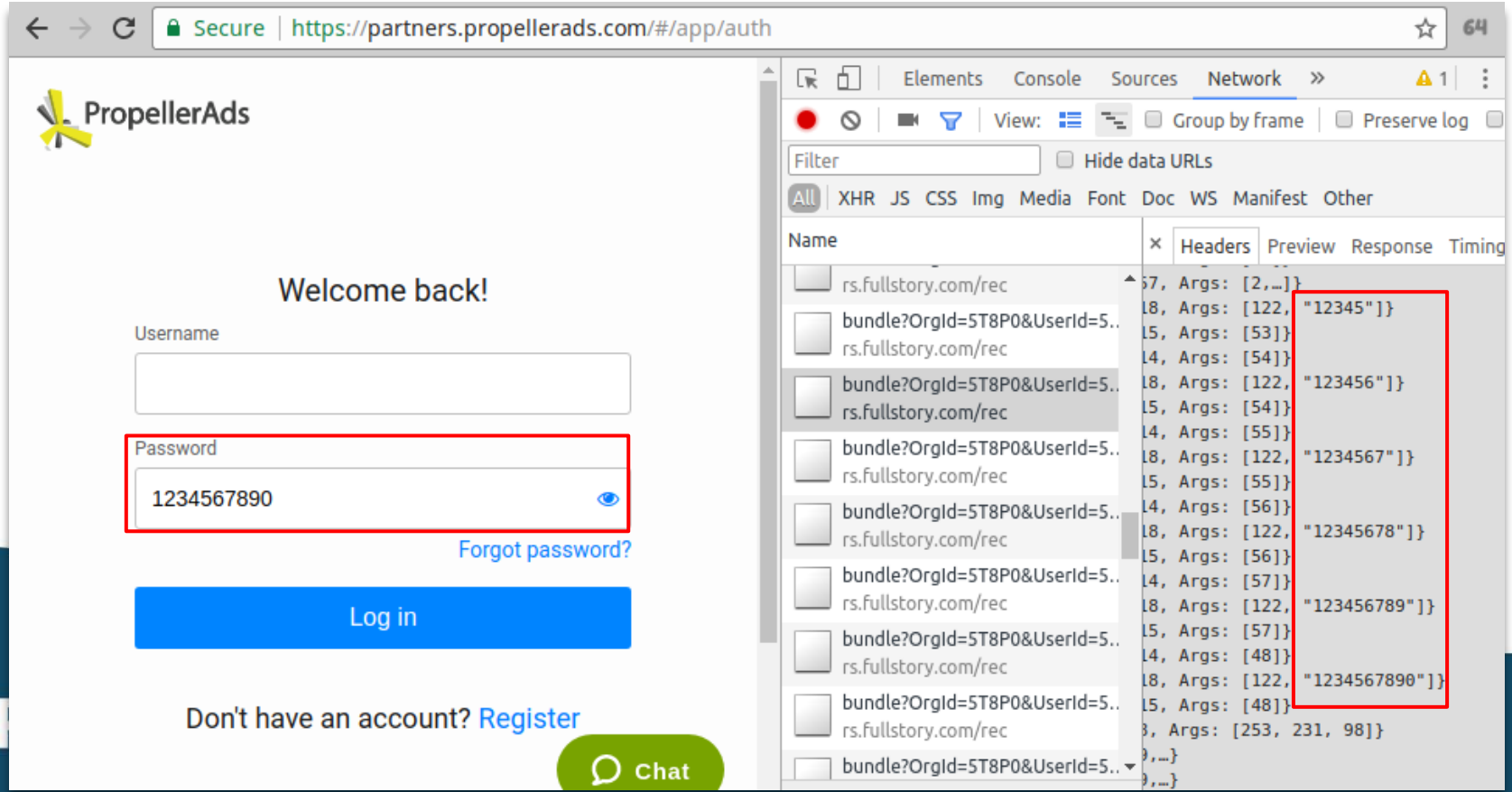
Credit Card details
(Bonobos)

