black hat®
USA 2021

August 4-5, 2021

BRIEFINGS

Use & Abuse of Personal Information

Alan J. Michaels, PhD
Kiernan B. George

Use & Abuse of Personal Information

Merriam-Webster

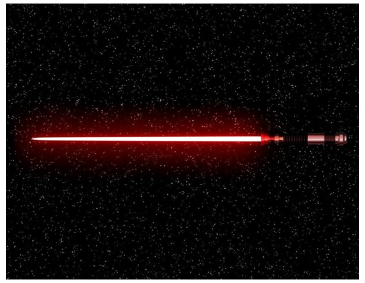
Use, /yo͞os/ : a method or manner of employing or applying something

Abuse, /əˈbyo͞os/ : to put to a *wrong* or *improper* use

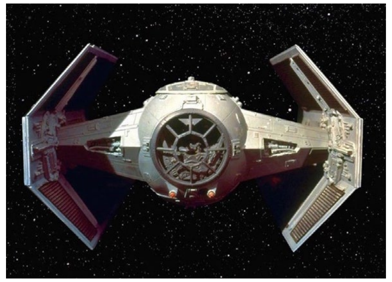
: to use *excessively*

Darth Vader
Sith Lord (ret.)
anakin@aol.com
DS: 01-407-939-5277
1977 Fortress Vader; Mustafar
Force affiliation: Dark

USE



Supplies



Transportation

THE IMPERIAL NEWS
IT IS A DARK TIME FOR THE REBELLION
ALTHOUGH THE DEATH STAR HAS BEEN DESTROYED, IMPERIAL TROOPS HAVE DRIVEN THE REBEL FORCES FROM THEIR HIDDEN BASE

News

ABUSE

BUY WITH CONFIDENCE
gerryWood AUTO GROUP
1 Month/1,000 Mile Warranty
100% of Labor and 100% Parts Covered

Used Car Warranty

Tax Scams

Excessive “News”

black hat
USA 2021

Use & Abuse Team

Alan J. Michaels, PhD

Kiernan B. George

VT | HUME CENTER FOR NATIONAL SECURITY AND TECHNOLOGY
VIRGINIA TECH.

ELECTRONIC SYSTEMS LAB

Lauren Anderson

Harrison Bui

Cara Dunnavant

Piper Hancock

Joe Harrison

Joshua Lyons

Maya Jackson

Clare Mathewes

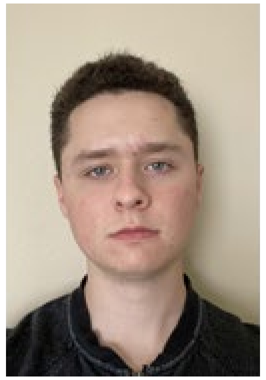
Lauren Maunder

Paul O'Donnell

Sarah Ramboyong

Allie Schliefer

Brian Timana-Gomez

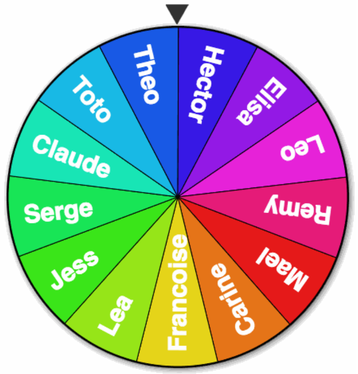

Brandon Vanek

Multiple semesters. Additional acknowledgements to Hume Center faculty: Chelsy Ables (animations), Deanna Clark / Tiasha Khan (program managers), and Ryan Chase (IT support)

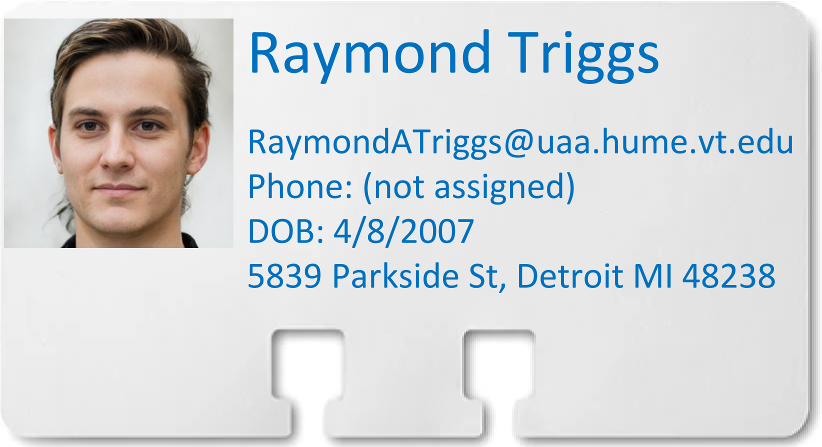
Personal Information

300 distinct fake identities based on domestic “averages”

This person does not exist



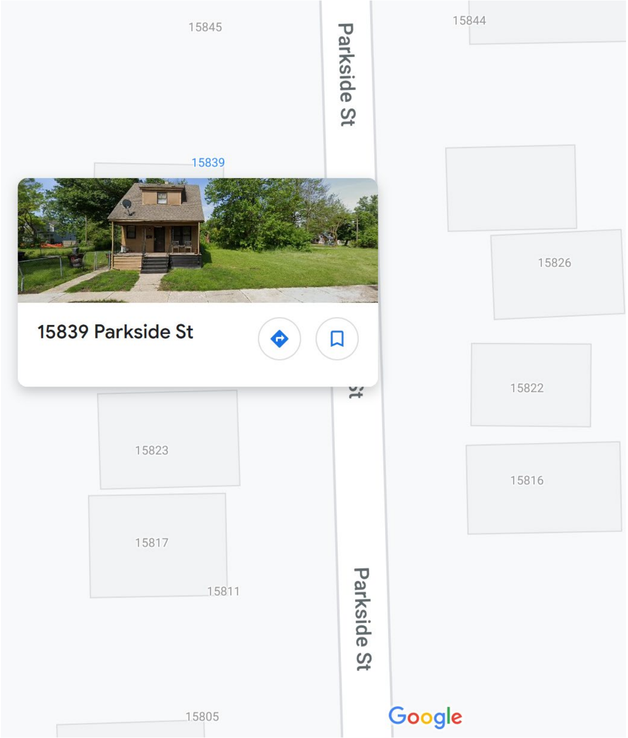
Random name generator



Raymond Triggs

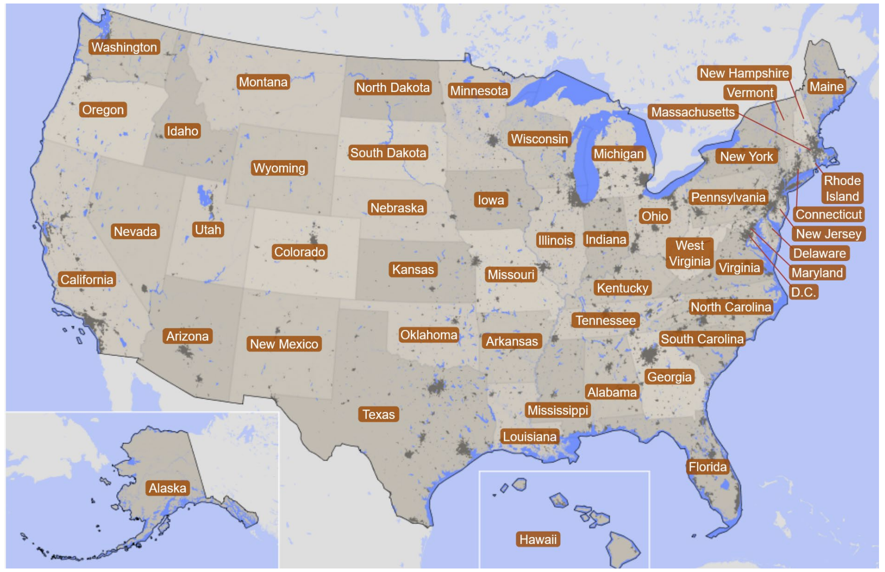
RaymondATriggs@uaa.hume.vt.edu
Phone: (not assigned)
DOB: 4/8/2007
5839 Parkside St, Detroit MI 48238

Identity



15839 Parkside St

Modifications of real addresses



Statistical atlas

Political affiliation

Organizations

Companies have been color coded according to their economic sectors:					
Consumer Staples/Defensive	Communication Services	Industrials	Political Organizations	Real Estate	Other
Online Retail/Cyclical	Hospitality	News/Media	Software/Technology	Restaurants	

Domestic Organizations						
7-11	CNET	Flikr	Kohls	Papa John's	Target	Yelp
ACLU	CNN	Food Lion	Kroger	Pepsi	The Guardian	Yidio
ACM	Coca Cola	Food Network	LA Times	PETA	Tiktok	YouTube/Google
Adidas	Collegiate Times	Fox	LinkedIn	Pinterest	Tim Kaine (VA Senator)	Zillow
Alesis	Communist Party	Free Movies	Lowes	Planned Parenthood	Toyota	Zoom
Amazon********	Consumer Report*	g2a	Lyft	Player Auctions*	Trip Advisor	
American Airlines	Costco	Glassdoor	Macy’s	Poshmark	Trulia	
Apple	Csgv.org	Go fund me	Marriot	Pro-Life Action League	Tumblr	
Atlanta JCC	CVS	Godaddy	Match	Putlocker	Twitch	
Autotrader	Dccc.org	Gop.gov	McDonalds	Quizlet	Twitter	
Autozone	Delta	Green Peace	Medium	Quora	Uber	
BBC	Denver Post	Groupon	Miami Herald	Realtor	UNICEF	
Bed Bath and Beyond	Discord	Healthline	Michaels	Reddit*	US News	
Best Buy*	Dollar tree	Hi5	Microsoft	Retail Me Not	USA Today	
Bleacher Report	Dominos	Hilton	Mitch McConnell	Roanoke Times	VA Citizens Defense League	
Breitbart	DonaldjTrump.com	Home Depot*	Motor Mile	Rotten Tomatoes	Walgreens	
bstock	Dunkin Donuts	Huffington Post	Moviesjoy	Safeway	Walmart	
Business Insider	Ebay	IKEA	MSN	Sheetz	Wayfair	
Carmax	eHarmony	IMDB	NAACP	Slack	WeatherBug	
Cars.com	ESPN	Indeed	Netflix	Spotify	WebMD	
Carvana	Etsy	Instagram	New York Times	Squarespace	Wendys*	
CD Keys	Expedia	Jimmy Johns	Newegg	Starbucks	Whatsapp	
Chicago Tribune	Exxon	Joebiden.com	Nike	Steam	Wish*	
Chick-Fil-A	Facebook	Kickass Torrents	NPR	StubHub	Wix	
Chipotle	Family Research Council	Kinguin	Panera	Taco Bell*	Yahoo News	

Foreign Organizations	
20 Minutes	PesaPal
Alibaba	Rakuten
Asos	RuTube
Badoo	Shaadi
Cookpad	Sputnik
Discovery Store	Stuff
Douban	Taringa
Goalzz (KOOORA)	Tokopedia
Hatena	Toutiao
Hudson Bay	Twoo
JD Sports	VZ
KrisShop	XING
Leboncoin	Yandex
Lefigaro	Yandex Disk
Millat Facebook	Zhanqi
Ouest France	

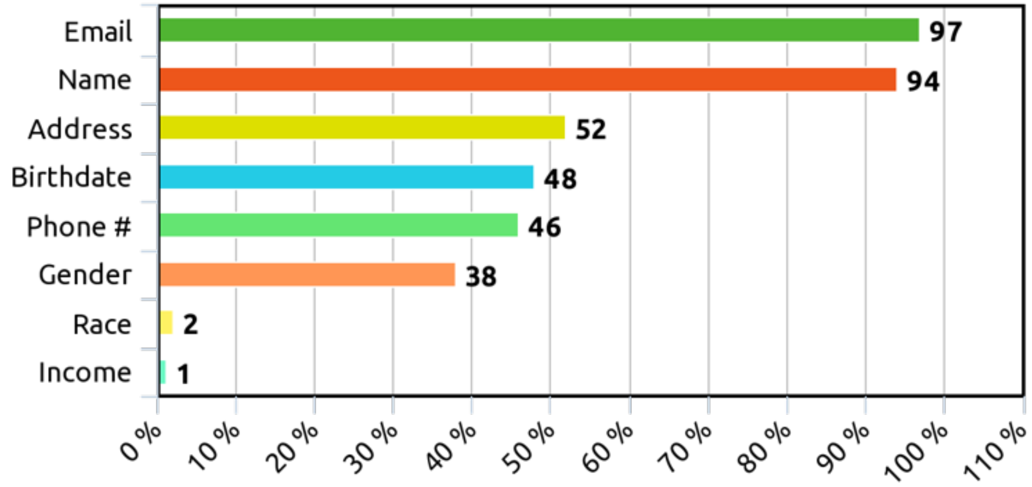
* indicates a financial transaction
Multiple asterisks indicate multiple transactions, each by a different fake identity

Personal Information

Mostly random assignment of 300 identities to a list of 185 companies

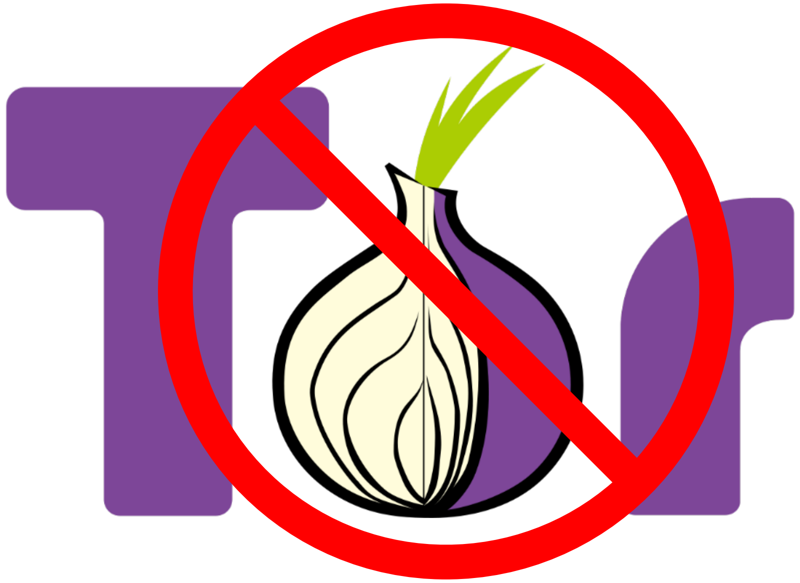
- **One-time transaction**: sign up for account, newsletter, request information, make purchase, etc
- Gave any *personal* information they would take during transaction

PII requested by accounts (%)



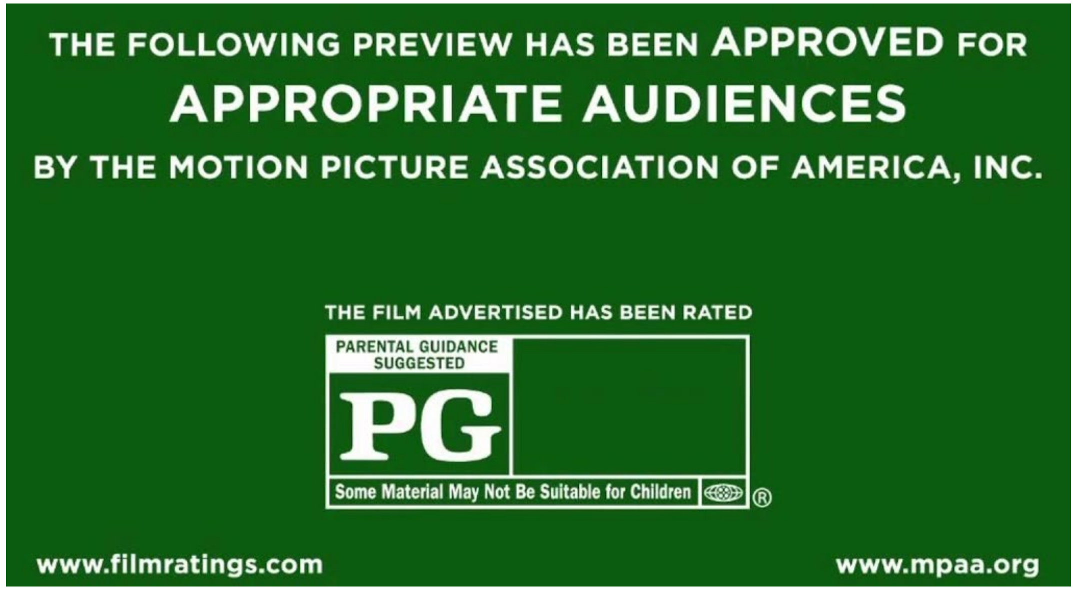
Constraints

SOCIAL SECURITY
123-45-6789
THIS NUMBER HAS BEEN ESTABLISHED FOR
JOHN Q PUBLIC
SIGNATURE



FEDERAL ELECTION COMMISSION
UNITED STATES OF AMERICA

Passive collection only



THE FOLLOWING PREVIEW HAS BEEN APPROVED FOR
APPROPRIATE AUDIENCES
BY THE MOTION PICTURE ASSOCIATION OF AMERICA, INC.

