Inferring User Behaviors from Log Data for Understanding Computer Security Decisions

Dr. Emilee Rader
Department of Media and Information
Michigan State University
emilee@msu.edu | msu.edu/~emilee

May 14, 2018

- Socio-technical systems: people * technology * information
- “Black boxes”: opaque about how inputs become outputs
- Three types of problems:

1. Privacy issues related to sensors and derived data
   - Emilee Rader and Janine Slaker. “The Importance of Visibility for Folk Theories of Sensor Data” *SOUPS 2017*. https://www.usenix.org/system/files/conference/soups2017/soups2017-rader.pdf
2. Algorithmic decision-making in social media (NSF Grant IIS-1217212)
   - Emilee Rader, Kelley Cotter and Janghee Cho. “Explanations as Mechanisms for Supporting Algorithmic Transparency”. *CHI 2018.* doi: 10.1145/3173574.3173677
3. Computer security decision-making about threats that are hard to be aware of and understand (NSF Grant CNS-1115926)
   - Rick Wash, Emilee Rader, and Chris Fennell. “Can People Self-Report Security Accurately? Agreement Between Self-Report and Behavioral Measures”. *CHI 2017.* doi: 10.1145/3025453.3025911

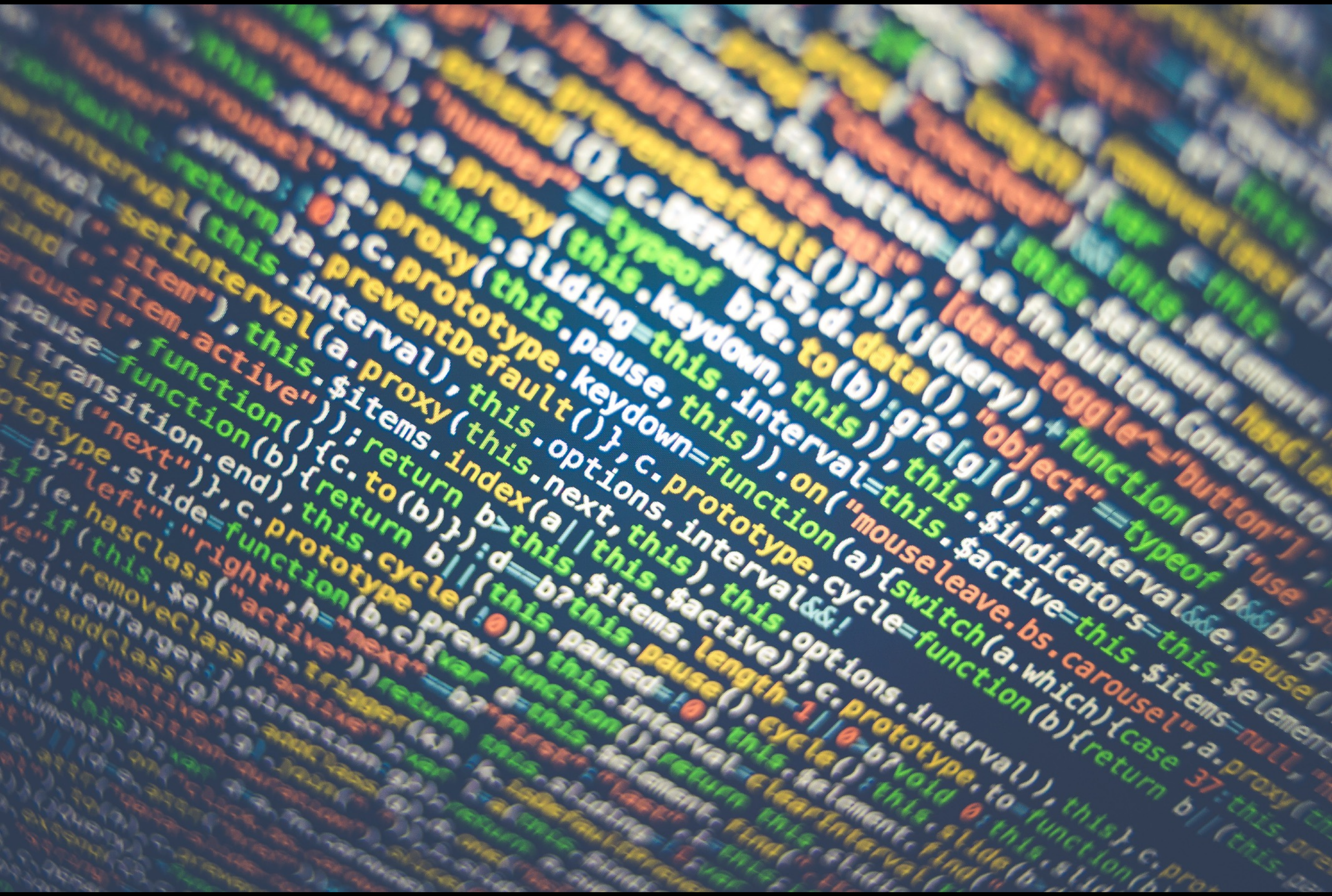


Photo by Markus Spiske — https://www.pexels.com/photo/full-frame-shot-of-multi-colored-pattern-330771/

Everyone faces security decisions on a daily basis...