Student data
(Gradescope)

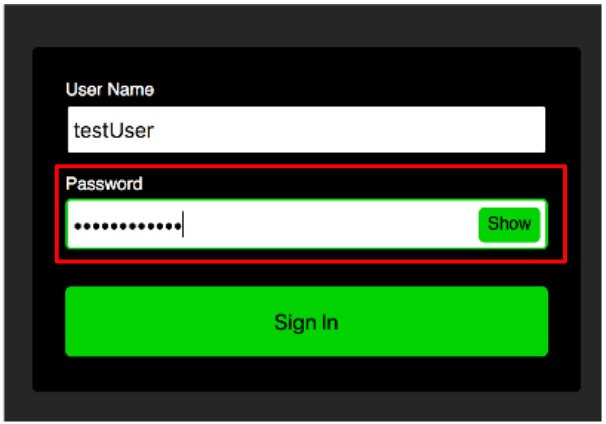
Health data
(Walgreens)

Purchase details
(Lenovo)

Passwords leak to Fullstory on PropellerAds



Filter bypass due to unexpected input types



`<input type="password">`

(automatically redacted with [type="password"] rule)

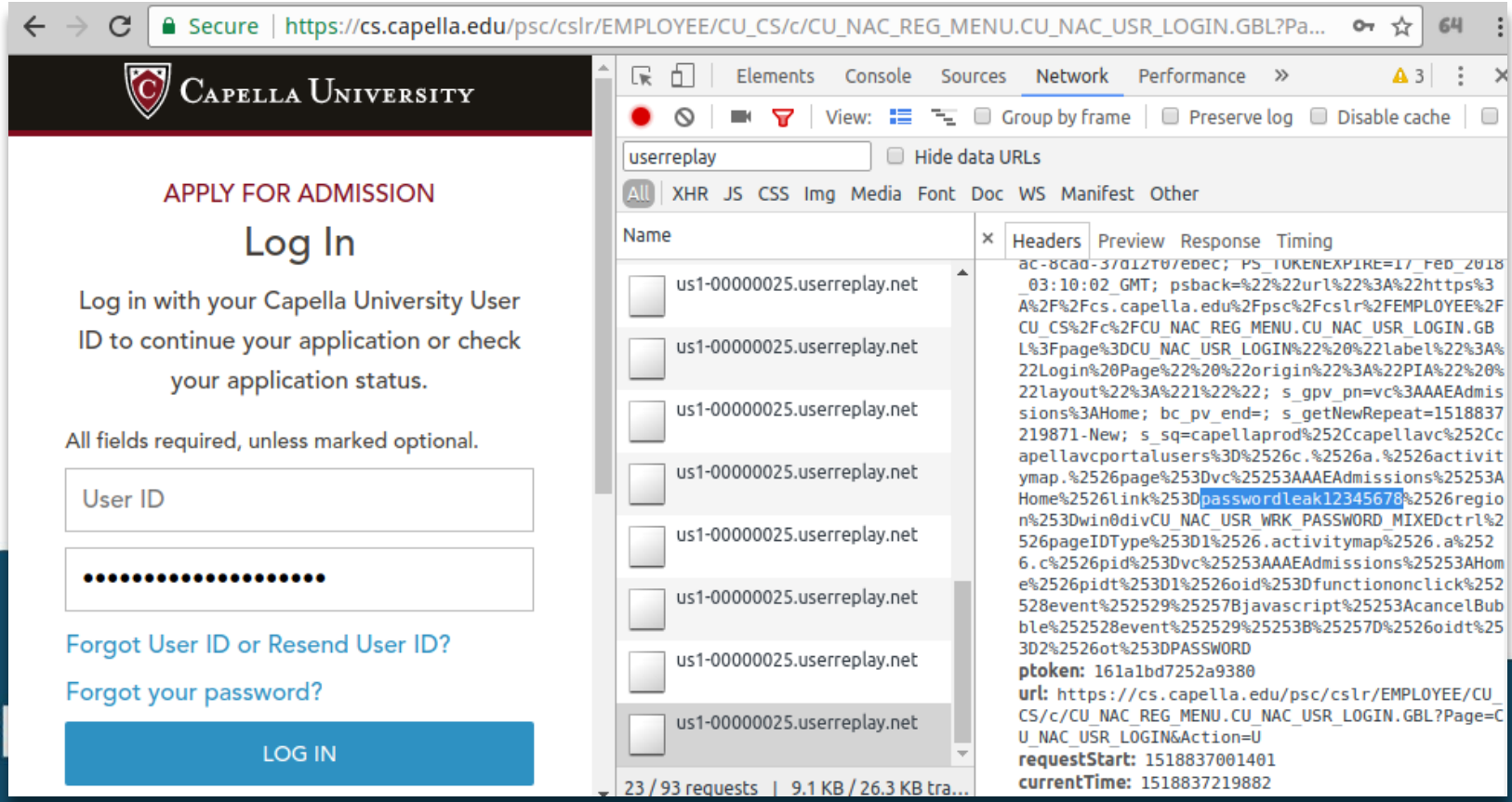
Show password clicked...