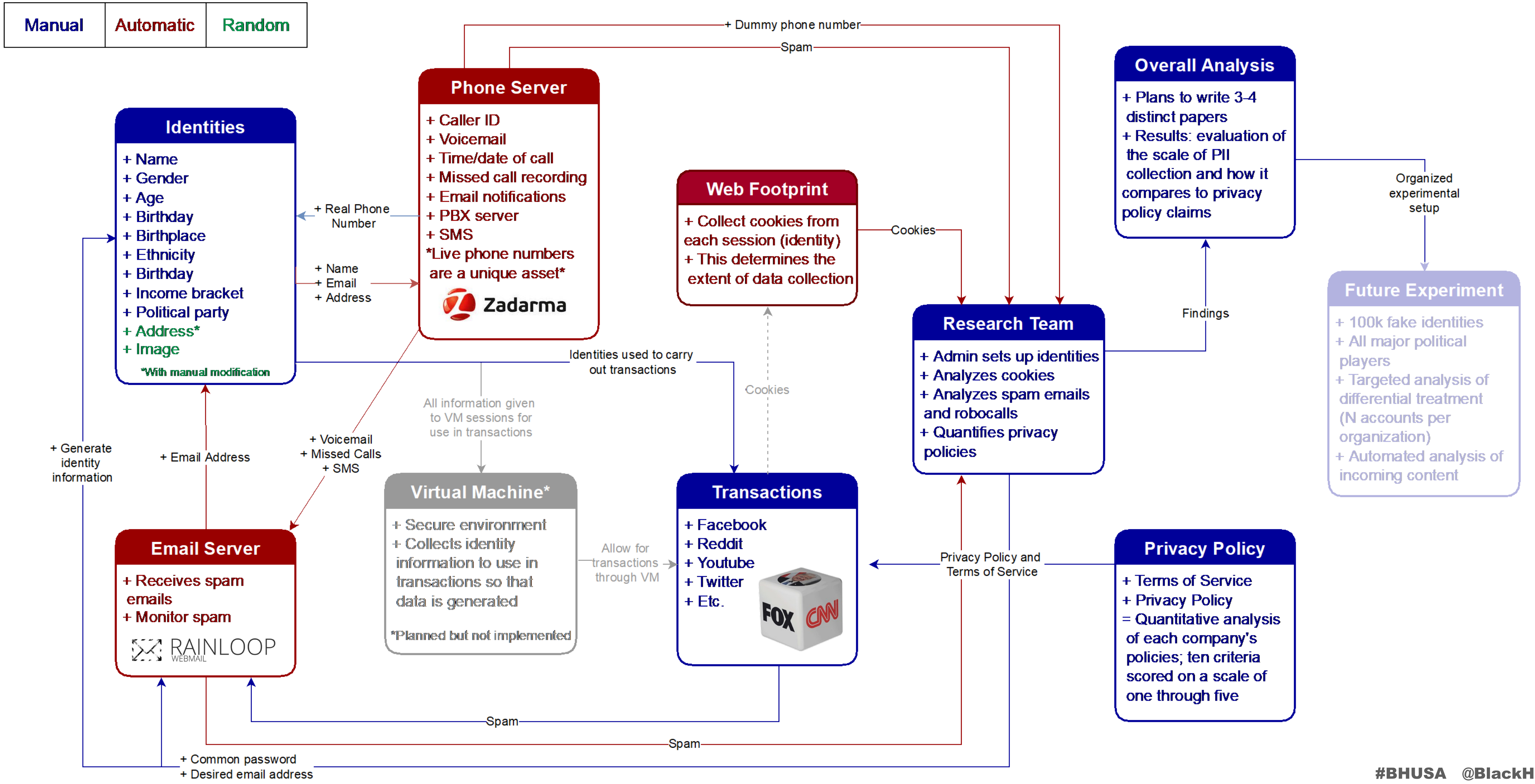
THE FILM ADVERTISED HAS BEEN RATED

PARENTAL GUIDANCE SUGGESTED
PG
Some Material May Not Be Suitable for Children

www.filmratings.com

www.mpaa.org

Experimental Setup



Manual	Automatic	Random

Identities

- + Name
- + Gender
- + Age
- + Birthday
- + Birthplace
- + Ethnicity
- + Birthday
- + Income bracket
- + Political party
- + Address*
- + Image

*With manual modification

Phone Server

- + Caller ID
- + Voicemail
- + Time/date of call
- + Missed call recording
- + Email notifications
- + PBX server
- + SMS

Live phone numbers are a unique asset

Zadarma

Web Footprint

- + Collect cookies from each session (identity)
- + This determines the extent of data collection

Overall Analysis

- + Plans to write 3-4 distinct papers
- + Results: evaluation of the scale of PII collection and how it compares to privacy policy claims

Future Experiment

- + 100k fake identities
- + All major political players
- + Targeted analysis of differential treatment (N accounts per organization)
- + Automated analysis of incoming content

Research Team

- + Admin sets up identities
- + Analyzes cookies
- + Analyzes spam emails and robocalls
- + Quantifies privacy policies

Virtual Machine*

- + Secure environment
- + Collects identity information to use in transactions so that data is generated

*Planned but not implemented

Transactions

- + Facebook
- + Reddit
- + Youtube
- + Twitter
- + Etc.

Email Server

- + Receives spam emails
- + Monitor spam

RAINLOOP WEBMAIL

Privacy Policy

- + Terms of Service
- + Privacy Policy
- = Quantitative analysis of each company's policies; ten criteria scored on a scale of one through five

+ Dummy phone number

Spam

+ Real Phone Number

+ Name
+ Email
+ Address

Cookies

Organized experimental setup

Findings

Identities used to carry out transactions

All information given to VM sessions for use in transactions

Cookies

+ Generate identity information

+ Email Address

+ Voicemail
+ Missed Calls
+ SMS

Allow for transactions through VM

Privacy Policy and Terms of Service

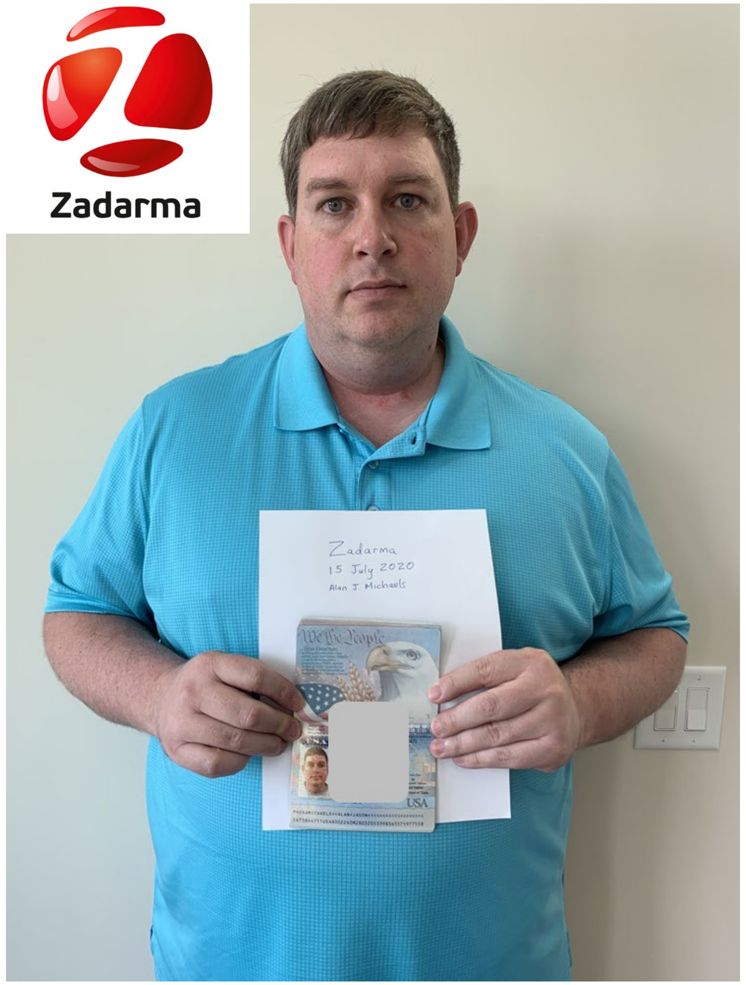
Spam

Spam

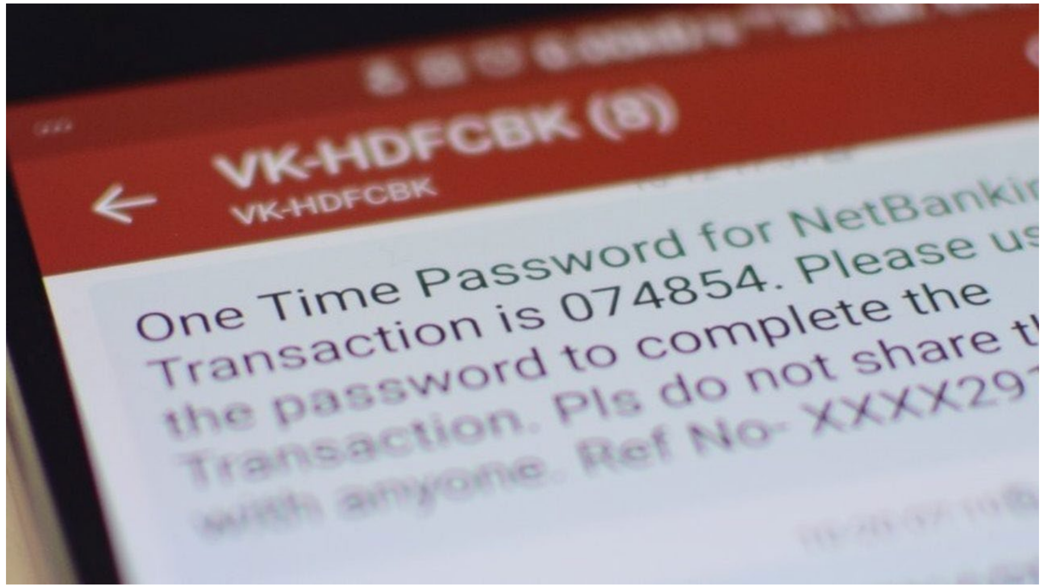
+ Common password
+ Desired email address

Lessons Learned

Iteration #6…



Erratic account verification methods


VK-HDFCBK (8)
VK-HDFCBK

One Time Password for NetBanki
Transaction is 074854. Please us
the password to complete the
Transaction. Pls do not share t
with anyone. Ref No- XXXX29

Financial transactions are a pain

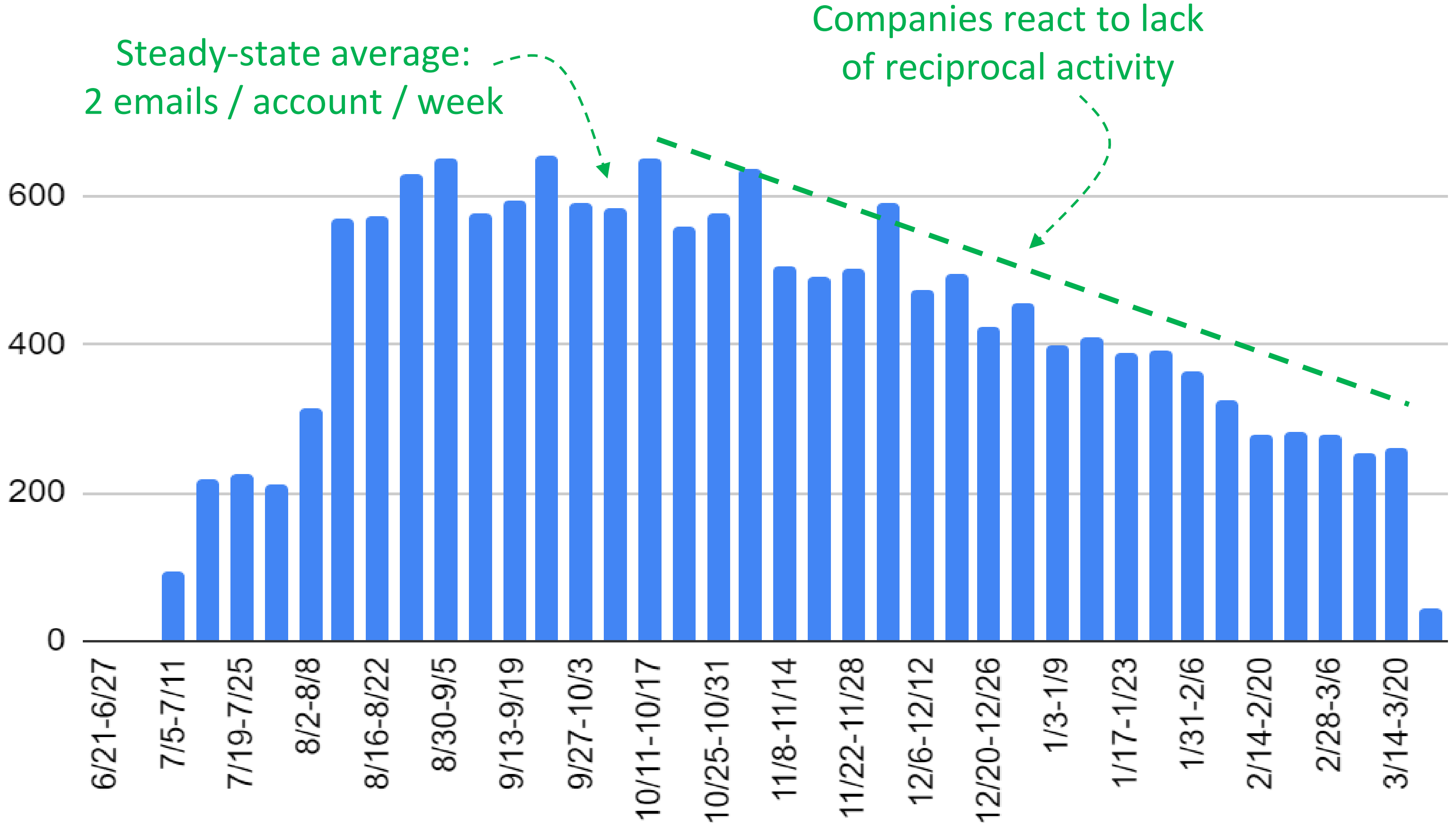
Experimental Setup

Received Data

Collection Results: Emails (1 of 2)

Received 16436 emails over 9 months

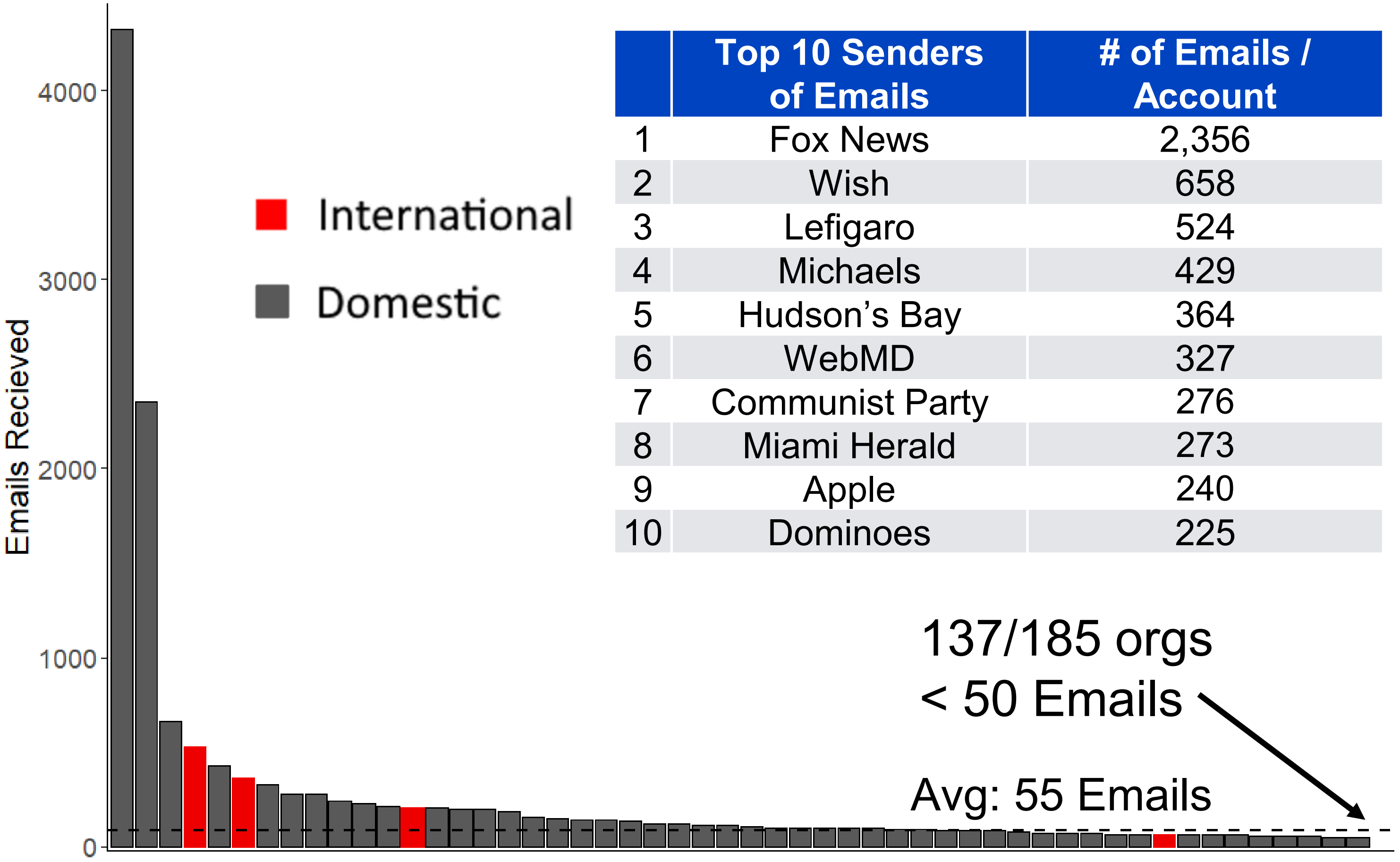
Emails received per week



Steady-state average: 2 emails / account / week

Companies react to lack of reciprocal activity

Emails by Organization



	Top 10 Senders of Emails	# of Emails / Account
1	Fox News	2,356
2	Wish	658
3	Lefigaro	524
4	Michaels	429
5	Hudson’s Bay	364
6	WebMD	327
7	Communist Party	276
8	Miami Herald	273
9	Apple	240
10	Dominoes	225