*****SPAM***** Security Alert Inbox x

Michigan State University thais.santos@univasf.ec Feb 22

to

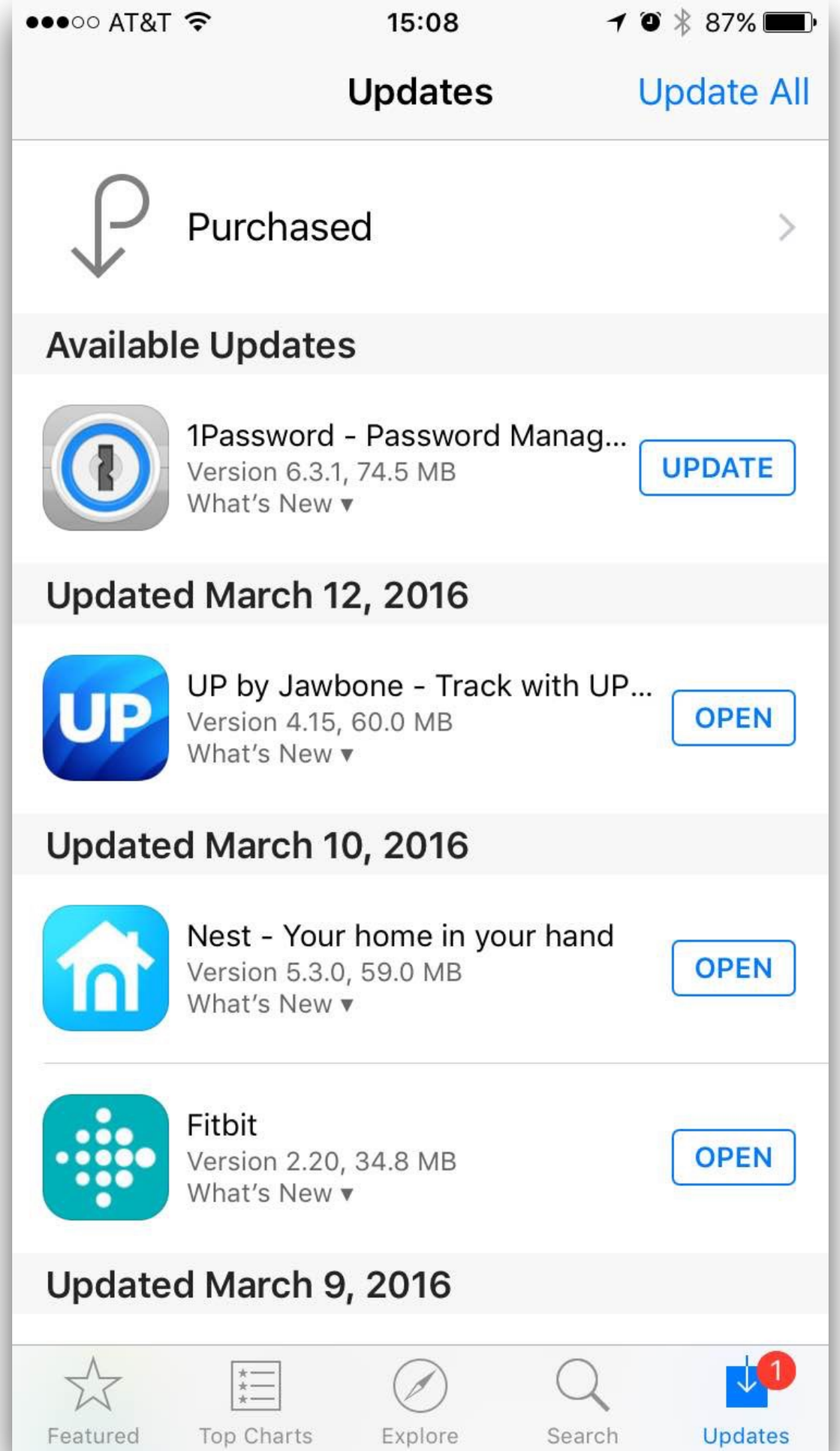
Helpdesk Support Center, Due to congestion in all MSU Net users accounts you need to update your account with our released F-Secure Internet Security 2016. New version of a better resource spam and virus. If you have not upgraded your account, kindly fill in the columns below.

Failure to comply with Cyber-security regulation your MSU NetID account will be temporarily blocked or suspended from our network and you may not be able to receive or send e-mail due to non-compliance.

Full Name*
MSU NetID*
Password*
Confirm Password*

Thank you for your co-operation

© Michigan State University Board of Trustees Inc. East Lansing, MI 48824
Copyright © 2016 msu.edu is an affirmative-action, equal-opportunity employer®


AT&T
15:08
87%
Updates
Update All
Purchased
Available Updates
1Password - Password Manag...
Version 6.3.1, 74.5 MB
What's New
UPDATE
Updated March 12, 2016
UP by Jawbone - Track with UP...
Version 4.15, 60.0 MB
What's New
OPEN
Updated March 10, 2016
Nest - Your home in your hand
Version 5.3.0, 59.0 MB
What's New
OPEN
Fitbit
Version 2.20, 34.8 MB
What's New
OPEN
Updated March 9, 2016
Featured
Top Charts
Explore
Search
Updates
1

Change Password

Help with this page

Select Password | Confirmation

Change your Password — Changing your Password periodically helps ensure the security of your account information.

*** Required field**

Change Password

Enter your current Password, then choose and confirm your new Password.

Your new Password:

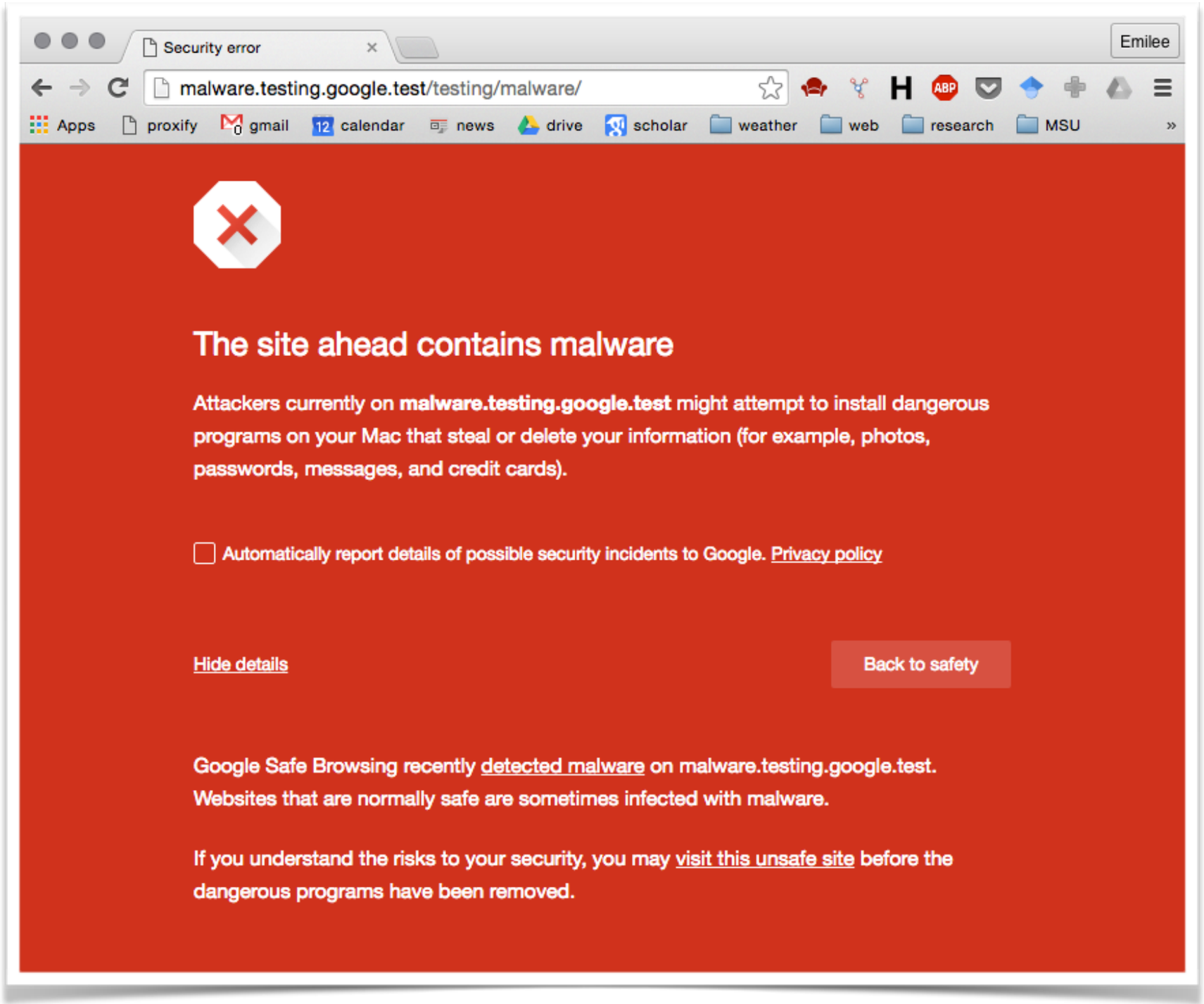
- Must be 8-32 characters long
- Must include at least two of the following elements:
  - At least one letter (upper or lowercase)
  - At least one number
  - At least one special character from the following: # $ % ' ^ , ( ) * + . : | = ? @ /] [_ ` { } \ ! ; - ~
- Must be different than your previous five Passwords
- Must not match your User ID
- Must not include more than 2 identical characters (for example: 111 or aaa)
- Must not include more than 2 consecutive characters (for example: 123 or abc)
- Must not use the name of the financial institution (for example: JPM, MORGAN, CHASE)
- Must not be a commonly used password (for example: password1)

Current Password *

New Password *

Confirm new Password *

Change Password


Security error
Emilee
malware.testing.google.test/testing/malware/
Apps
proxify
gmail
calendar
news
drive
scholar
weather
web
research
MSU
The site ahead contains malware
Attackers currently on malware.testing.google.test might attempt to install dangerous programs on your Mac that steal or delete your information (for example, photos, passwords, messages, and credit cards).
Automatically report details of possible security incidents to Google. Privacy policy
Hide details
Back to safety
Google Safe Browsing recently detected malware on malware.testing.google.test. Websites that are normally safe are sometimes infected with malware.
If you understand the risks to your security, you may visit this unsafe site before the dangerous programs have been removed.

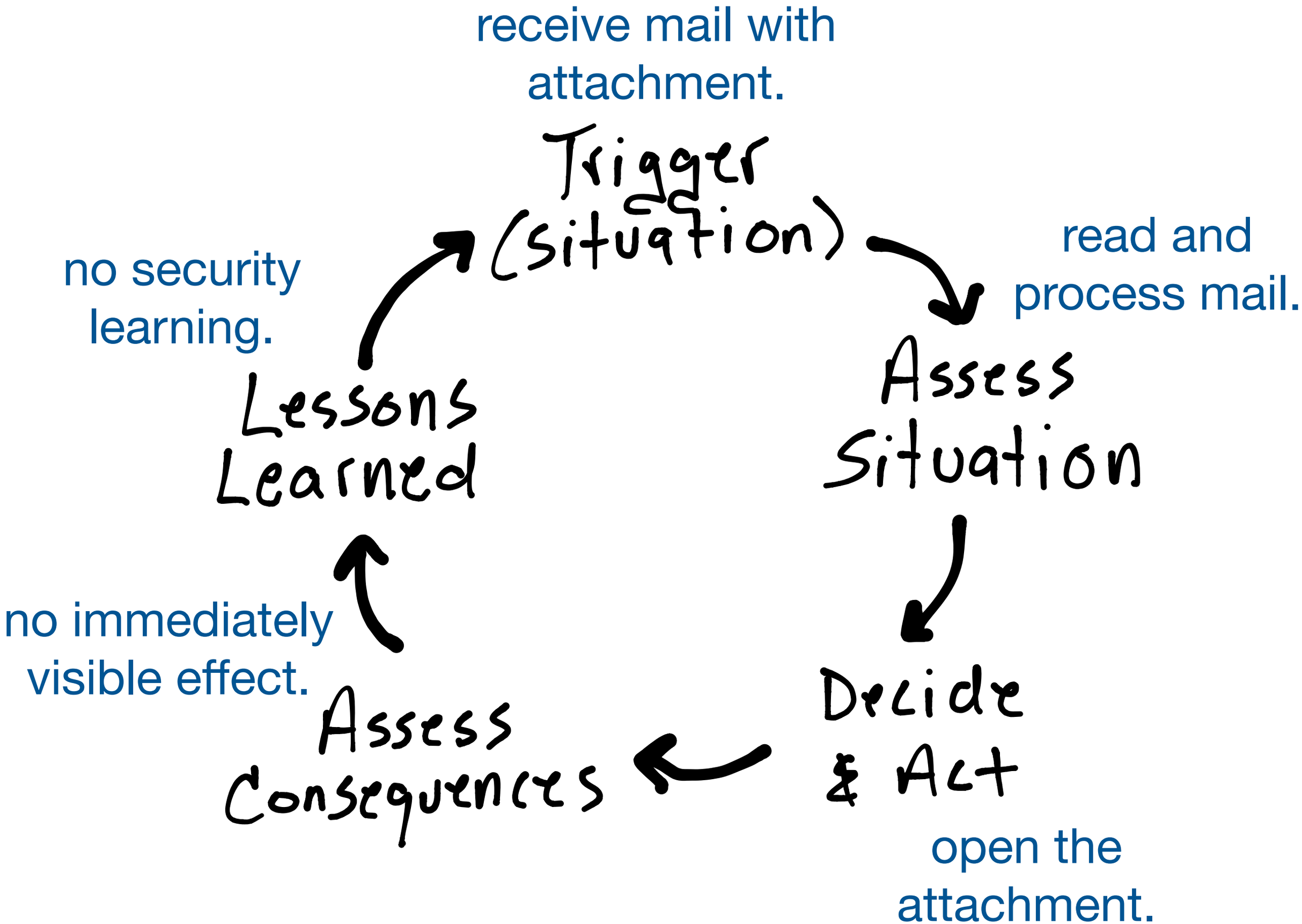
everyday computer users: people without training in computer science or security who use computing technology and the Internet