`<input type="text">`

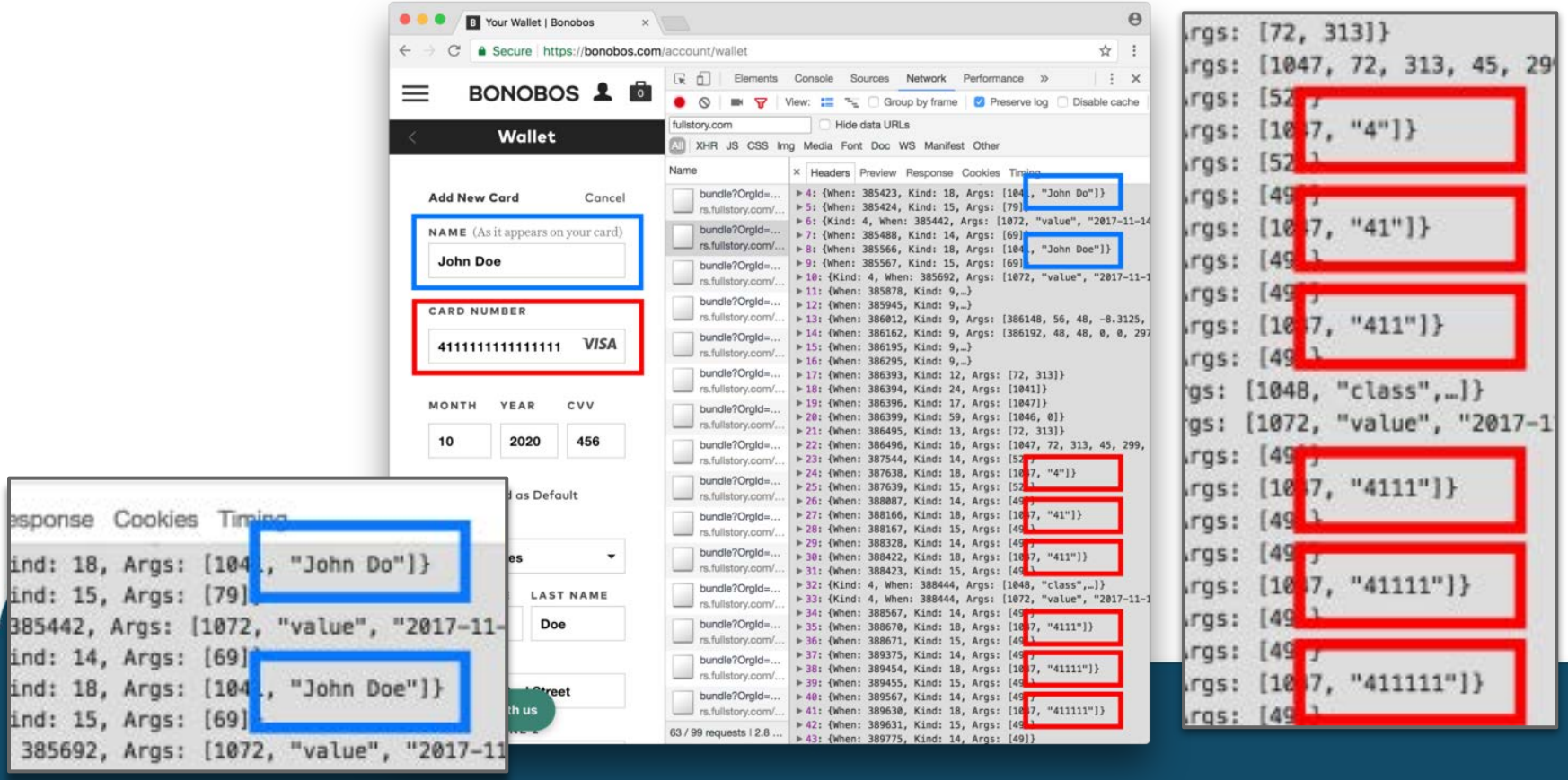
(automatic redaction fails)