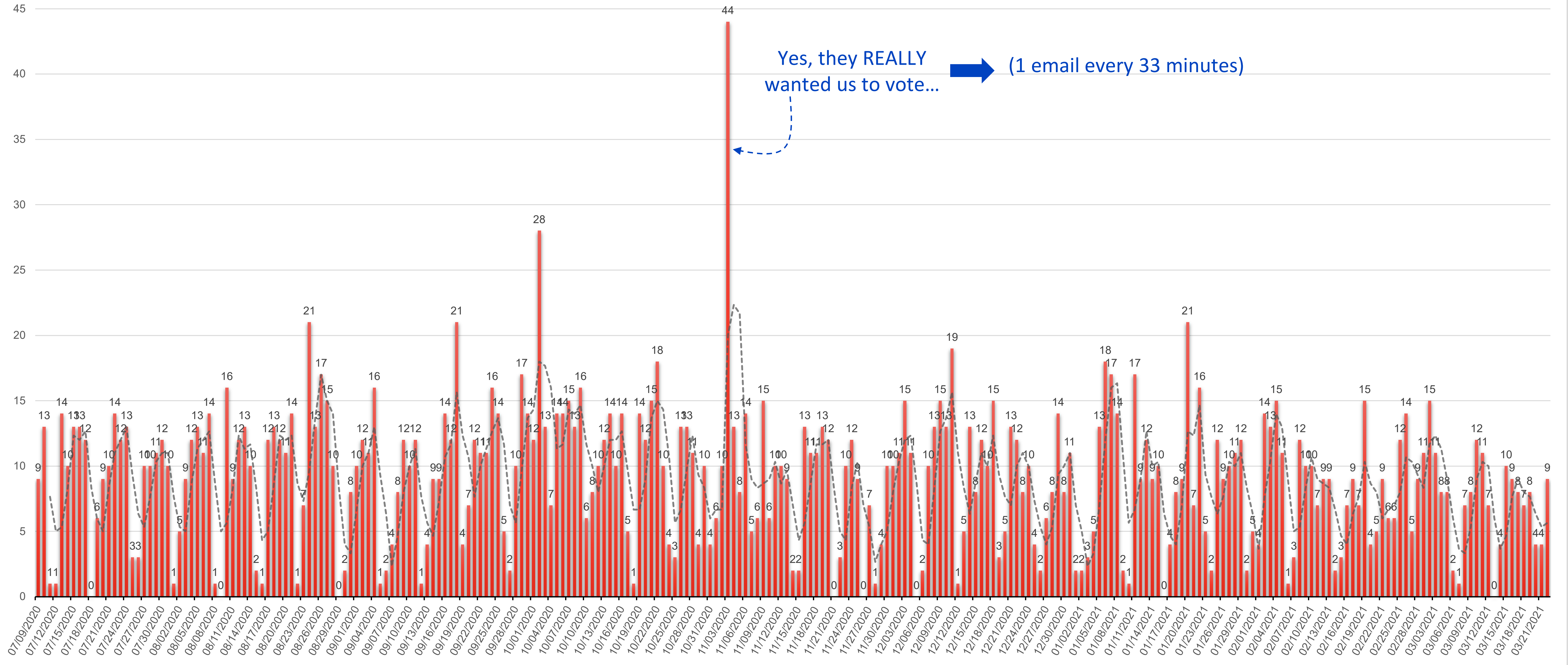
137/185 orgs < 50 Emails

Avg: 55 Emails

Collection Results: Emails (2 of 2)

Most prolific email source was Fox News with 2356 emails in < 9 months (~9 / day)

Daily Emails to ID_072 (Fox News)



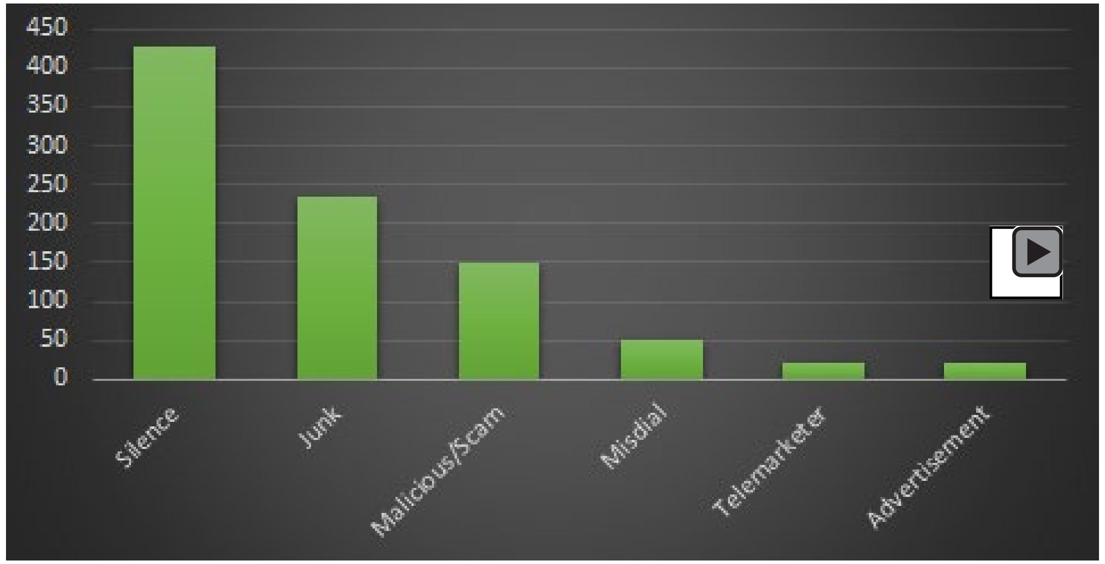
Yes, they REALLY wanted us to vote…

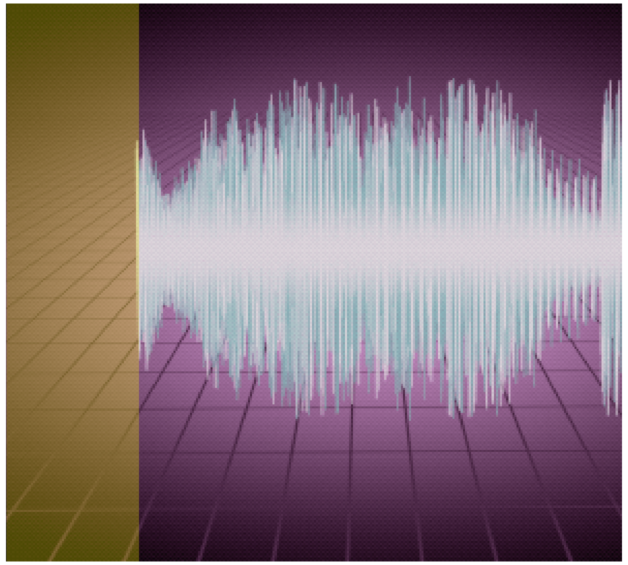
(1 email every 33 minutes)

Collection Results: Voicemails

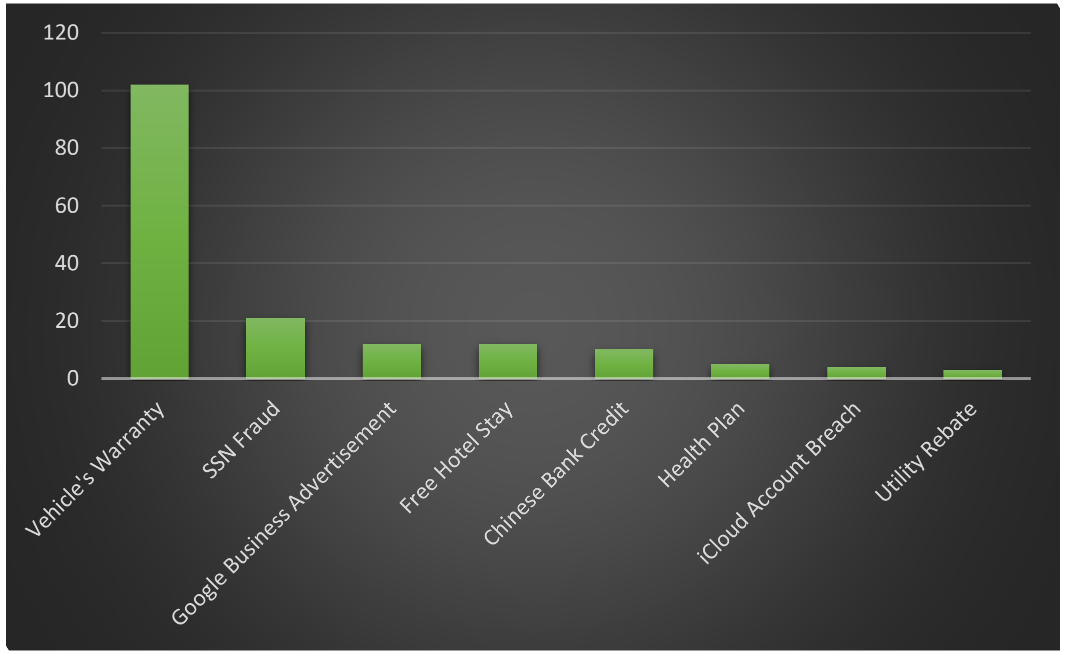
Received 3482 phonecalls, 949 voicemails over 9 months (150 phone lines)