A large proportion of attacks on the Internet **target vulnerabilities in end users** rather than vulnerabilities in technology (*Symantec*)

The majority of computers are compromised using vulnerabilities **for which a security update was available** but had not yet been installed (*Microsoft*)

A system's security depends on the choices made by its users.

One way to influence users' choices is to influence what they know about security.



Adapted from: Marsick VJ, Watkins KE. Informal and incidental learning. New Dir Adult Contin Educ 2001; 25–34.

Locky ransomware activity ticks up

Locky is now one of the most commonly seen types of ransomware

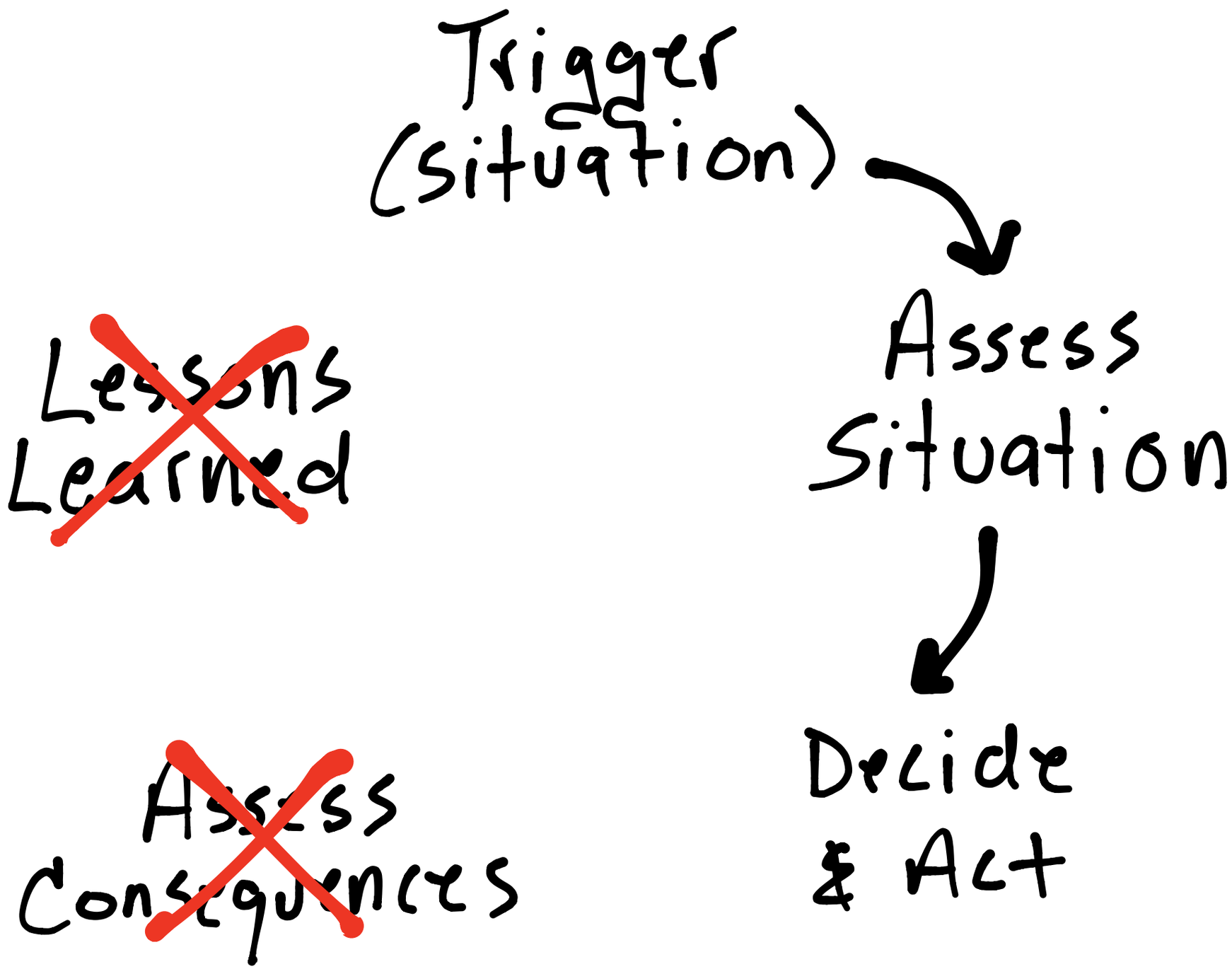
```
!!! IMPORTANT INFORMATION !!!!

All of your files are encrypted with RSA-2048 and AES-128 ciphers.
More information about the RSA and AES can be found here:
    http://en.wikipedia.org/wiki/RSA_(cryptosystem)
    http://en.wikipedia.org/wiki/Advanced_Encryption_Standard

Decrypting of your files is only possible with the private key and decrypt program, which is on our secret server.
To receive your private key follow one of the links:
    1. http://twbers4hmi6dx65f.tor2web.org/[redacted]
    2. http://twbers4hmi6dx65f.onion.to/[redacted]
    3. http://twbers4hmi6dx65f.onion.cab/[redacted]

If all of this addresses are not available, follow these steps:
    1. Download and install Tor Browser: https://www.torproject.org/download/download-easy.html
    2. After a successful installation, run the browser and wait for initialization.
    3. Type in the address bar: twbers4hmi6dx65f.onion/[redacted]
    4. Follow the instructions on the site.

!!! Your personal identification ID: [redacted] !!!
```