Demo from: https://codepen.io/geoffreyrose/pen/HKDkB

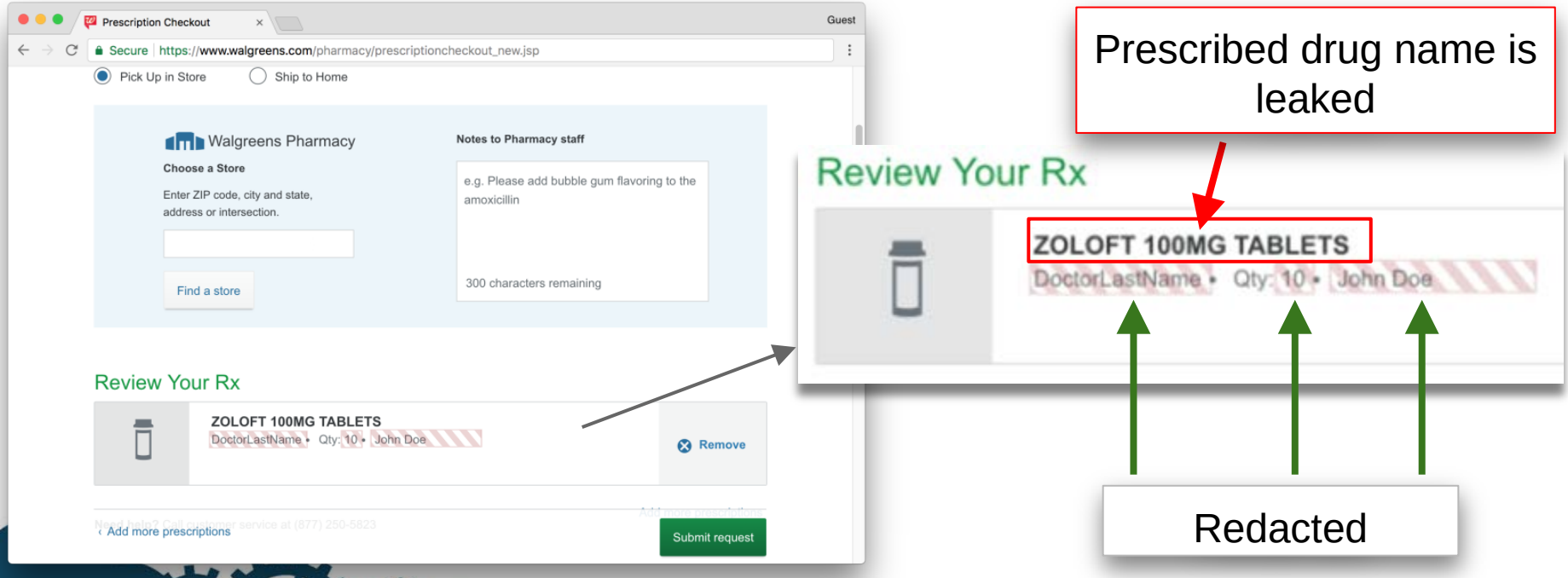
Passwords leak to Userreplay on Capella University