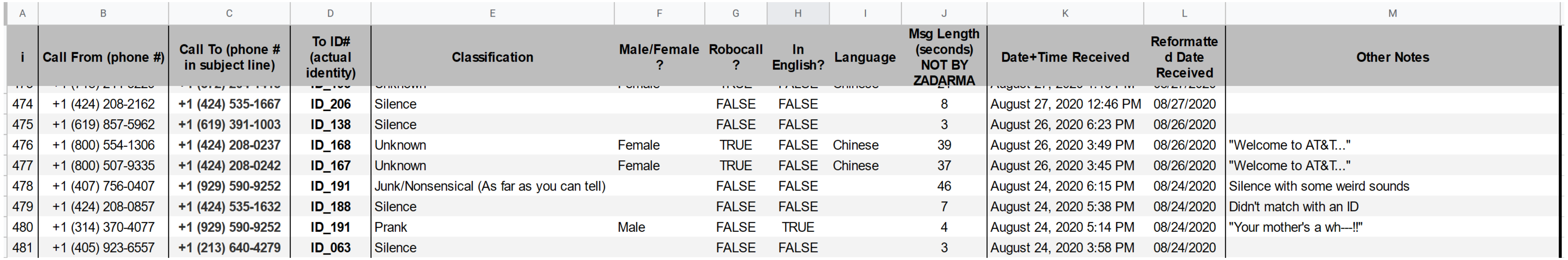
Types of Voicemail





Spam Call Breakdown

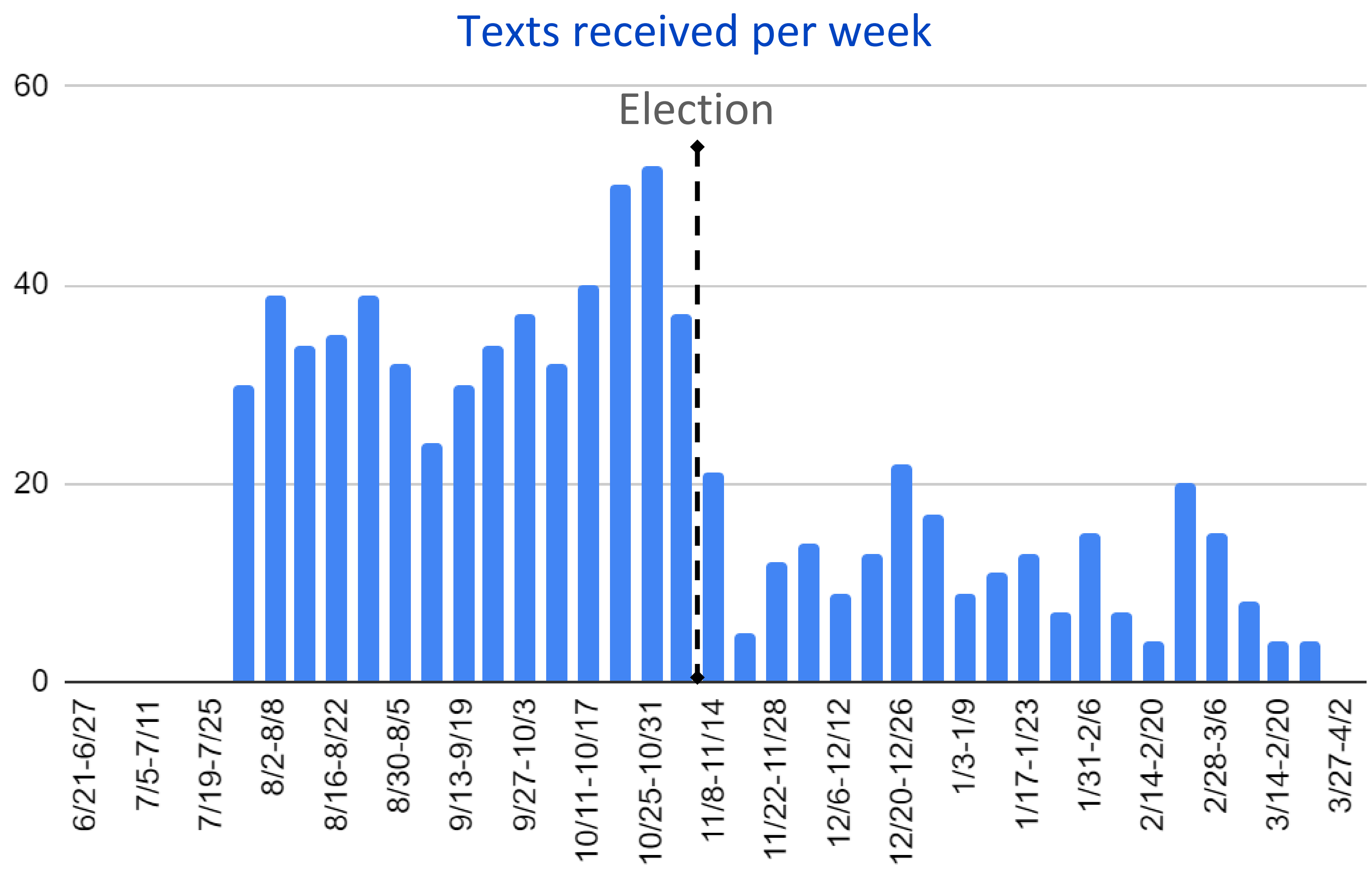


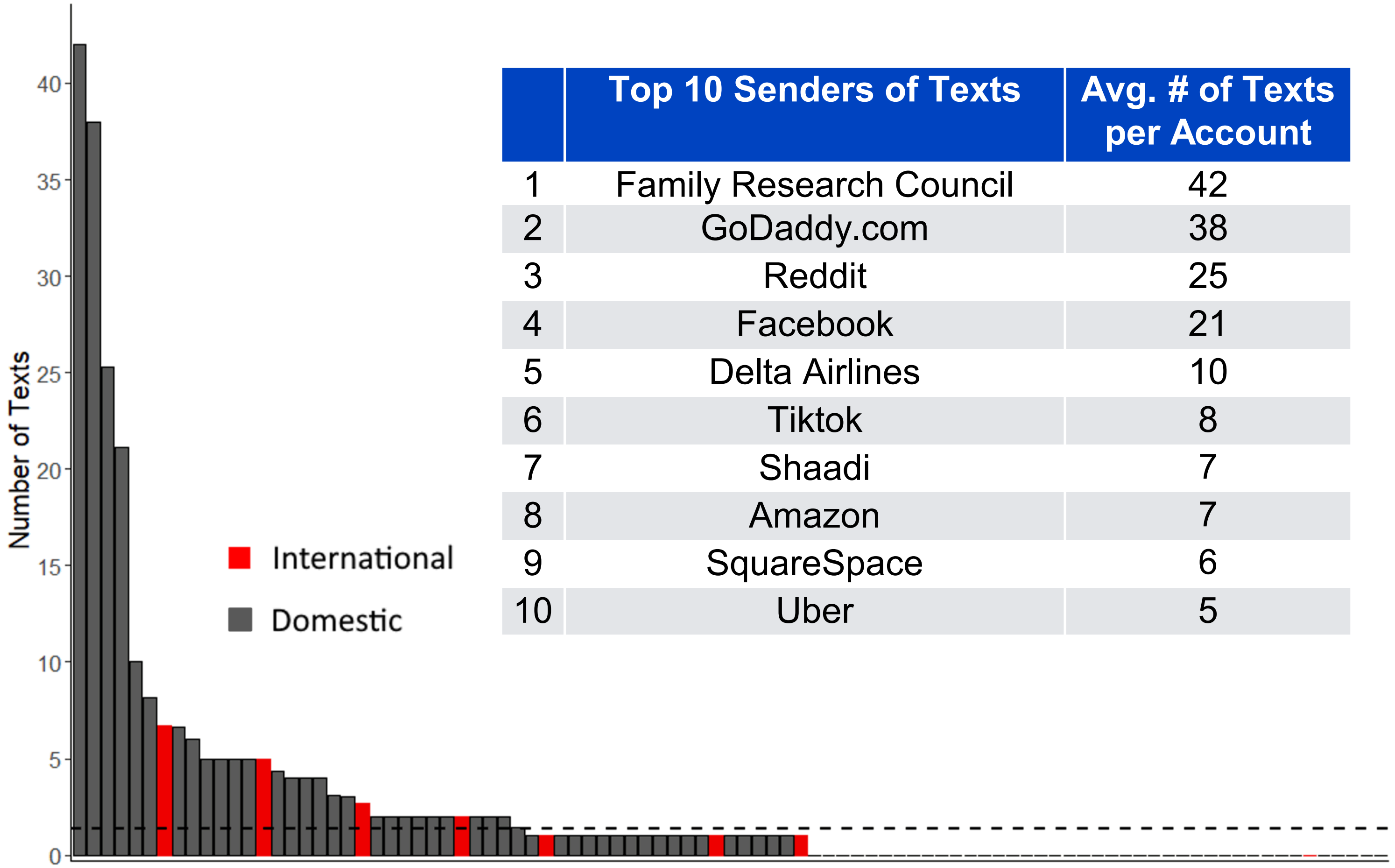


A	B	C	D	E	F	G	H	I	J	K	L	M
i	Call From (phone #)	Call To (phone # in subject line)	To ID# (actual identity)	Classification	Male/Female ?	Robocall ?	In English?	Language	Msg Length (seconds) NOT BY ZADARMA	Date+Time Received	Reformatted Date Received	Other Notes
474	+1 (424) 208-2162	**+1 (424) 535-1667**	**ID_206**	Silence		FALSE	FALSE		8	August 27, 2020 12:46 PM	08/27/2020	
475	+1 (619) 857-5962	**+1 (619) 391-1003**	**ID_138**	Silence		FALSE	FALSE		3	August 26, 2020 6:23 PM	08/26/2020	
476	+1 (800) 554-1306	**+1 (424) 208-0237**	**ID_168**	Unknown	Female	TRUE	FALSE	Chinese	39	August 26, 2020 3:49 PM	08/26/2020	"Welcome to AT&T..."
477	+1 (800) 507-9335	**+1 (424) 208-0242**	**ID_167**	Unknown	Female	TRUE	FALSE	Chinese	37	August 26, 2020 3:45 PM	08/26/2020	"Welcome to AT&T..."
478	+1 (407) 756-0407	**+1 (929) 590-9252**	**ID_191**	Junk/Nonsensical (As far as you can tell)		FALSE	FALSE		46	August 24, 2020 6:15 PM	08/24/2020	Silence with some weird sounds
479	+1 (424) 208-0857	**+1 (424) 535-1632**	**ID_188**	Silence		FALSE	FALSE		7	August 24, 2020 5:38 PM	08/24/2020	Didn't match with an ID
480	+1 (314) 370-4077	**+1 (929) 590-9252**	**ID_191**	Prank	Male	FALSE	TRUE		4	August 24, 2020 5:14 PM	08/24/2020	"Your mother's a wh---!!"
481	+1 (405) 923-6557	**+1 (213) 640-4279**	**ID_063**	Silence		FALSE	FALSE		3	August 24, 2020 3:58 PM	08/24/2020	

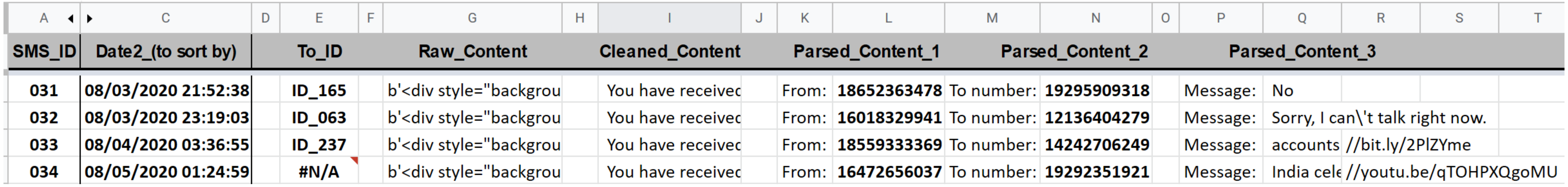
Collection Results: Texts

Received 774 texts over 9 months (150 phone lines)





	Top 10 Senders of Texts	Avg. # of Texts per Account
1	Family Research Council	42
2	GoDaddy.com	38
3	Reddit	25
4	Facebook	21
5	Delta Airlines	10
6	Tiktok	8
7	Shaadi	7
8	Amazon	7
9	SquareSpace	6
10	Uber	5



A ◂ ▸ C D E F G H I J K L M N O P Q R S T

SMS_ID	Date2_(to sort by)	To_ID	Raw_Content	Cleaned_Content	Parsed_Content_1		Parsed_Content_2		Parsed_Content_3	
031	08/03/2020 21:52:38	ID_165	b'<div style="backgrou	You have receivec	From:	18652363478	To number:	19295909318	Message:	No
032	08/03/2020 23:19:03	ID_063	b'<div style="backgrou	You have receivec	From:	16018329941	To number:	12136404279	Message:	Sorry, I can\'t talk right now.
033	08/04/2020 03:36:55	ID_237	b'<div style="backgrou	You have receivec	From:	18559333369	To number:	14242706249	Message:	accounts //bit.ly/2PlZYme
034	08/05/2020 01:24:59	#N/A	b'<div style="backgrou	You have receivec	From:	16472656037	To number:	19292351921	Message:	India cele //youtu.be/qTOHPXQgoMU

All’s Fair in Love & PII

For the most part, companies hoard PII*: no evidence of sharing from 290/300 accounts

Communist Party USA → PEOPLE'S WORLD

Communist Party USA → unemployedsolidarity.org

ID_054

B STOCK ↓ Etsy

ID_082

JD

Finish Line

ID_297/299

aaron brothers

CUSTOM FRAMING AT Michaels

ID_111

Cookie history shows limited connection to information selling / harvesting

ID_052 ↔ TRUMP VICTORY

campaigns.rnchq.com
victory.d.trump.com

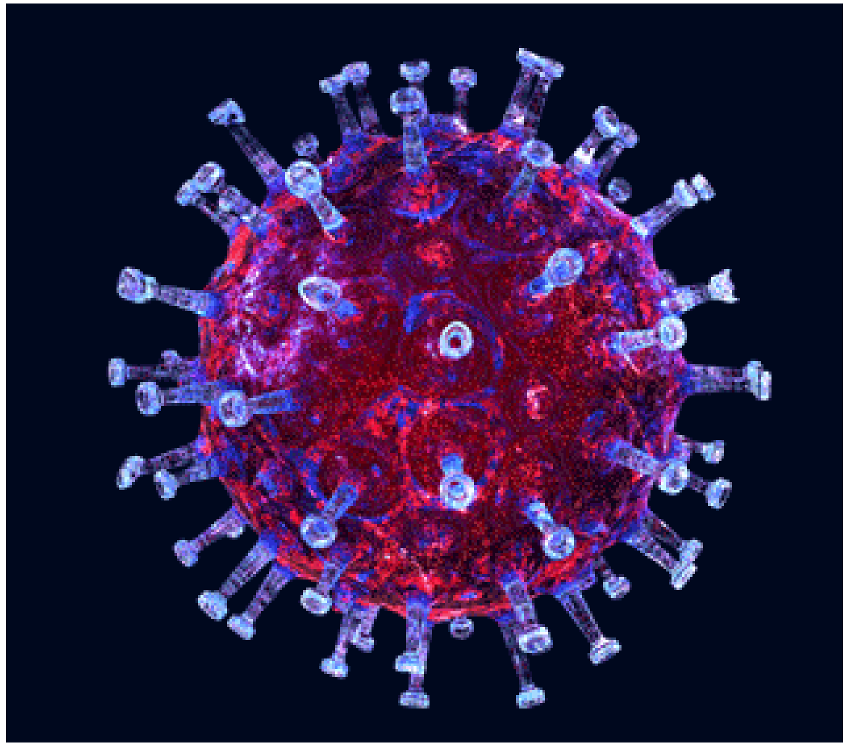
TikTok

ID_058 ↔ DCCC dccc.org

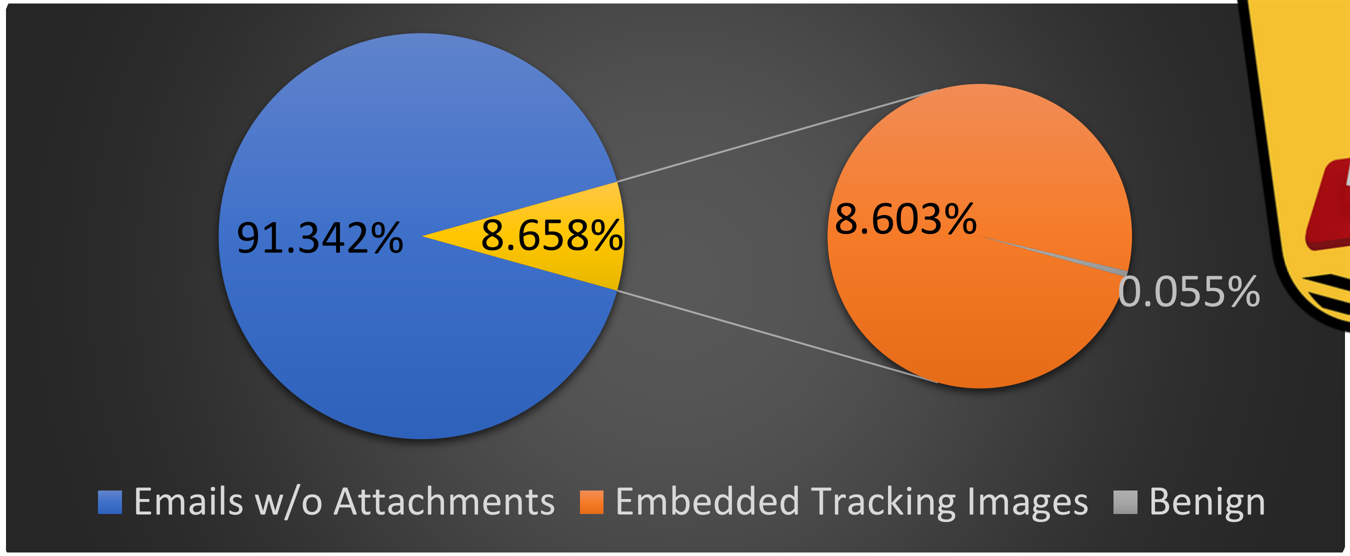
UNITE THE COUNTRY

*ID_156 and ID_179 excised due to erroneous signup process

Malicious Content



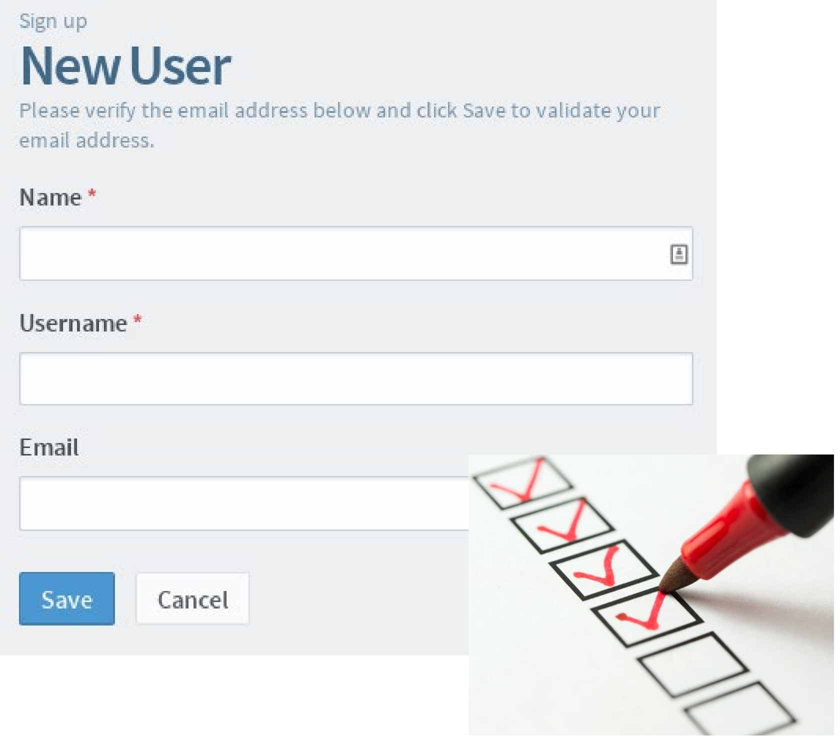
- Number of emails containing attachments: 1423
- Number of attachments with malicious attachments: 0



Types of Attachments

Lessons Learned

Account signup, validation, and maintenance processes



Sign up

New User

Please verify the email address below and click Save to validate your email address.

Name *

Username *

Email

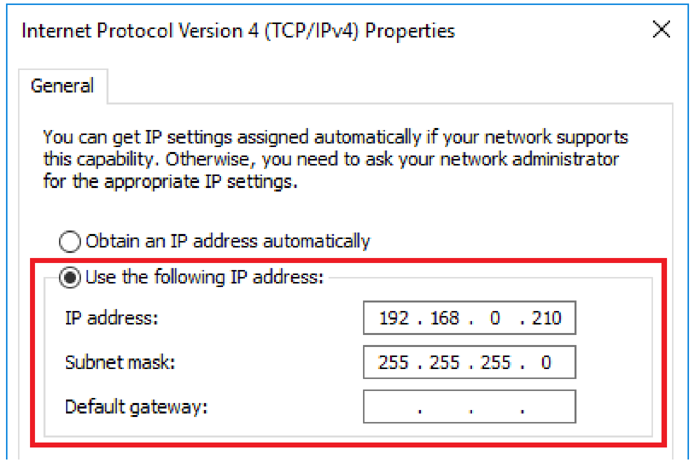
Save | Cancel

Rejected two virtual phone numbers

Only accepts local (Chinese) phone numbers

Limitations of university servers and infrastructure

Static IP



Internet Protocol Version 4 (TCP/IPv4) Properties

General

You can get IP settings assigned automatically if your network supports this capability. Otherwise, you need to ask your network administrator for the appropriate IP settings.