A ransomware program called Locky has quickly become one of the most common types of malware seen in spam. Credit: McAfee

2 COMMENTS

Jeremy Kirk
IDG News Service Mar 10, 2016 3:50 AM

Source: http://www.pcworld.com/article/3042580/security/locky-ransomware-activity-ticks-up.html



Adapted from: Marsick VJ, Watkins KE. Informal and incidental learning. New Dir Adult Contin Educ 2001; 25–34.

The challenge: how to connect what people think and know about security, with the outcomes of the choices they make!

How did we study this?

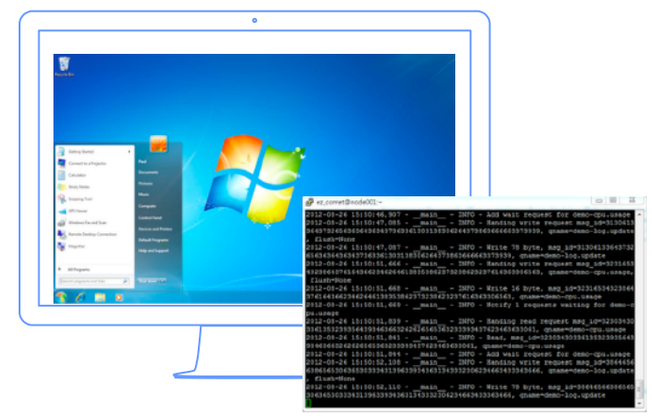
- Custom software development
  - Windows app (C# and PowerShell)
  - Web browser plugins for Firefox and Chrome (JavaScript)
  - Server software (PHP)
  - LOTS of analysis scripts (Python, MySQL, R)
- Six-week data collection
  - 134 university students (excluding CS and Engineering)
  - 53% Women, 46% Men
  - $70 compensation

How did we study this?

Participants

Pre-Survey

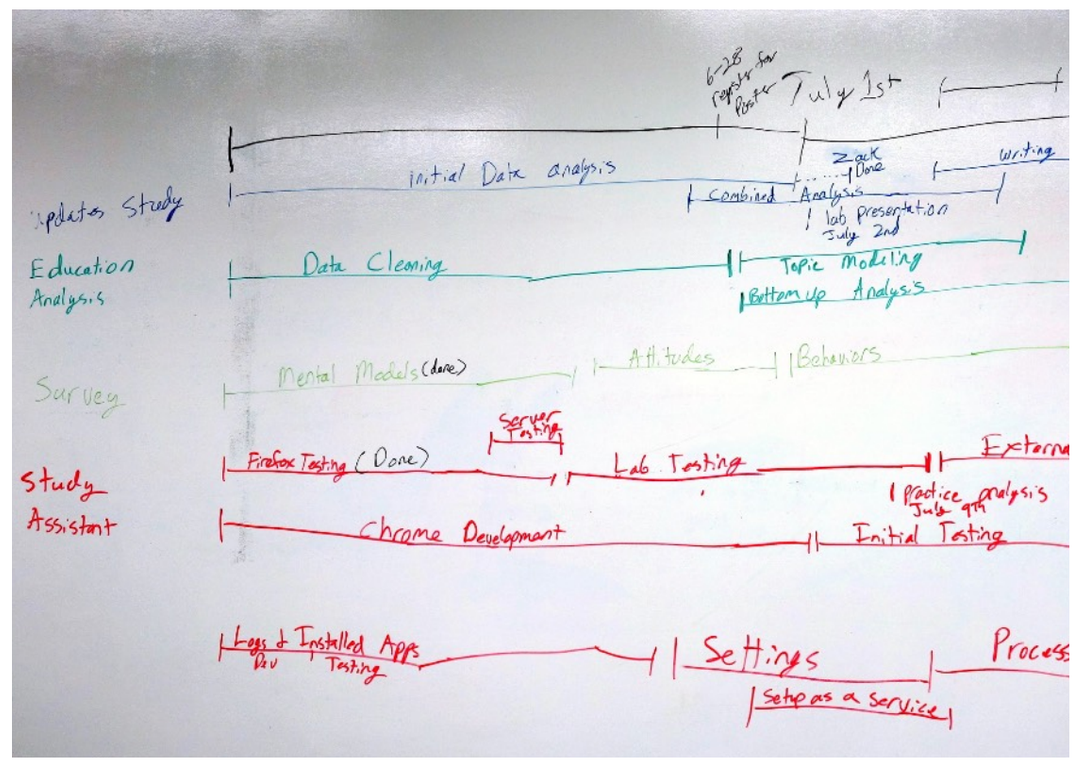
Custom Logging Software



Post Survey


Delaying updates
Does when the update becomes available impact install time
What causes People to Delay Updates
Difference between types of updates
General attitudes toward updates
Annoying so don't update (annoyed + rebellious) sketched out?
Auto-update attitude
Suspicious
faithful
uncertain
Why do People turn off Auto-updates?
Annoying, but important and/or updates anyway (annoyed + faithful)
timing
random
schedulers
post-poners
Annoying but doesn't know/care if they update or not (annoyed + unaware)


