



# 2023 PRIVACY & IOT RESEARCH WORKSHOP

Shubhangi Waldiya, Jonah  
Melgosa

# TEAM PEPPER



**Shubhangi Waldiya**

**Major: Computer Science**

**Attends: El Camino College  
& South High School**



**Jonah Melgosa**

**Major: Industrial Engineering**

**Attends: Long Beach City  
College**

AROUND 82 PERCENT OF  
WEB TRAFFIC CONTAINS  
THIRD-PARTY SCRIPTS BY  
GOOGLE, AND ALMOST HALF  
OF THEM ARE CONFIRMED  
TO BE TRACKING YOU

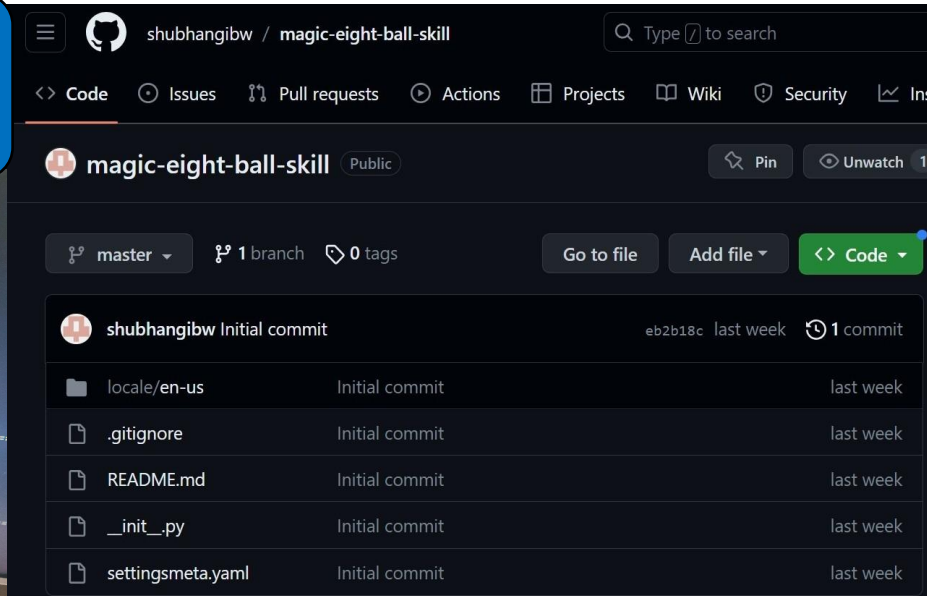
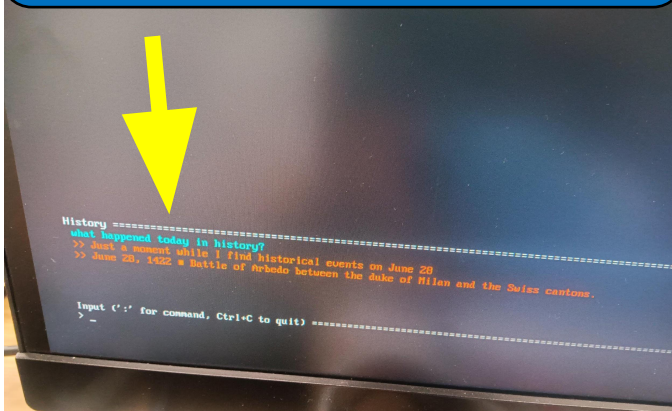
*(Source:  
<https://www.cloudwards.net/data-privacy-statistics/>)*

# INTRODUCTION

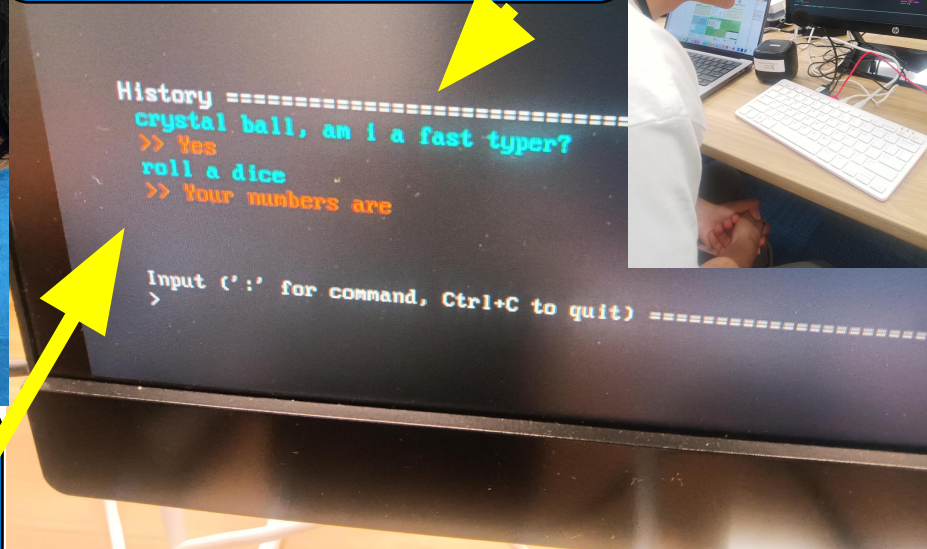
The growth of technology has led to an increased concern about privacy. Through this two-week program we learned:

- ✓ About Mycroft skills
- ✓ Tracking & how to prevent it
- ✓ What data is collected in Alexa vs Mycroft
- ✓ Data collected by TikTok on the web versus mobile
- ✓ About ads and how to make our own ad blocker

Today In History skill  
downloaded from Mycroft.ai



Crystal ball skill  
downloaded from GitHub



Dice skill downloaded  
from Mycroft.ai

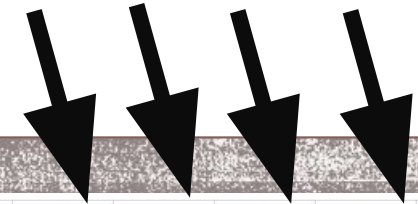
# MYCROFT SKILLS

- Installation of skills from Mycroft.ai and GitHub
  - Dice
  - Crystal ball
- Created our own skills
  - Magic eight ball
  - Hunger

# TRACKING IN ALEXA VS MYCROFT

	G	H	I	J	K	L	M	N	O	P	Q	R
1	tcp_stream	hostname	path	headers	pii_types	package_name	piholeblocklist_d	stopad_abp_blo	moaab_abp_blo	firebog_advertising_tracking_abp_block_decision		
2		9	del-ta-g7g.amazon.com		[]		0	0	0	0		
3		8	mlis.amazon.com		[]		0	0	0	0		
4		21	mlis.amazon.com		[]		0	0	0	0		
5		27	mlis.amazon.com		[]		0	0	0	0		
6		7	mlis.amazon.com		[]		0	0	0	0		
7		37	mlis.amazon.com		[]		0	0	0	0		
8		11	mlis.amazon.com		[]		0	0	0	0		
9		20	mlis.amazon.com		[]		0	0	0	0		
10		30	mlis.amazon.com		[]		0	0	0	0		
11		15	mlis.amazon.com		[]		0	0	0	0		
12		22	mlis.amazon.com		[]		0	0	0	0		
13		44	mlis.amazon.com		[]		0	0	0	0		
14		6	mlis.amazon.com		[]		0	0	0	0		
15		17	api.amazonalexa.com		[]		0	0	0	0		
16		49	mlis.amazon.com		[]		0	0	0	0		
17		28	mlis.amazon.com		[]		0	0	0	0		
18		2	mlis.amazon.com		[]		0	0	0	0		

**In Mycroft, there were no "1"s in the following columns, indicating no tracking**



	G	H	I	J	K	L	M	N	O	P	Q	R
1	tcp_stream	hostname	path	headers	pii_types	package_name	piholeblocklist_d	stopad_abp_blo	moaab_abp_blo	firebog_advertising_tracking_abp_block_decision		
2		3	api.mycroft.ai		[]		0	0	0	0		
3		0	api.mycroft.ai		[]		0	0	0	0		
4		7	mimic-api.mycroft.ai		[]		0	0	0	0		
5		10	api.mycroft.ai		[]		0	0	0	0		
6		12	api.mycroft.ai		[]		0	0	0	0		
7		5	api.mycroft.ai		[]		0	0	0	0		
8		4	api.mycroft.ai		[]		0	0	0	0		
9		8	api.mycroft.ai		[]		0	0	0	0		
10		13	api.mycroft.ai		[]		0	0	0	0		
11		9	api.mycroft.ai		[]		0	0	0	0		
12		1	api.mycroft.ai		[]		0	0	0	0		
13		14	api.mycroft.ai		[]		0	0	0	0		
14		2	api.mycroft.ai		[]		0	0	0	0		

**In Alexa, there were also no "1"s however, if the data was collected for longer, there would be more "1"s than in Mycroft**



*Adblock Plus*



# TRACKING & HOW TO PREVENT IT



*Ublock Origin*

Advertising

❖ Behavior

❖ Contextual

Tracking

❖ Stateful

❖ First party cookies

❖ Third party cookies

❖ Stateless

❖ Fingerprinting



*Opera  
Browser*

Methods of prevention:

□ Ad blockers

□ Privacy browsers



*AdGuard*

*Brave  
Browser*



# DIY AD BLOCKER

The blank space  
seen here is a side  
effect of manually  
blocking ads



## Steps:

1. Right click on website
2. Click on Network tab
3. Catch me if you can

## Pros:

- Will work everytime the website is opened on that browser
- Learn about Network Tab

## Cons:

- Website may still maintain the ad slot (the gray area).
- May cause some frustration





# TRACKING IN TIKTOK

## Web vs Mobile

- Lesser ads
- More ads
- More difficulty changing interest
- Easier to change interest
- Less data types used for tracking
- More data types used for tracking