○ Obtain an IP address automatically

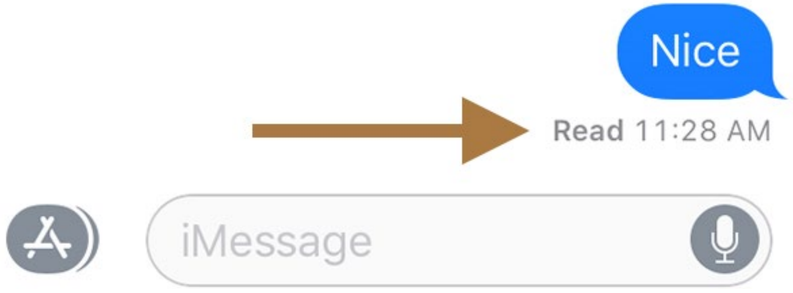
● Use the following IP address:

IP address: 192 . 168 . 0 . 210

Subnet mask: 255 . 255 . 255 . 0

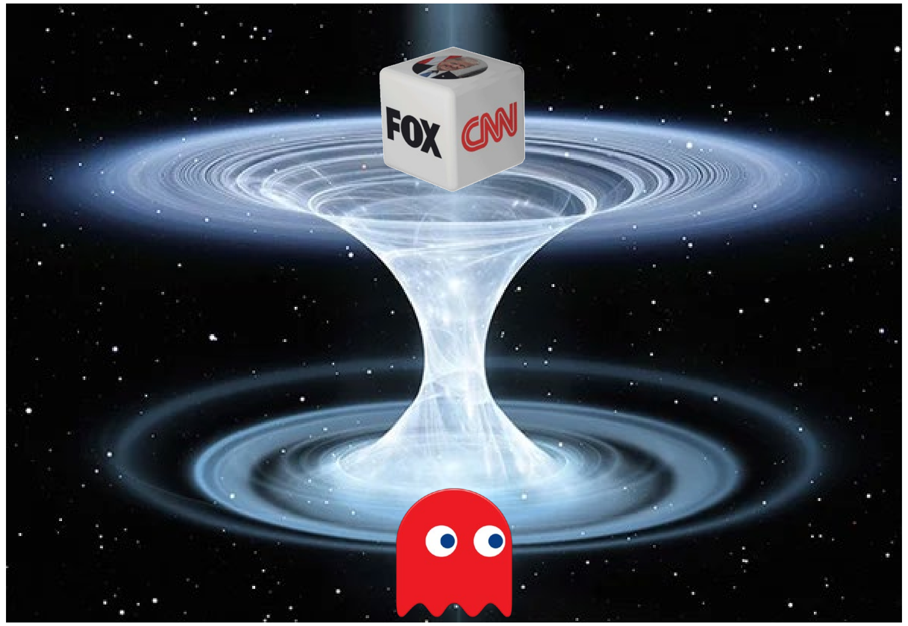
Default gateway: . . .

Lack of recipient activity (e.g., read receipts, social media entropy) is clear indicator of ghost account



Nice

Read 11:28 AM



Beyond Raw Numbers...

Practical Considerations:

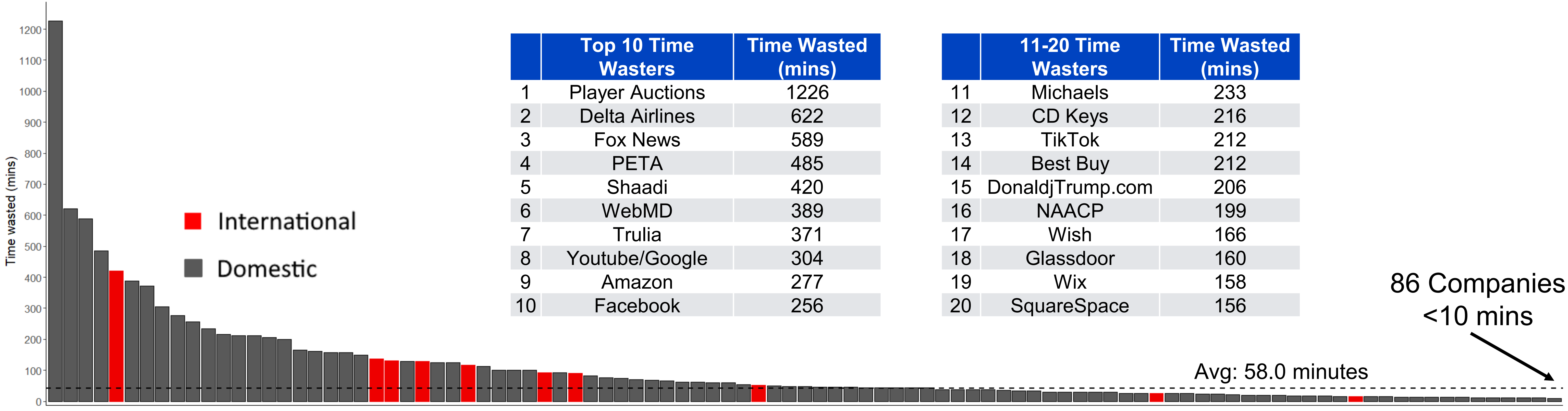
- What's the impact?
- Is there a political divide?
- Sharing by news sources?
- Foreign interest?
- Industry behaviors?
- Sharing as function of:
  - ~~Gender~~
  - ~~Race~~
  - ~~Location / address~~
- How relevant is the content?
- Does the organization stop when I ask them to?
- Can we predict how much an organization will send us?

Quantifying the Impact

Death by ~~1000~~ 1226 papercuts

Let’s estimate the amount of time wasted over 9 months if we actually paid attention to the glut of information sent to us from a one-time interaction.

- Voicemails = 5 minute distraction
- Texts = 1 minute distraction
- Email = 15 second distraction



86 Companies <10 mins

	Top 10 Time Wasters	Time Wasted (mins)
1	Player Auctions	1226
2	Delta Airlines	622
3	Fox News	589
4	PETA	485
5	Shaadi	420
6	WebMD	389
7	Trulia	371
8	Youtube/Google	304
9	Amazon	277
10	Facebook	256

	11-20 Time Wasters	Time Wasted (mins)
11	Michaels	233
12	CD Keys	216
13	TikTok	212
14	Best Buy	212
15	DonaldjTrump.com	206
16	NAACP	199
17	Wish	166
18	Glassdoor	160
19	Wix	158
20	SquareSpace	156

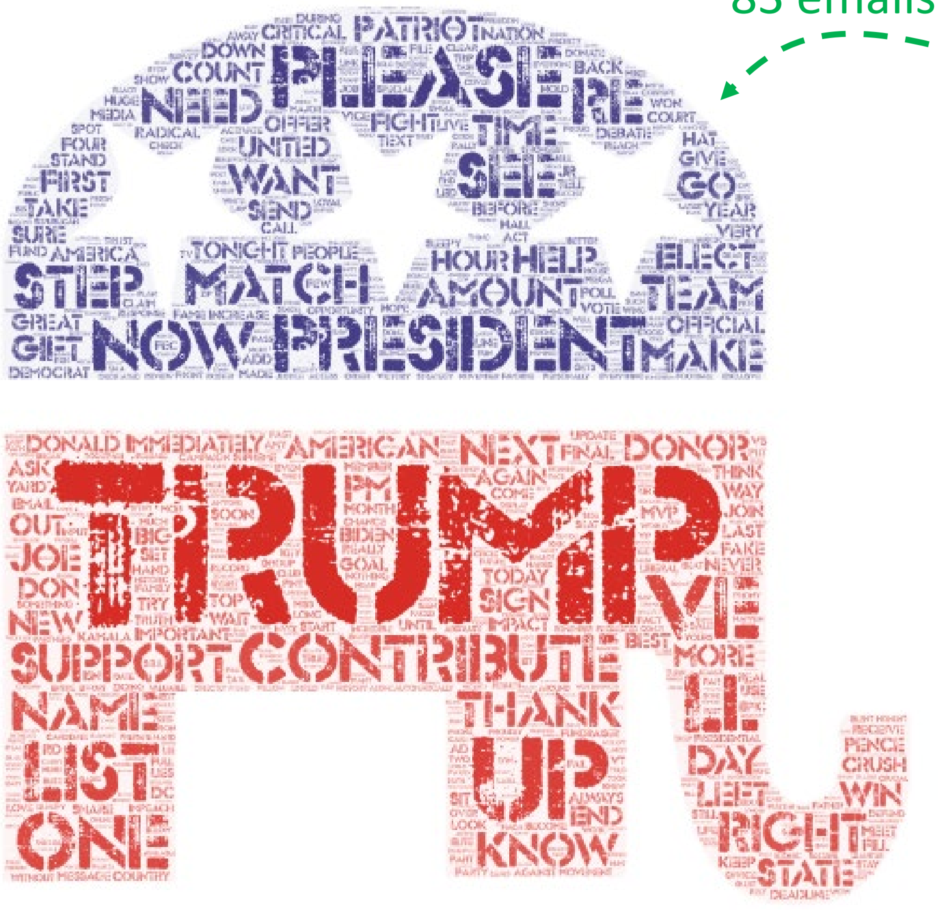
Measuring Politics: Quantity and Content