Cron: Check Install
Registration
Initial Survey
Quotients
Subject
US
Not Eligible
Display Install Instructions
Ping Server
DB Table
Redmine Entry
Survey Done; Need Software
Email US
Rejoice!
Email Subject
Activate Browser Plugin(s)
Install Browser Plugins
Update Status to Admin
Data Collection Normal
Software Partial Install
Inserting DB


July 1st
initial Data analysis
Zack Done
Writing
Updates Study
Combined Analysis
lab presentation July 2nd
Education Analysis
Data Cleaning
Topic Modeling
Bottom up Analysis
Survey
Mental Models (done)
Attitudes
Behaviors
Server Testing
Study Assistant
Firefox Testing (Done)
Lab Testing
Practice analysis July 4th
Chrome Development
Initial Testing
Logs & Installed Apps
Testing
Settings
Setup as a Service
Process

Custom Web Browser Extensions

- What is a browser extension, anyway?
- Data we collected:
  - all URLs visited
  - download events
  - installed plugins and extensions
  - all passwords (hashed!) and the webpage visits they were associated with
  - from that we reconstructed browsing sessions

about 774,000 visits
to 300,000 different distinct URLs
14,000 downloads
24,000 password entries
150,000 browser add-ons

Custom Windows App

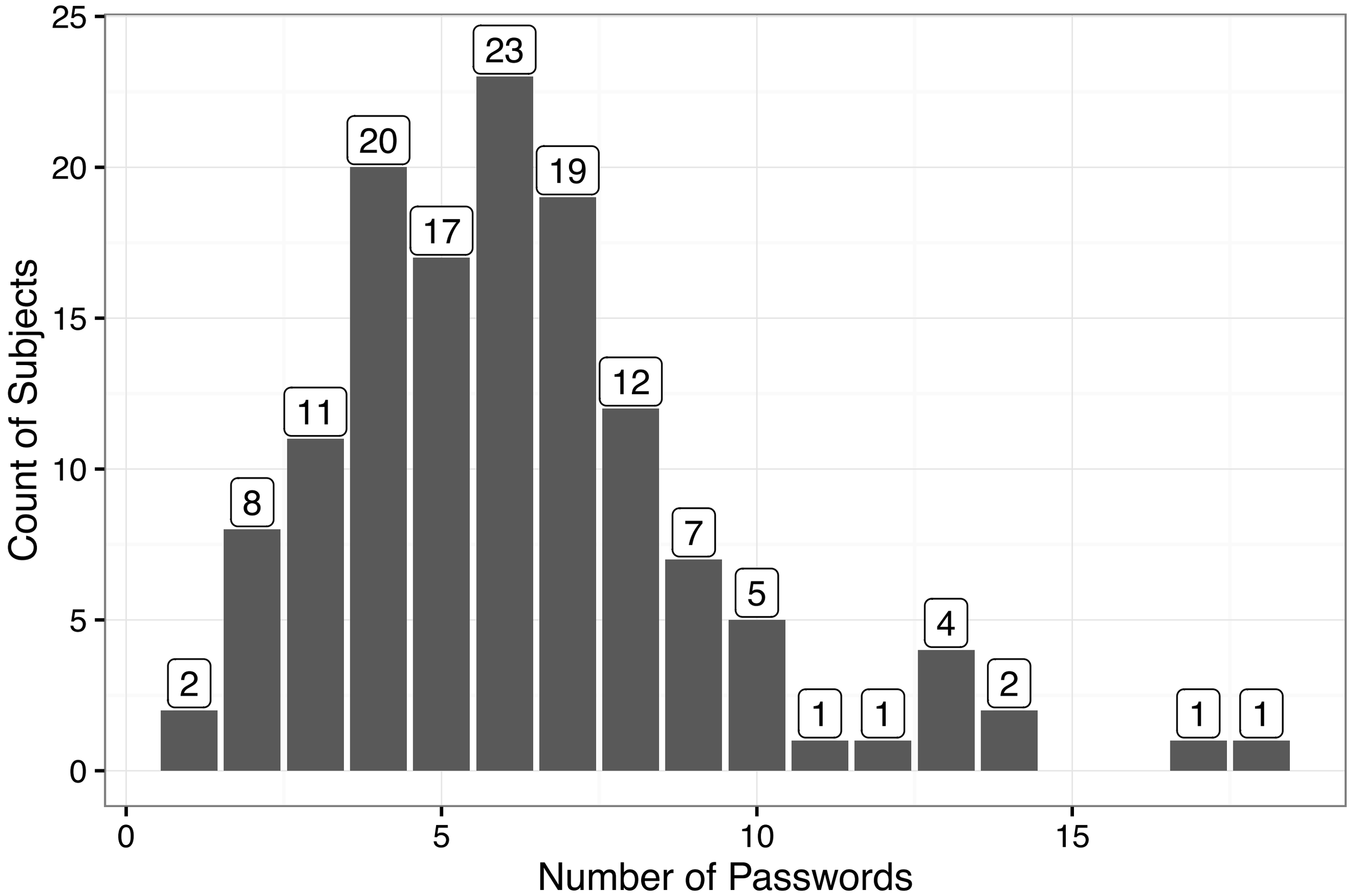
- Windows can log a lot of stuff for developers…
- We turned all those logs on and collected data from them:
  - all processes that ran on the participants' computers
  - software installed
  - security settings
  - wifi and firewall logs
  - logon log
  - hardware and OS information
  - Windows (software) update information
  - crashes and shutdowns
  - and more…