CC#, CVV and name leak to Fullstory on Bonobos



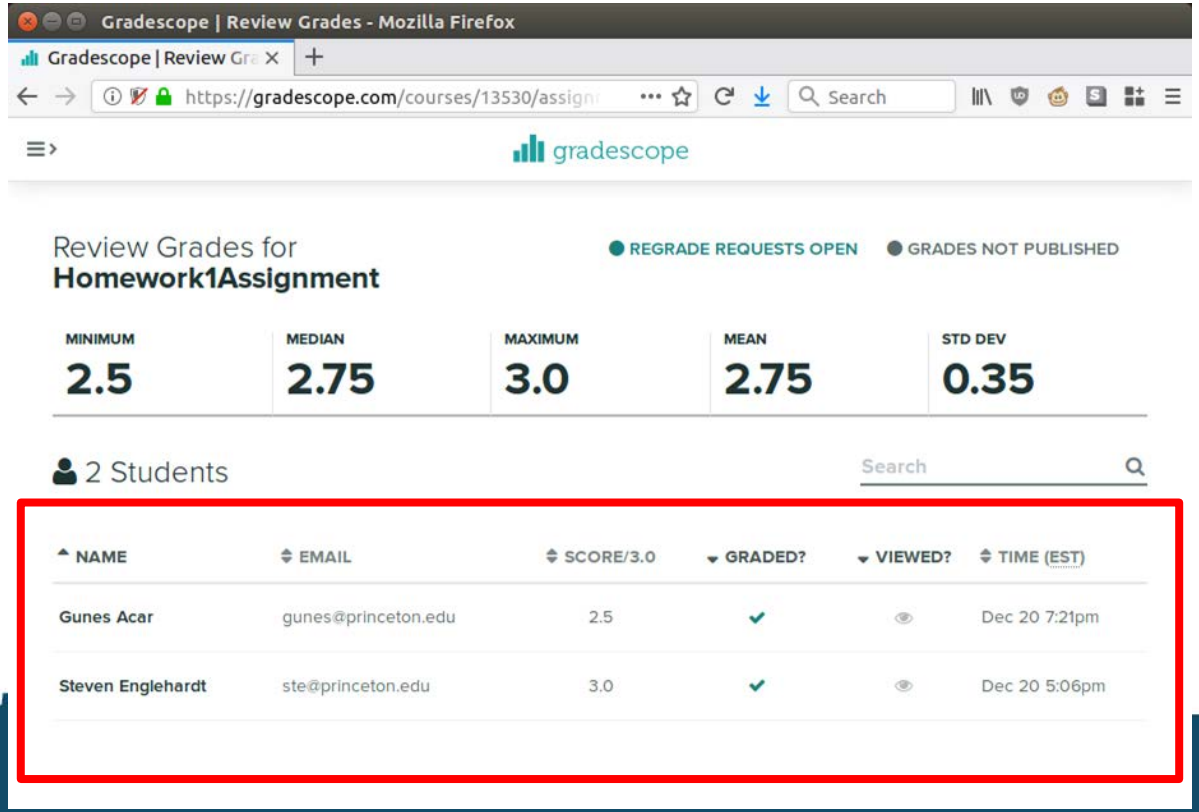
Prescription details leak to Fullstory on Walgreens



Student details leak to Fullstory on Gradescope

Recordings included:

- Student names
- Student emails
- Student grades
- Professor comments



PRIVACYCON

Session recordings are widespread

- 14+ analytics company offer recording services
  - Present on 99,174 of the top 1 million sites
- Evidence of recording on 7,918 sites.
  - Likely a lower bound as recording scripts sample users

Session recording present on ~1 - 10% of the top 1 million sites. We found several severe PII leaks after manually reviewing ~30 sites.