Total emails: 647
Total voicemails: 49
Total SMS: 47

Total emails: 357
Total voicemails: 42
Total SMS: 3

83 emails

94 emails

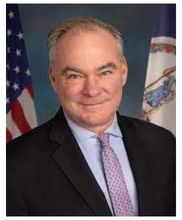


www.gop.gov

www.donaldjtrump.com

www.joebiden.com

www.democrats.org



Family Research Council

Mitch McConnell

Tim Kaine

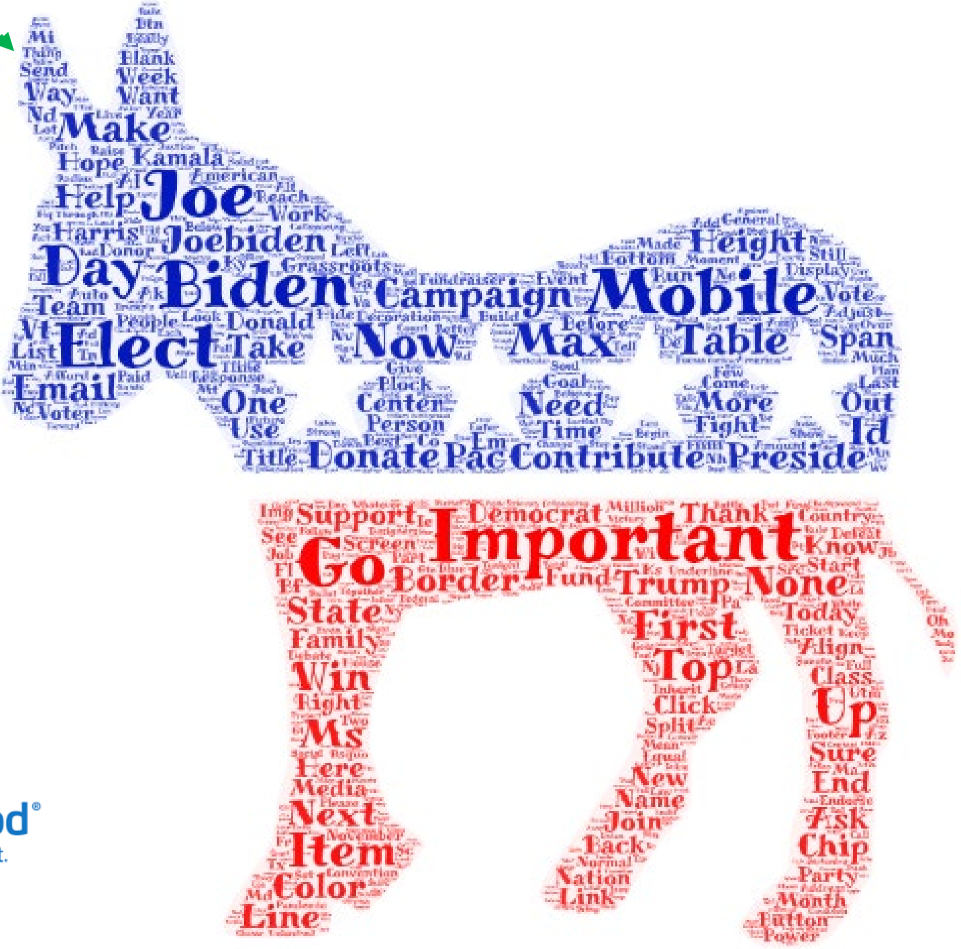
DCCC.org

Pro-Life Action League

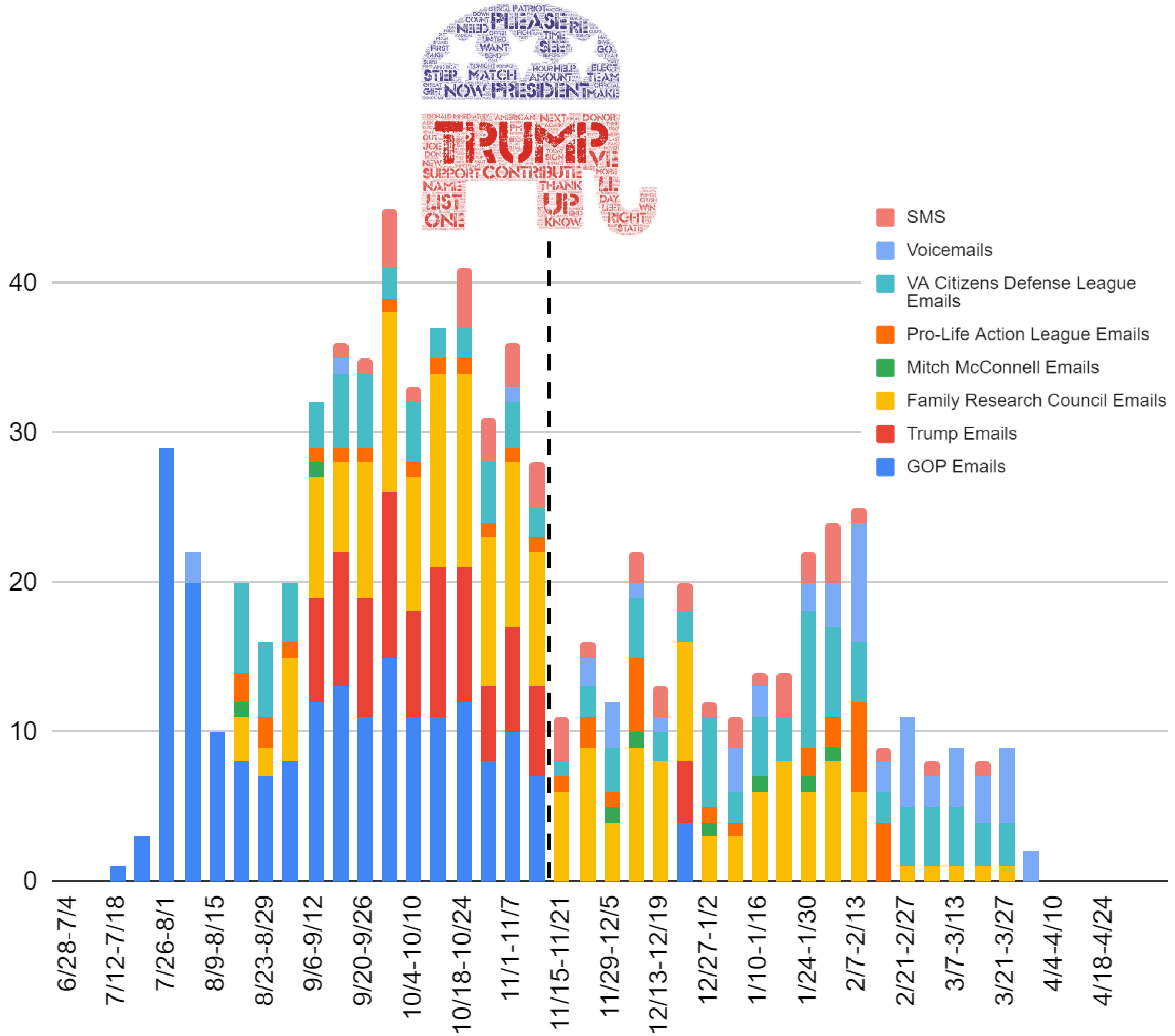
VA Citizens Defense League

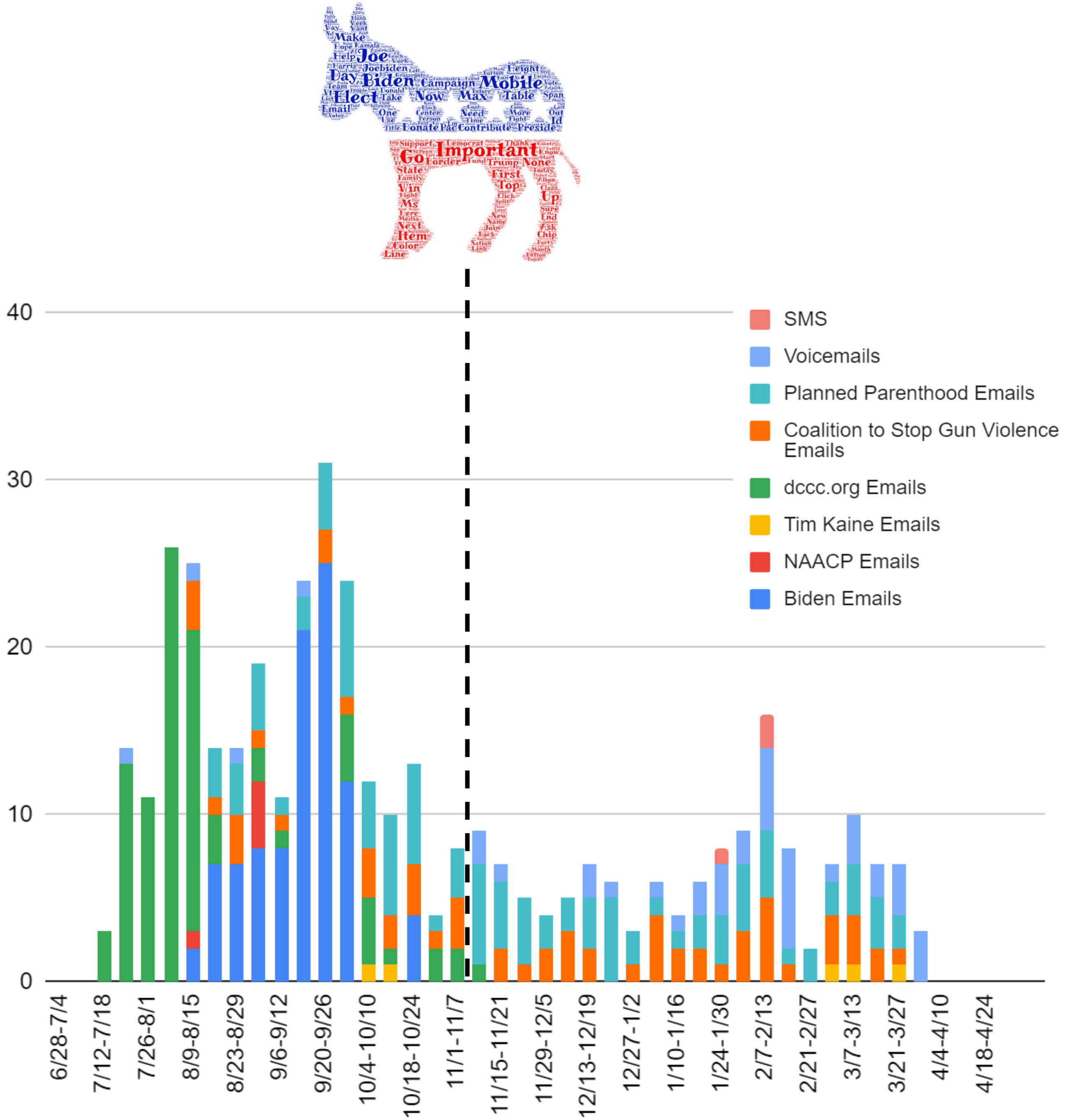
Coalition to Stop Gun Violence

Planned Parenthood

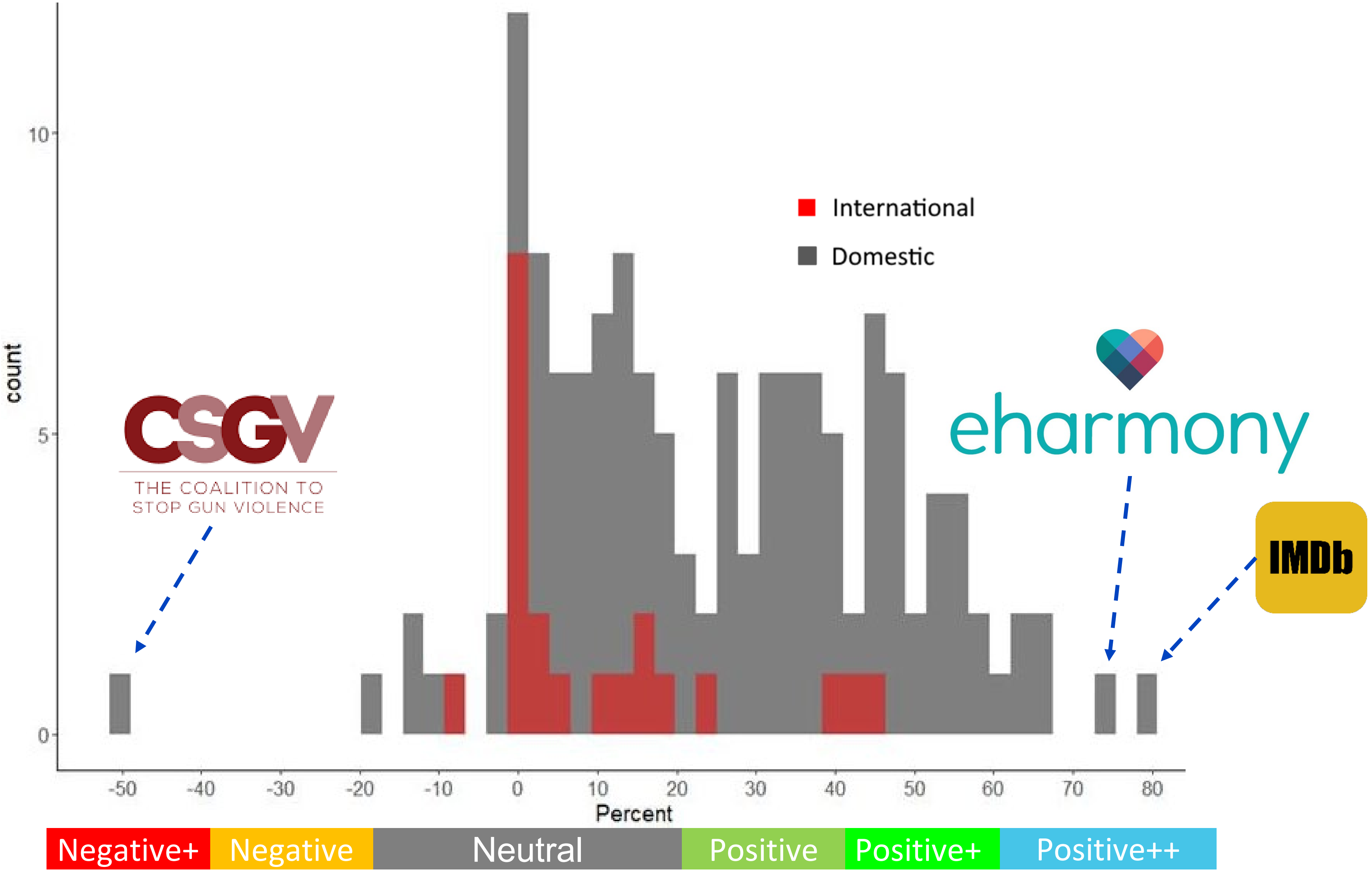


Political Volume





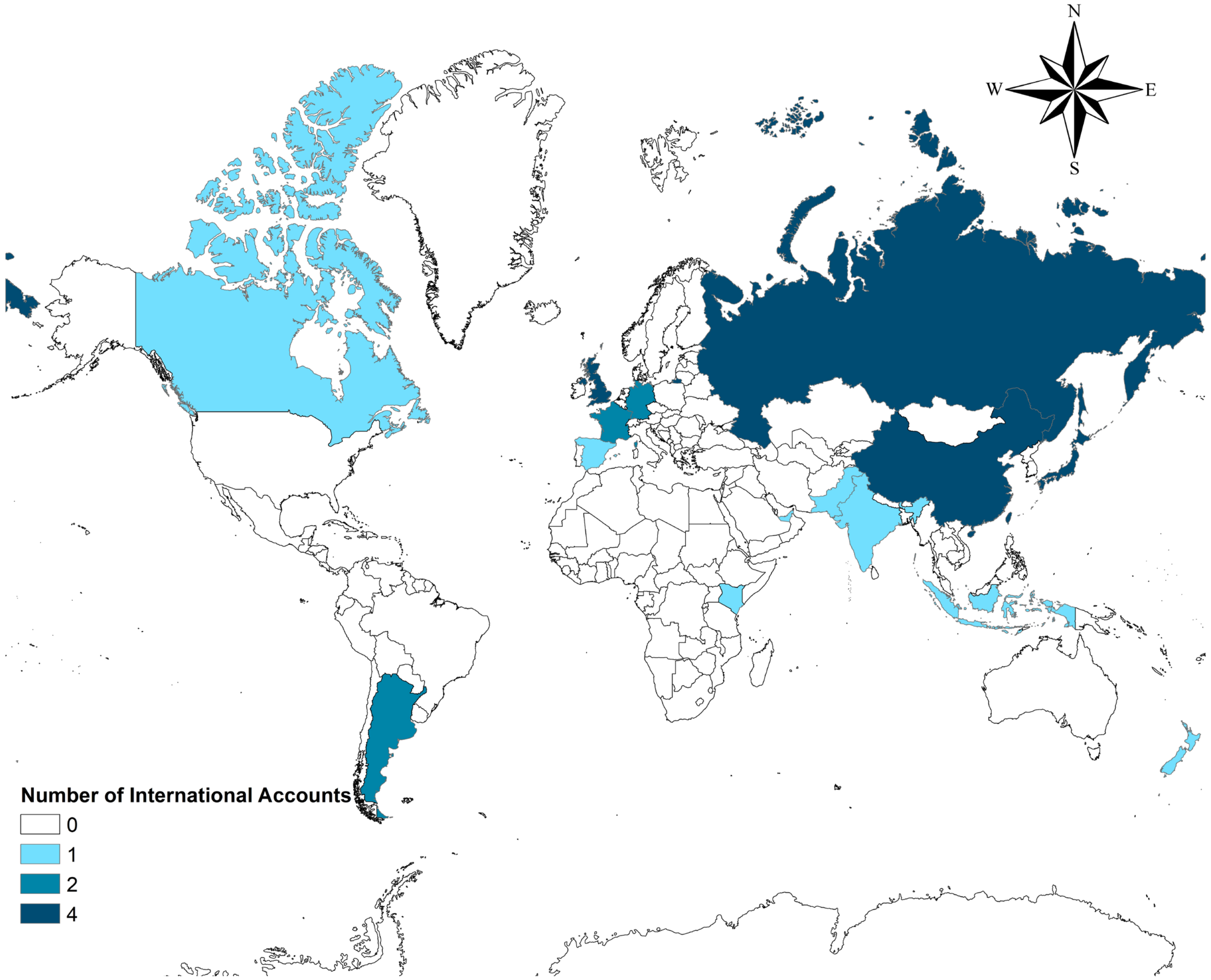
Sentiment Analysis

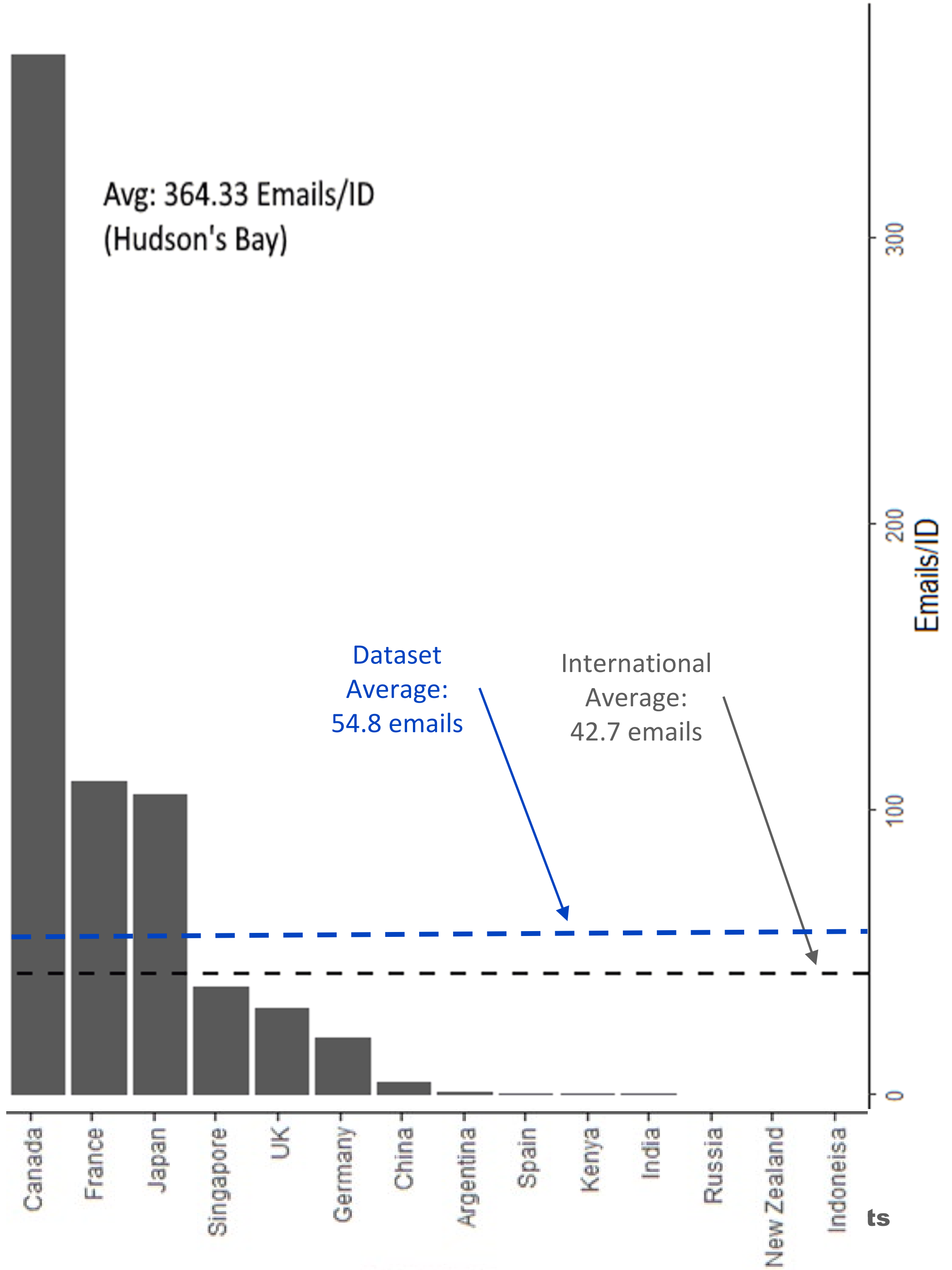


Analysis based on: Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

Foreign Interest

Country	Company
Argentina	Taringa
Argentina	Twoo (Formarly Sonico)
Canada	Hudson's Bay
China	Alibaba
China	Douban
China	Toutiao
China	Zhanqi
France	20 Minutos
France	Leboncoin
France	Lefigaro
France	Ouest France
Germany	VZ
Germany	XING
India	Shaadi
Indonesia	Tokopedia
Japan	Cookpad
Japan	Hatena
Japan	Rakuten
Kenya	PesaPal
New Zealand	Stuff
Pakistan	Millat Facebook
Russia	RuTube
Russia	Sputnik (World News)
Russia	Yandex
Russia	Yandex Disk
Spain	20 Minutos
Singapore	KrisShop
UK	Asos
UK	Badoo
UK	Discovery Store
UK	JD Sports
United Arab Emirates	Goalzzz (KOOORA)




Avg: 364.33 Emails/ID (Hudson's Bay)

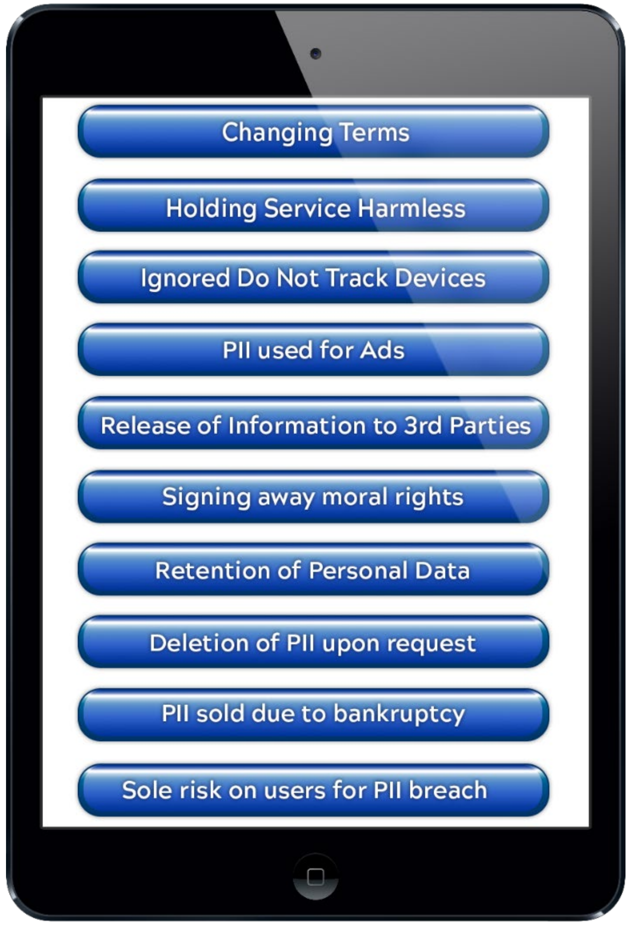
Dataset Average: 54.8 emails

International Average: 42.7 emails

Can we predict the level of *Use* & *Abuse*?