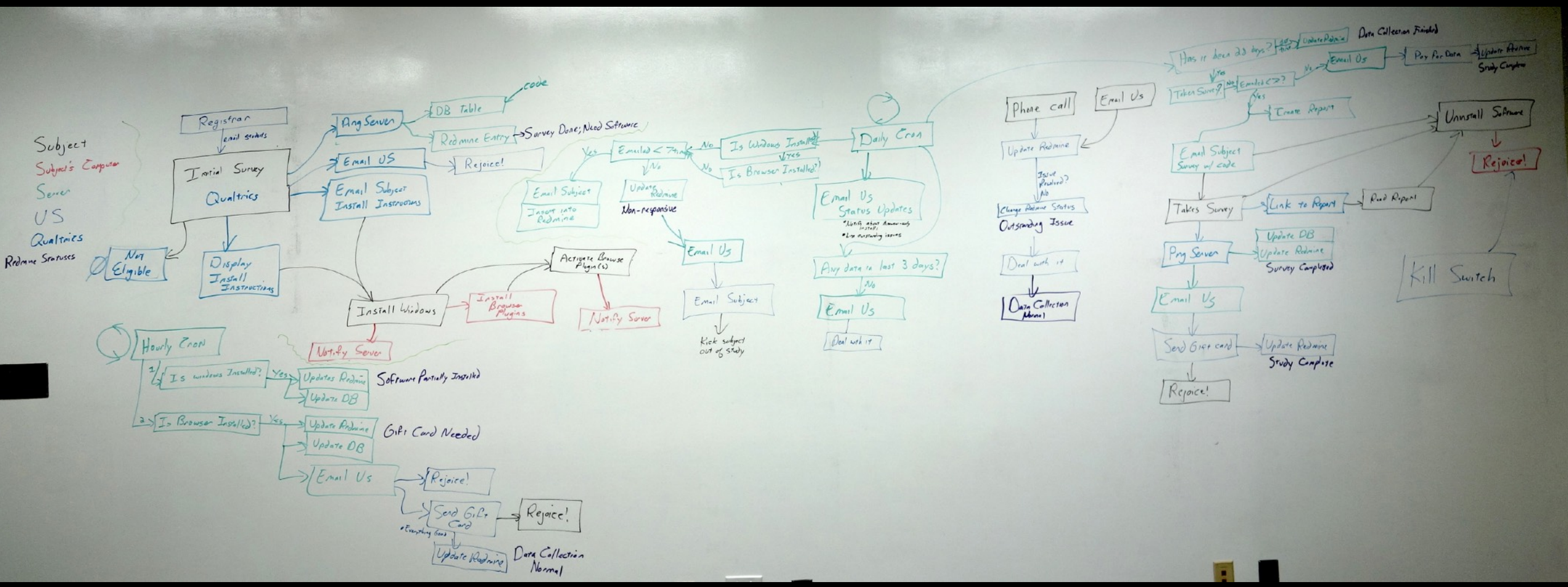
1.5 million installed applications
11 million processes run
120,000 wifi connections
70,000 windows updates installed

Server Software and Database

- Why did we need a server application?
  - Link browser plugin data and windows app data with participant survey data
  - Process the data and store it in the database
- Why a backend database?
  - Well, what's the alternative?
  - Think about it as lots of spreadsheets that reference each other…

security	65568	29.56 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
server_log	4661592	33.74 GB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
study_monitor_log	2488	144.00 KB	InnoDB	2016-09-05 16:33:58	latin1_swedish_ci
study_summary	189	48.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
tab_fork	89126	32.56 MB	InnoDB	2016-09-05 16:33:58	latin1_swedish_ci
users	190	48.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
visit_facts	965808	88.61 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
visits	964761	103.62 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_backup_log	218	1.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_computer_hardware	366	80.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_current_desktops	1983	320.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_defender_settings	211	64.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_error_reporting_log	426743	669.00 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_firewall	33970	7.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_group_user	324090	709.00 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_installed_applications	1470924	448.98 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_installed_apps	27135	14.52 MB	InnoDB	2016-09-05 16:33:58	latin1_swedish_ci
win_installed_products	0	16.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_log	60	16.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_logical_disks	22921	2.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_logon_log	384444	1.64 GB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_ms_antimalware_log	83140	56.59 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_operating_system	530	160.00 KB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_pnp_log	13857	11.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_power_log	809542	350.95 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_process_log	273644	10.60 GB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_processor_log	1245107	178.72 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_security_products	139768	15.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_security_settings	34563	5.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_server_log	701409	86.03 GB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_status	7559	1.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_study_assistant_log	326345	1013.00 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_sys_restore_log	17663	8.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_update_api	84394	19.55 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_update_client_event_log	333268	1.15 GB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_update_wmi	44121	5.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_user_account	34979	4.52 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci
win_wifi_log	655464	483.00 MB	InnoDB	2016-09-05 16:33:58	utf8_general_ci


Count of Subjects
Number of Passwords
0
5
10
15
20
25
2
8
11
20
17
23
19
12
7
5
1
1
4
2
1
1


Subject
Subject's Computer
Server
US
Qualtrics
Redmine Statuses
Registran
email sent
Initial Survey
Qualtrics
Not Eligible
Display Install Instructions
Ping Server
DB Table
code
Redmine Entry
Survey Done; Need Software
Email US
Rejoice!
Email Subject
Install Instructions
Install Windows
Install Browser Plugins
Activate Browse Plugin(s)
Notify Server
Notify Server
Hourly Cron
1
Is windows Installed?
Yes
Updates Redmine
Update DB
Software Partially Installed
2
Is Browser Installed?
Yes
Update Redmine
Update DB
Gift Card Needed
Email Us
Rejoice!
Send Gift Card
Rejoice!
Everything Good
Update Redmine
Data Collection Normal
Emailed < 7 times?
No
Is Windows Installed?
Yes
No
Is Browser Installed?
Yes
Email Subject
Insert into Redmine
Update Redmine
Non-responsive
Email Us
Email Subject
Kick subject out of study
Daily Cron
Email Us
Status Updates
Notify about Automatically Installs
Log outstanding issues
Any data in last 3 days?
No
Email Us
Deal with it
Phone call
Email Us
Update Redmine
Issue Resolved?
No
Change Redmine Status
Outstanding Issue
Deal with it
Data Collection Normal
Has it been 28 days?
Update Redmine
Data Collection Finished
Taken Survey?
No
Emailed < 7?
Yes
Email Us
Pay for Data
Update Redmine
Study Complete
Yes
Create Report
Email Subject
Survey w/ code
Takes Survey
Link to Report
Read Report
Ping Server
Update DB
Update Redmine
Survey Completed
Email Us
Send Gift card
Update Redmine
Study Complete
Rejoice!
Unnstall Software
Rejoice!
Kill Switch