→ **How many more leaks are out there?**

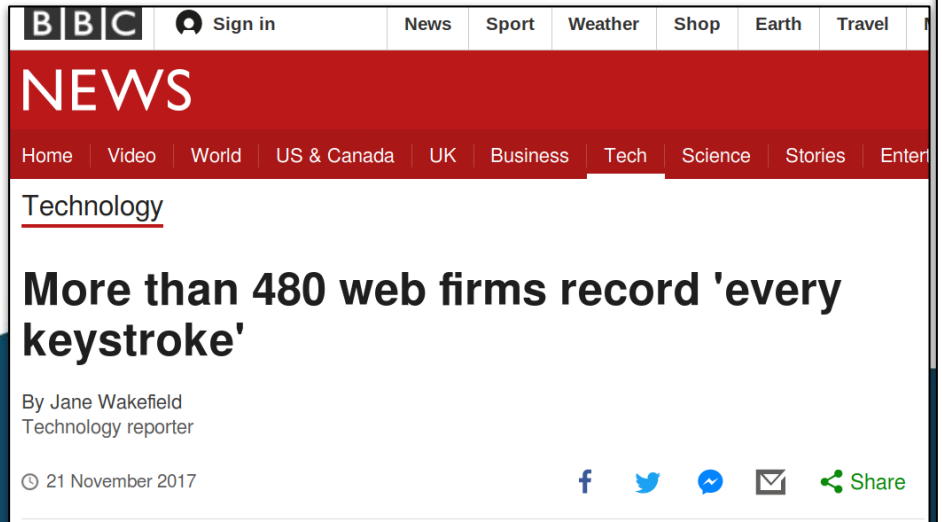
Takeaways

- Recordings contain sensitive information
  - incl. HIPAA, FERPA protected data
- Redaction is difficult and brittle
- Are users comfortable being watched?

Needing to ask “Is this legal?” should give you pause.

Updates

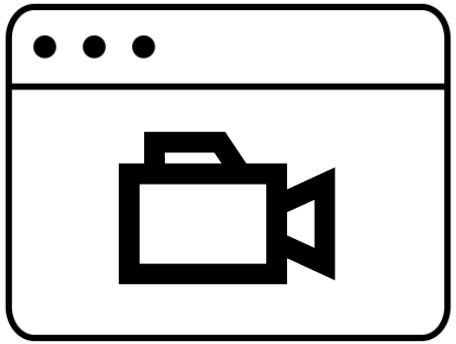
- Walgreens, Gradescope and Bonobos removed session replay script



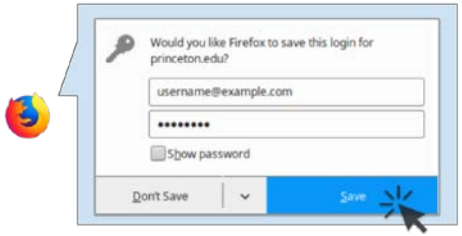
No boundaries series: Privacy vulnerabilities arising from directly embedded third parties

https://freedom-to-tinker.com/tag/noboundaries/

Session Replay



Autofill abuse



And more....

“No boundaries: Exfiltration of personal data by session-replay scripts” (freedom-to-tinker.com)
“No boundaries for user identities: Web trackers exploit browser login managers” (freedom-to-tinker.com)

Thanks for listening!

Contact:
gunes@princeton.edu
https://**gunesacar.net**

Blog post: https://freedom-to-tinker.com/2017/11/15/no-boundaries-exfiltration-of-personal-data-by-session-replay-scripts/

No boundaries series: https://freedom-to-tinker.com/tag/noboundaries/

Princeton Web Transparency & Accountability Project: https://webtap.princeton.edu/

PRIVACYCON

Image assets from the Noun Project: mouse click by Tomas Knopp, Keyboard by Arthur Shlain, recording by Guru, browser windows by DTDesign, HTML File by Burak Kucukparmaksiz, Credit by Por Suppasit, Student by Nimal Raj, Password by BomSymbols, Medicine by UNiCORN, Purchase by Ayub Irawan