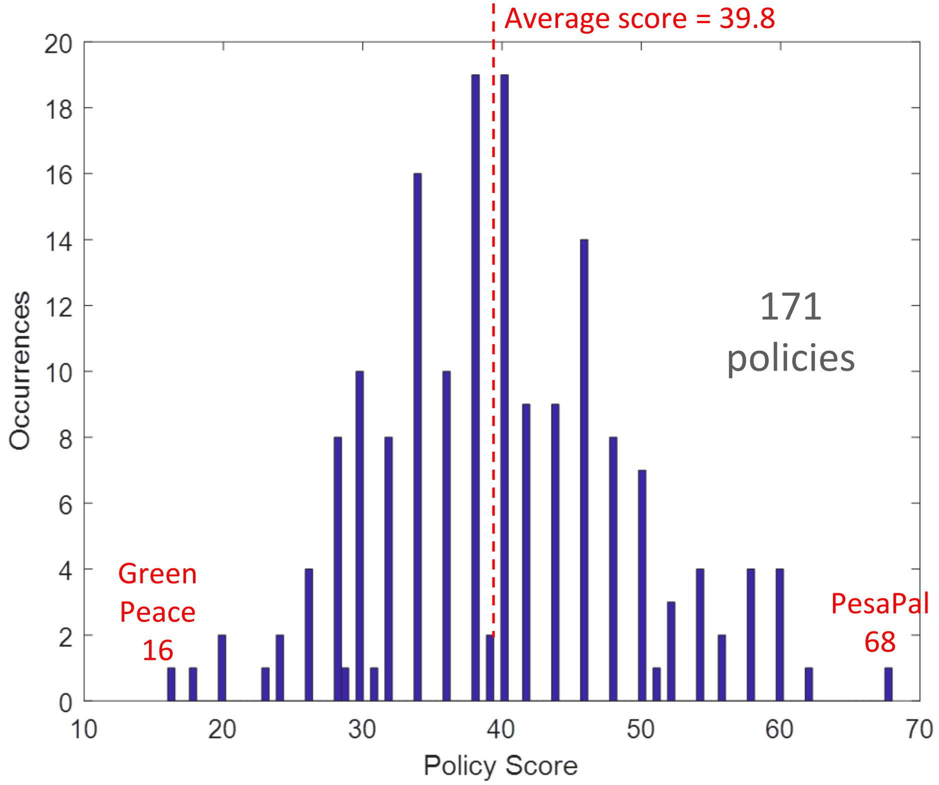
171 policies

Policy Scoring Rubric



- Changing Terms
- Holding Service Harmless
- Ignored Do Not Track Devices
- PII used for Ads
- Release of Information to 3rd Parties
- Signing away moral rights
- Retention of Personal Data
- Deletion of PII upon request
- PII sold due to bankruptcy
- Sole risk on users for PII breach

Scores: 0-100



	Top 10
1	PesaPal
2	ACM
3	RetailMeNot
3	Shelor Motor Mile
3	Mitch McConnell
3	Ebay
7	Safeway
7	RuTube
7	Roanoke Times
7	Ouest France

	Bottom 10
171	Green Peace
170	ACLU
168	DCCC.Org
168	Delta Airlines
167	Yelp
165	GOP.gov
165	US News
161	Badoo
161	Cars.com
161	Chicago Tribune

"What's a cookie? In real life, it's a delicious baked good. You can find lots of delicious cookie recipes on Pinterest." (Pinterest)

"Should we happen to get acquired (again!) or go out of business (no way), the transfer of assets from us to our buyer may very well include user information." (Tumblr)

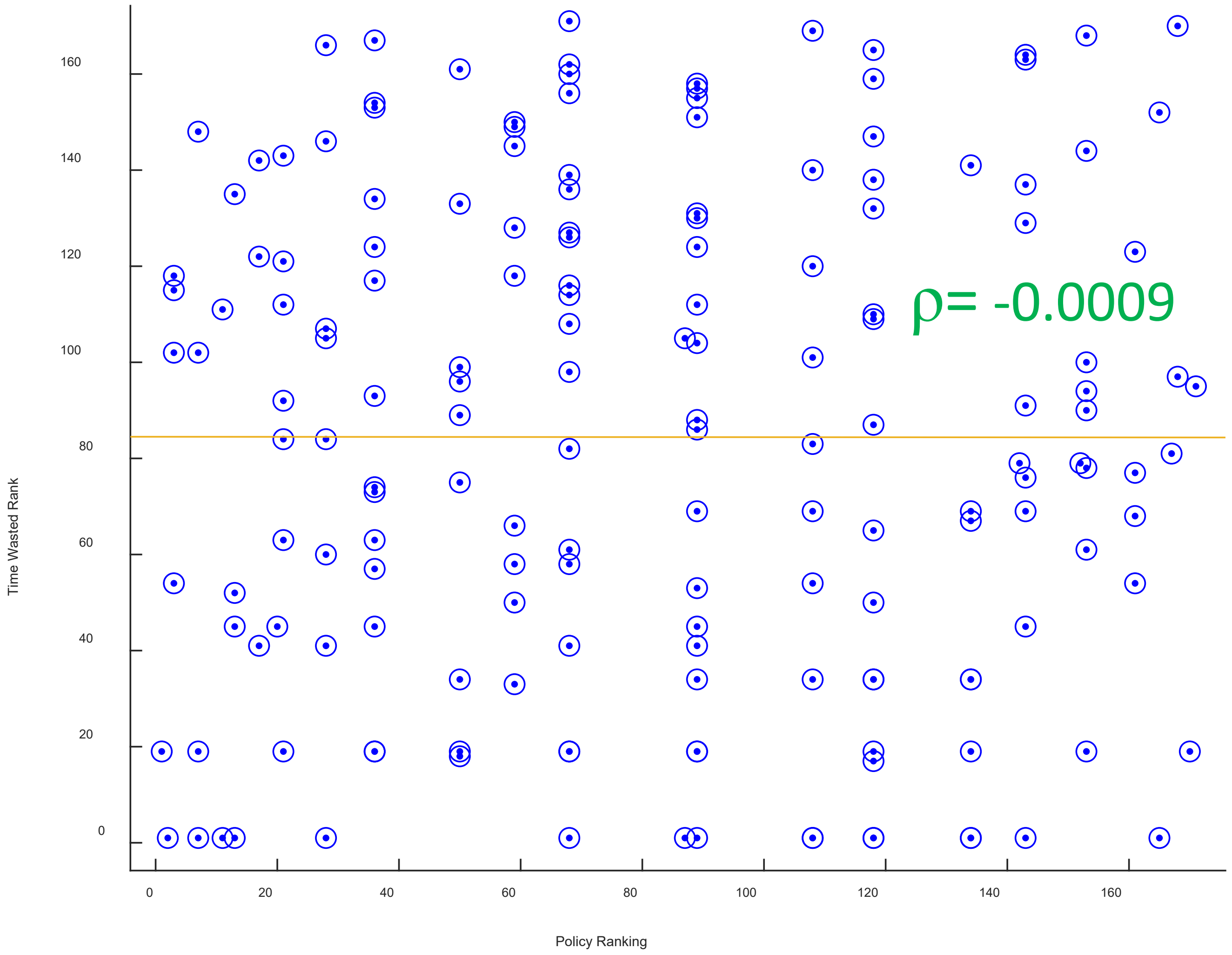
"If you do not understand the terms in this section or elsewhere in this agreement, please consult a lawyer for clarification before accessing or using the services." (TripAdvisor) [11th/172 highest grade level policy / 23rd longest]

Policy Rubric

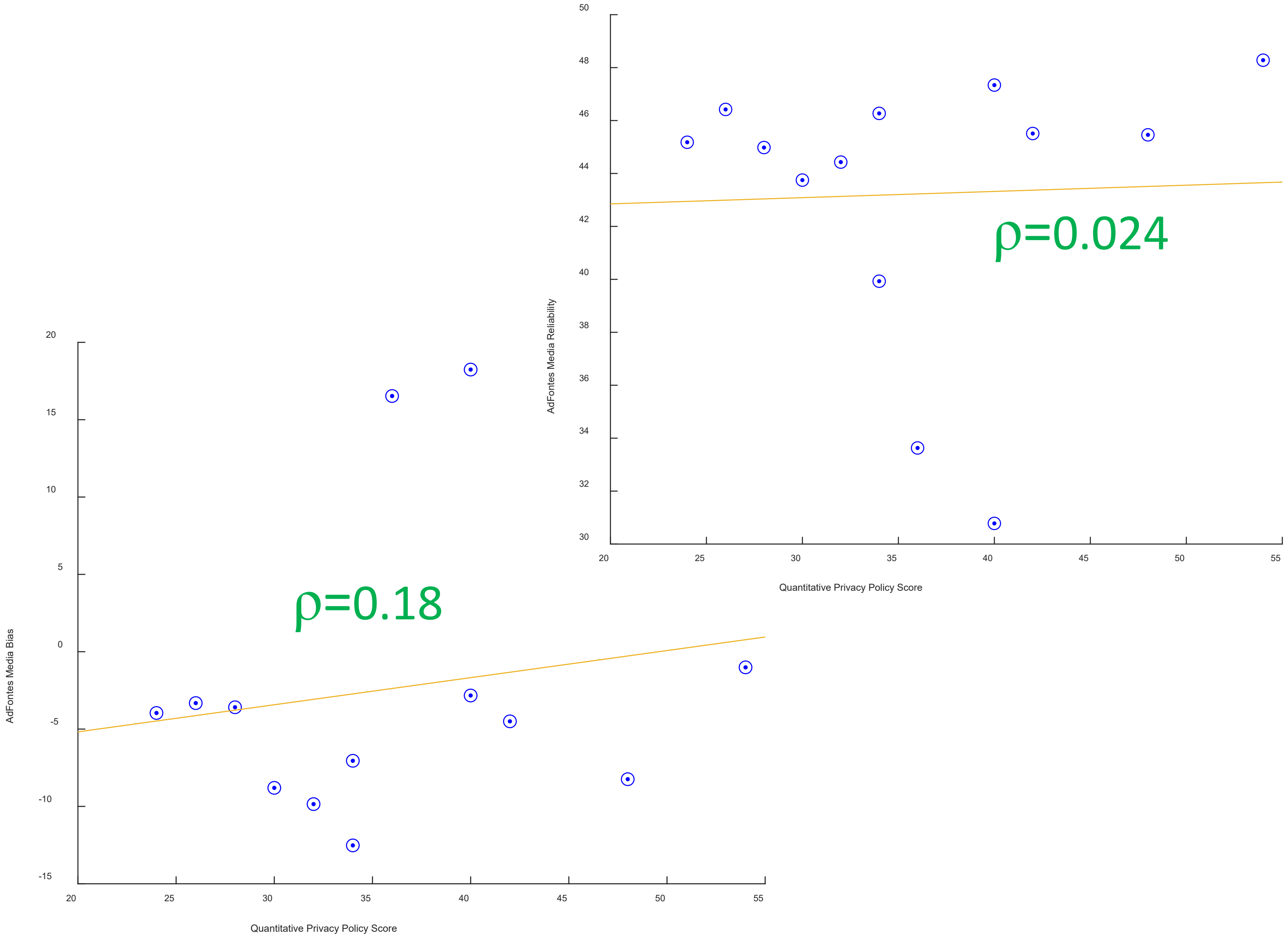
Category	Changing Terms	Holding Service Harmless	Ignores Do Not Track (DNT) Devices	Personal Identifiable Information (PII) used for Ads	Release of information to third parties	Signing away moral rights	Retention of Personal Data	Deletion of PII upon request	Information being sold due to Bankruptcy	Puts sole risk on users for liabilities
1	Change privacy policy with changes applicable retroactively	User must defend the service against any claims/costs/liabilities if any lawsuit arises	Does not acknowledge or mention DNT signals	The service internally collects any available information of the user to sell and/or create targeted ads	The service consistently sells/distributes PII to associated third parties for any purpose including undefined "business purposes"	A complete dismissal of these rights and liability of suit when the user agrees to a particular privacy policy	Full retention of all data indefinitely after a user deactivates their account	The service does not offer such a feature or continues to retain information despite a request from the user	The company/service will sell and contribute all stored customer data as a result of being bought out or merging with another company	Puts total risk on the user for any liabilities, and the aforementioned service is not held accountable
2	Change privacy policy without notification, but changes are forward looking	User is responsible for defending the service in cases where the user violated the companies privacy policy	Complete recognition of these signals and denies the user the right to the website and/or continues to track the user without notification	Collects a significant amount of PII possibly including address, contacts, and any site activity. Does not collect all available PII but more than as specified in Category 3	The service collects and sends PII to third parties in order for them to sell advertisements or for defined "business purposes"	The user obtains some say over their content; however, the particular service maintains most of the control	Service holds information for as long as they deem necessary/after a predefined extended period of time longer than a year	User is unable to request or delete any information; however, the service will allow less information to be collected	User is notified of acquisition; however, no action can be taken by the user to limit data being transferred	User maintains soles risk on every aspect of site; however, service can be held liable to distribute cash compensation up to twenty dollars or in extreme cases
3	Claims to give notice, but provides vague distribution details	User is responsible for defending the service in cases where the user violated others rights/broke the law, not from policy violation	Acknowledges DNT signals and continues to track only due to lack of infrastructure to support these settings / lack of standard	Collects a 'normal' amount of PII including name, email address, log data, general location data as ascertained from IP address, etc.	The service only releases information to third parties if the user requests a service or more information from the initial website	Rights are waived; however, the privacy policy places some liability on the company and users maintain almost equal control	Service temporarily holds a reduced quantity of information or retains PII in case of potential reactivation	A user is able to request their information; however, they are unable to delete any information or request to delete is not honored	In merger or asset sales, data is sent to receiving company under the pretense of equivalent or improved privacy standards	The user and the service are mutually responsible. The service uses good faith to ensure data security and information accuracy. Will not claim responsibility on negligence
4	Clear notification of changes in privacy policy	User is responsible for defending the service in cases where the user violated others; however, service can remain accountable if the service played any role in the digression	Acknowledges DNT signals and complies; however, the service does not allow full access to all of the present features	Service provides a menu to disable all but necessary cookies, and collects a normal/less than normal amount of PII as defined above	The service releases PII only with previous consent from the user in order to show the user more relevant content	Waiving moral rights is optional; however, the service still has the final say over user content on the service	Information is stored after deletion of account only to comply with applicable regulations. A scheduled deletion is still in place with no intention of prolonged storage	A user is able to request to delete all their information; however, they may not be able to delete most of their information, only some	User is notified that there data is forfeit due to bankruptcy/merger; however, they may only be able to delete certain aspects of their PII. Some will be transferred over to the acquisition company	Data breaches caused by the user are not protected; however, if the service experiences a breach in their databases or in any other circumstance, the user is not held liable. User is protected on service negligence
5	User permitted to opt out of privacy policy changes/ allows for extensive copies of previous policies to ensure changes	Service assumes risk and takes liability away from the user if a lawsuit arises	Service complies with DNT signals and allows user access to the full features of the service	Minimal to no PII is collected or used for internal targeted products or services. The user still has access to the full features of the product	The service releases little to no information to third parties regardless of user consent and maintains internal consistency with the user's PII	The privacy policy states you are not required to sign away your moral rights	Either a user can delete all PII upon deactivation or request, or companies collect no user PII (in which case retention is impossible)	A user is able to request all their information and delete upon request with assurance from the service that the information will be rightly processed	Either all information is forfeit and not sold to the takeover company, or the user is notified and has an opportunity to delete their data before it is sold	The service is completely responsible for breaches on their end and/or not all risk is placed on the user

Correlations with Privacy Policies

No obvious correlations with media “bias”, “reliability”, policy scores, or “Lifetime Wasted”
- Small sample sizes