	A	B	C	D	E	F
58	browser_language	en-us	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
59	browser_name	mozilla	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
60	browser_online	TRUE	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
61	browser_platform	linux aarch64	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
62	browser_version	5.0 (x11; cros aarch64 13597.84.0) ar	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
63	channel	tiktok_web	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
64	cookie_enabled	TRUE	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
65	count	1	query string	2	1076	0 webcast.us.tiktok.com
66	device_id	7231975087179204142	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
67	device_platform	web_pc	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
68	focus_state	FALSE	query string	5	1076	0 tiktok.com;;webcast.us.tiktok.com
69	from_page	typ	query string	8	1076	0.01 tiktok.com;;webcast.us.tiktok.com
70	history_len		query string	6	1076	0.01 tiktok.com;;webcast.us.tiktok.com
71	is_fullscreen	FALSE	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
72	is_page_visible	TRUE	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
73	os	linux	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com
74	page		query string	1	1076	0 webcast.us.tiktok.com
75	priority_region	us	query string	49	1076	0.05 tiktok.com;;webcast.us.tiktok.com

Figure 1: csv file for web

The data types highlighted in yellow can be used for tracking. There are more in the mobile version

	A	B	C	D	E	F
24	ssmix	a	query string	266	875	0.3 api16-core-useast5.us.tiktokv.com;;a
25	device_type	pixel 3	query string	273	875	0.31 api16-core-useast5.us.tiktokv.com;;a
26	device_brand	google	query string	268	875	0.31 api16-core-useast5.us.tiktokv.com;;a
27	language	en	query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
28	os_api		29 query string	268	875	0.31 api16-core-useast5.us.tiktokv.com;;a
29	os_version		10 query string	269	875	0.31 api16-core-useast5.us.tiktokv.com;;a
30	manifest_version_code	2022903040	query string	268	875	0.31 api16-core-useast5.us.tiktokv.com;;a
31	resolution	1080*2028	query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
32	dpi		440 query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
33	update_version_code	2022903040	query string	273	875	0.31 api16-core-useast5.us.tiktokv.com;;a
34	_rticket	1686177164307	query string	1	875	0 api16-normal-useast5.us.tiktokv.com;;a
35	current_region	us	query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
36	app_type	normal	query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
37	mc_mnc		310260 query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
38	timezone_name	america/los_angeles	query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
39	carrier_region_v2		310 query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
40	residence	us	query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a
41	app_language	en	query string	267	875	0.31 api16-core-useast5.us.tiktokv.com;;a

Figure 2: csv file for mobile

# CHALLENGES, RESULTS & OVERALL EXPERIENCE

## Challenges

- Trying to work on this presentation simultaneously
- Setting up Mycroft
- Collecting data from Mycroft and Alexa
  - Naming the files for the .py scripts to run
- Collecting TikTok data
  - The files didn't stop collecting data after export

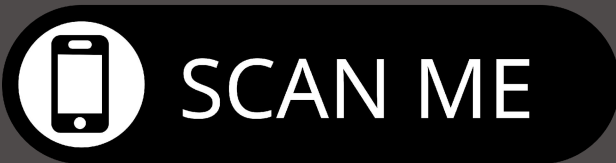
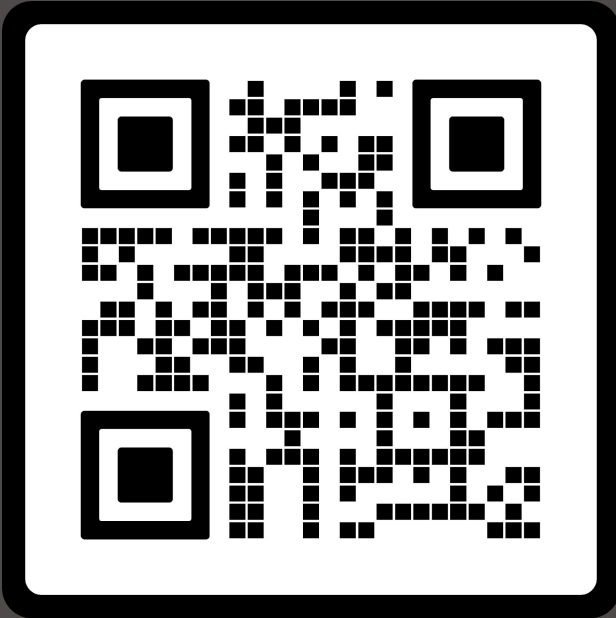
## Results

- Learned more about ads, tracking and how to prevent it
- Knowledge of linux commands, using PuTTY and ssh
- Hands on experience with network analysis, collecting data, running scripts, creating Mycroft skills

## Overall Experience

- Got to learn from PhD research students and professors with Proper Data
- Fun, enjoyable, interesting
- Would recommend to my peers for next year

Jonah's LinkedIn:



**Shubhangi's email:**

[shubhangiwaldiya@gmail.com](mailto:shubhangiwaldiya@gmail.com)