Privacy and Ethics Issues

Informed Consent

- IRB approval for “spyware”
- Multiple users on a single machine
- Giving people the ability to turn off the data collection
- What is the right amount to compensate people?

Privacy and Log Data

- Logging browsing activity
  - sensitive activities
  - illegal activities
- Logging passwords
  - risk of compromise
  - password reuse

Privacy and Log Data

- Logging Windows operating system data
  - software update state
  - installed software and versions
  - anti-virus installed, in use?
  - time spent doing certain activities

Anonymization

- "Data can be perfectly useful or perfectly anonymous but never both" —Paul Ohm
- What does "identifiable" data look like?
- What log data might be identifiable?
- What might participants not want us to infer about them?

Sharing and Reproducibility

- Our dataset is a snapshot in time
- Our custom software is brittle
- Risk of re-identification
- How to share code, datasets?
- How to prevent unintended uses?
- Long-term storage issues

OSF | Influencing Mental Mode ×

Emilee

Secure | https://osf.io/m8svp/

Apps | proxify | gmail | calendar | drive | D2L | scholar | ravelry | Merriam-Webster | weather | MSU | misc | knitting | Other Bookmarks

OSF HOME ▾ Search Support

https://osf.io/m8svp/

Influencing Mental Models of ... | Files | Wiki | Analytics | Registrations

Influencing Mental Models of Computer Security

Public | 0 | ...

Contributors: Emilee Rader, Rick Wash

Date created: 2016-10-20 12:05 PM | Last Updated: 2016-10-28 06:43 PM

Category: Project

Description: This project investigates how mental models of computer security are formed, how ideas and information about computer security are incorporated into mental models, and how they are transmitted from person to person. It measures the prevalence of different mental models and correlates them with logs of actual security behaviors. Through these investigations, this project seeks to characterize the reasons that many everyday computer users choose not to act securely — a question which is one of the biggest challenges of computer security.

Files

Filter

Name	Modified
Influencing Mental Models of Comp...	
– OSF Storage	
– Replication Materials for SOUPS ...	
– OSF Storage	
+ analysis files	
codebook.csv	2016-10-28 10:5...
+ data collection code	

Citation

osf.io/m8svp

Components

Replication Materials for SOUPS 2016 paper "Understanding Password Choices: How Frequently Entered Passwords are Re-used Across Websites"

Rader & Wash

Replication Materials for CHI 2017 paper "Can People Self-Report

What did we learn?

Current technologies make it difficult for individuals to learn about security:

- Automating the install of software updates makes it harder for people to learn how to make decisions about updates because there are fewer opportunities to learn [SOUPS 2014].
- More knowledge about security or technical issues is not associated with more secure behavior [SOUPS 2015].
- People can only accurately self-report security behaviors that are discrete and have visible outcomes [CHI 2017].

What did we learn?

People generalize security learning from one system to other, technically unrelated systems:

- Negative experiences with software updates create spillover, or a refusal to install even unrelated updates [CHI 2014].
- People re-use passwords they must enter frequently on many other websites, most likely because it is easiest to recall [SOUPS 2016].

References

[CHI 2014] Vaniea, K., Rader, E., and Wash, R. “Betrayed By Updates: How Negative Experiences Affect Future Security”. DOI: 10.1145/2556288.2557275

[SOUPS 2014] Wash, R., Rader, E., Vaniea, K, and Rizor, M. “Out of the Loop: How Automated Software Updates Cause Unintended Security Consequences”. https://www.usenix.org/system/files/soups14-paper-wash.pdf

[SOUPS 2015] Wash R. and Rader, E. “Too Much Knowledge? Security Beliefs and Protective Behaviors Among US Internet Users”. https://www.usenix.org/system/files/conference/soups2015/soups15-paper-wash.pdf

[SOUPS 2016] Wash, R., Rader, E., Berman, R., and Wellmer, Z. “Understanding Password Choices: How Frequently Entered Passwords are Re-used Across Websites”. https://www.usenix.org/system/files/conference/soups2016/soups2016-paper-wash.pdf

[CHI 2017] Wash, R., Rader, E., and Fennell, C. “Can People Self-Report Security Accurately? Agreement Between Self-Report and Behavioral Measures”. DOI: 10.1145/3025453.3025911

How did I learn to do all this stuff?

- A long time ago, I took a couple of programming courses
- To learn, I relied a LOT on code other people had written
- Worked with (or near!) people who knew more than me and asked a LOT of questions
- Came up with projects that were interesting enough to me that I needed to learn these things
- Made a lot of mistakes, learned from them, got better
- A lot of this is learning about how to organize the work and what I should do myself vs. what I should hire or find collaborators to do…

Thank you!

Dr. Emilee Rader
Department of Media and Information
Michigan State University
emilee@msu.edu | msu.edu/~emilee

This material is based upon work supported by the National Science Foundation under Grants CNS-1115926, CNS-1116544

Special thanks to collaborators and co-authors on this work: Rick Wash, Brandon Brooks, Nate Zemanek, Chris Fennell, Kami Vaniea, Michelle Rizor, Katie Hoban, and the rest of the BITLab team.