ρ= -0.0009

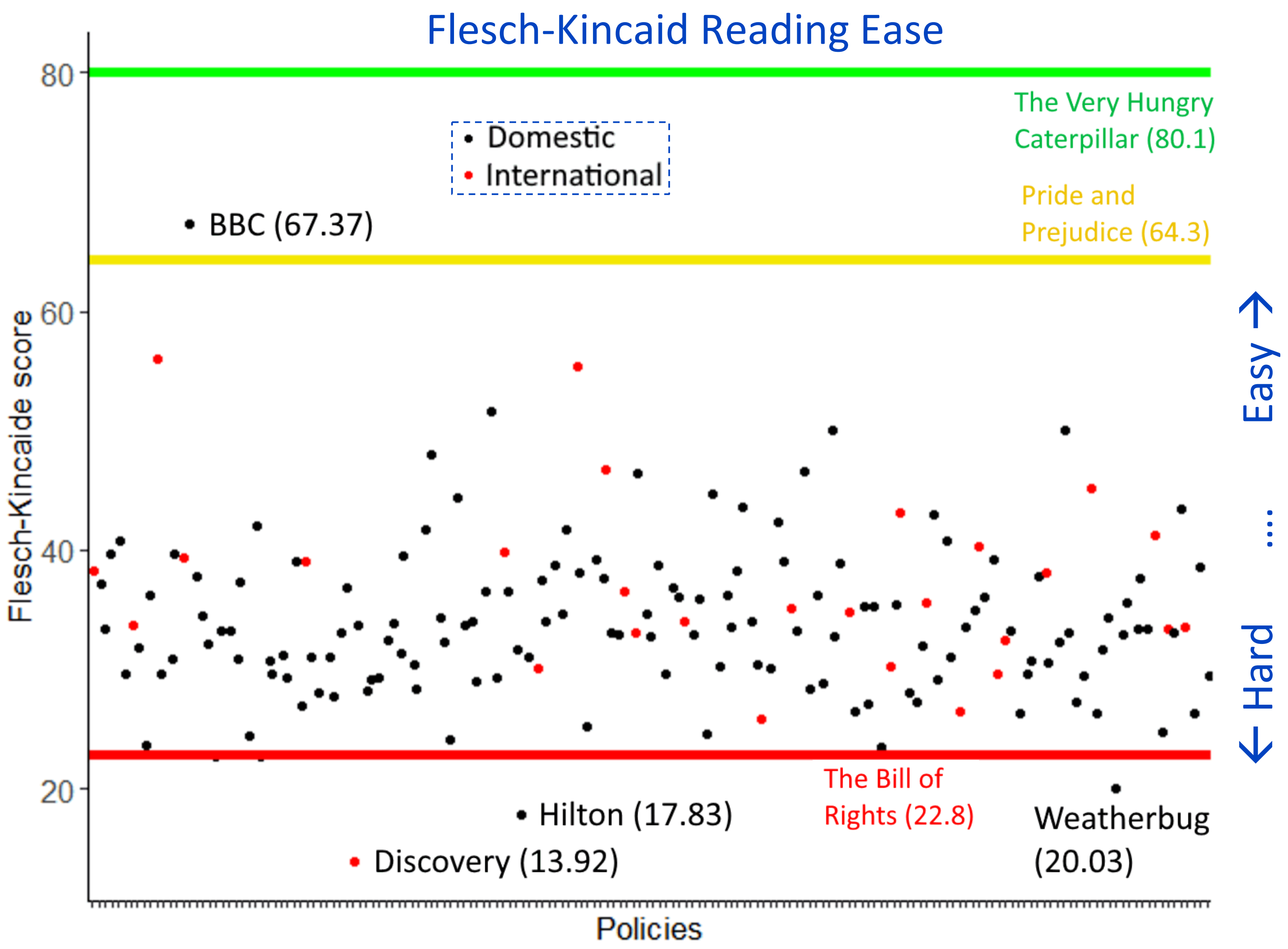


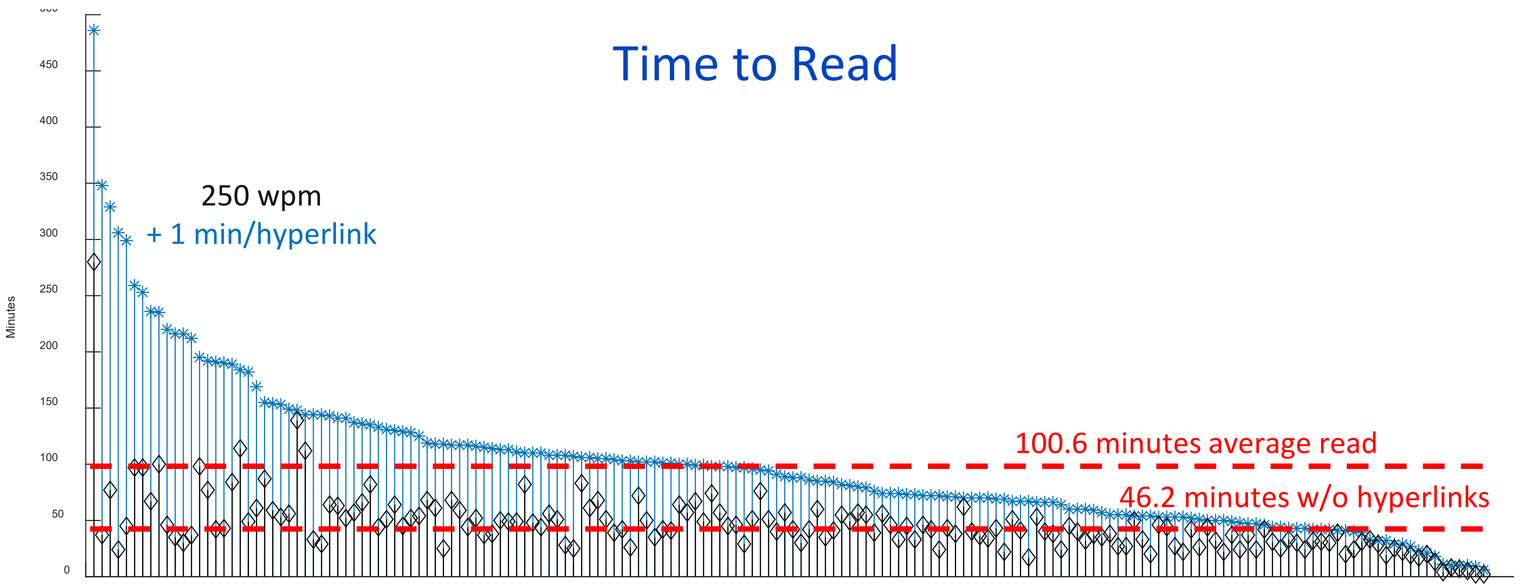
ρ=0.18

ρ=0.024

AdFontes Media *Bias* and *Reliability* metrics: https://www.adfontesmedia.com/

Policy Analysis: Quantitative

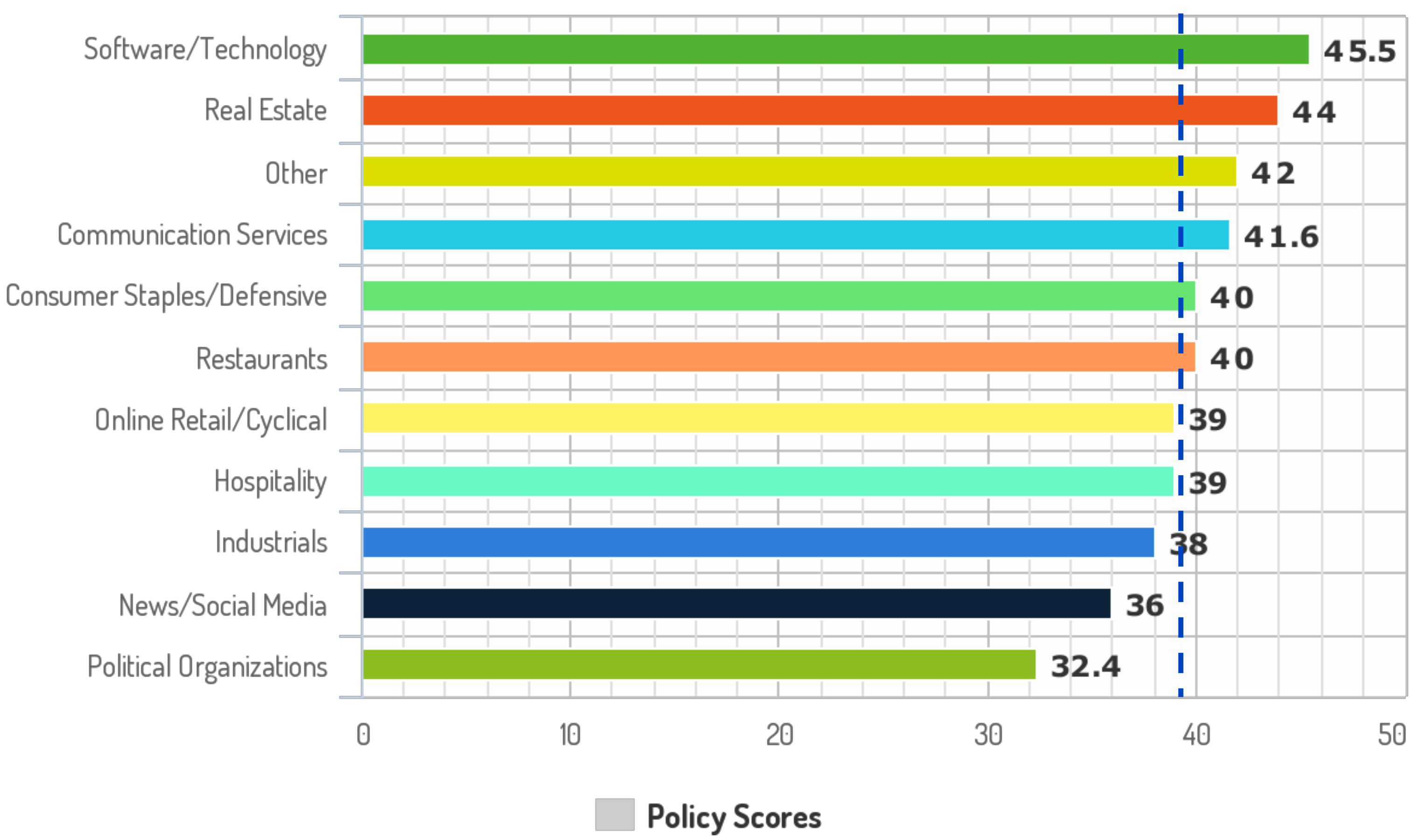




	Top 8 (Longest)		Bottom 8 (Shortest)
1	Indeed	172	Pro-Life Action League
2	OuestFrance	171	Mitch McConnell
3	XING	170	GreenPeace
4	Microsoft	169	Family Research Council
5	Twitter	168	PesaPal
6	Safeway	167	GOP
7	RetailMeNot	166	NAACP
8	JD-Sports	165	VA Citizen's Def League

Policy Analysis

Average Policy Score by Industry



Political Policy Scores

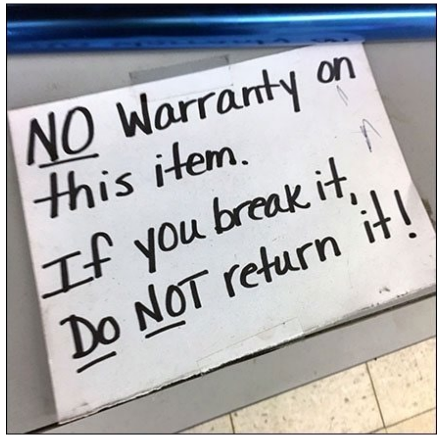
		Top 7		Bottom 7	
60	1	Mitch McConnell	16	Green Peace	16
50	2	Planned Parenthood	15	ACLU	18
40	3	UNICEF	14	Dccc.org	20
38	4	Donald J Trump	13	GOP.gov	24
38	4	NAACP	12	Pro-Life Action League	26
38	4	VA Citizens Defense League	10	PETA	28
34	7	Joe Biden	10	Family Research Council	28

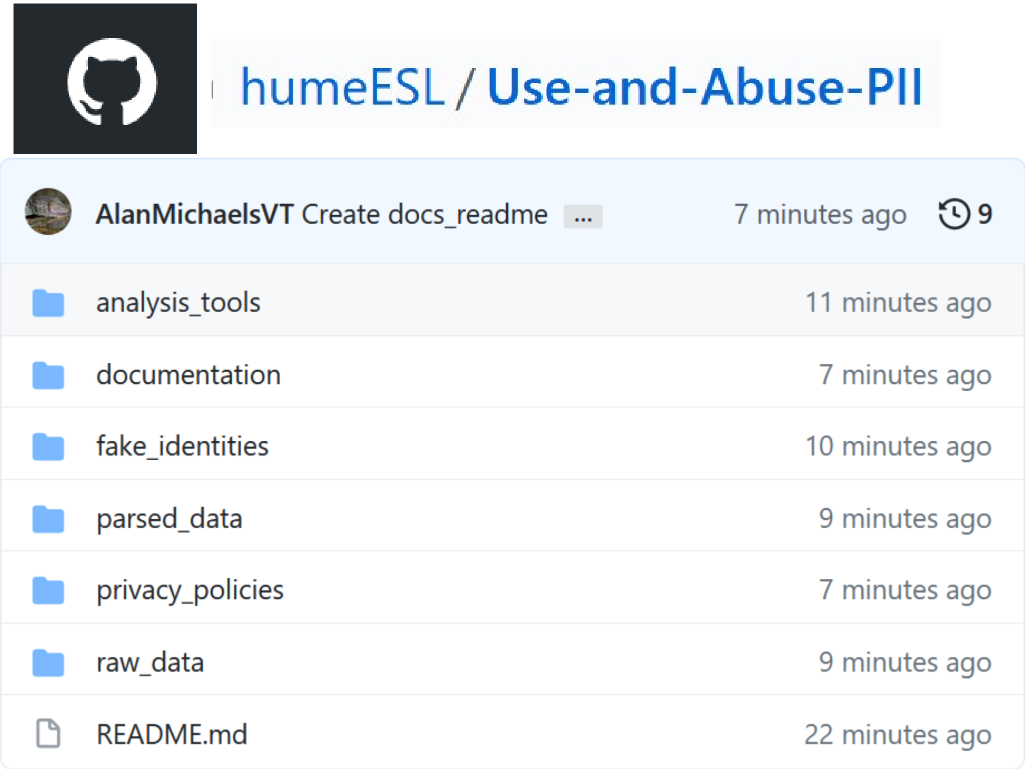
Open Source Dataset

Whitepaper posted to Blackhat website

Dataset offered to other researchers:

- Raw data*:
  - 16436 emails
  - 3482 phonecalls
  - 949 voicemails
  - 774 texts
  - *purchase data and private emails used for testing excised
- Information for 300 fake identities
  - Assignments to 185 entities, manufactured biases
  - Passwords omitted; all accounts disabled
- 171 privacy policies with scoring rubrics
- Various scripts and tools for automating analysis





humeESL / Use-and-Abuse-PII

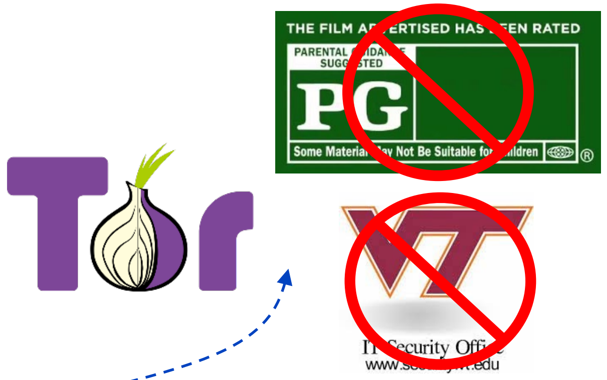
AlanMichaelsVT Create docs_readme ...	7 minutes ago 9
analysis_tools	11 minutes ago
documentation	7 minutes ago
fake_identities	10 minutes ago
parsed_data	9 minutes ago
privacy_policies	7 minutes ago
raw_data	9 minutes ago
README.md	22 minutes ago

https://humeESL.github.io/Use-and-Abuse-PII/

Key Takeaways

Small scale experiment intended testing feasibility of PII traceability

- Addition of phone services ($$) offered rich insights into phone/text behaviors
- Respected companies do not generally share PII
- Prototyped a quantitative scoring method for privacy policies and terms of service
- Observe trends for PII use & abuse within industries and specific outliers
- Surprisingly little domestic political interest from foreign accounts
- Collection mechanisms very constrained



Future experiment with 100k identities; seek answers to:

- Where is bulk spam activity and malicious content originating?
- Can we improve traceability of personal information, particularly phone/SMS?
- Can we improve anonymized tracking and cookie tracing activity?
- Best ways to stimulate "activity" in automated fashion?
- Do companies treat account holders differently as a function of gender, race, age, geography, or any other measurable criteria?
- Special interest questions: politics, timing, OSINT fusion
- Is a crowd sourced interface / open sourced analysis feasible?

HUME CENTER FOR NATIONAL SECURITY AND TECHNOLOGY
VIRGINIA TECH™