Digital Safety for Activists and Organizers

03.12.2023 aequa Berlin, DE
Emily Gorcenski

Agenda

1. Introductions
    a. House rules
    b. A bit about the hosts
    c. A bit about you!
2. Understanding Threats
    a. Terms
        i. **Workshop: understanding your Sphere**
    b. Exploring Threat Modeling
    c. Exploring Threat Actors
        i. State actors
        ii. Employers
        iii. Non-state actors
        iv. Comrades
    d. Homework: cataloging your attack surface

10 minute break

3. Understanding Protections
    a. Your communications
        i. Workshop: setting up Signal
        ii. Workshop: understanding VPNs
    b. Your phone
    c. Your accounts
        i. Workshop: set up 2FA
    d. Your identity
        i. Workshop: your public footprint
    e. Harassment defense and anti-doxxing
    f. What does not work
4. Digital Safety Advocacy
    a. EU – eIDAS
    b. US – KOSA

Introductions

About Emily Gorcenski

- 5+ years in Berlin
- Technologist / data scientist
- Writer
- Anti-fascist activist
- Created First Vigil, a project tracking white supremacist and neo-nazi court cases in the United States

Housekeeping:

- Content warning: some of the topics might be disturbing and deal with surveillance, police and prison, and internet harassment.
- Exits are at the front and back of the space.
- Facilities are in the back, let us know if you need barrierfrei access.
- Be kind to each other.

Understanding Threats

Key Terms

Digital safety – the extension of the protections of the **personal sphere** to the digital space

Attack surface – a collection of **vulnerabilities** or **vectors** that can be exploited by a threat actor

Threat model – an assessment of potential **risks**, **harms**, and **actors** that can affect you, together with strategies for **mitigating** or **avoiding** them

Threat actor – an individual or entity that can cause harm, be it through **targeted**, **random**, or **negligent** actions

Workshop! Understand your personal sphere

Instructions – take some stickies and a marker.

One each sticky, write something in your personal sphere you want to protect. Why do you protect it? How do you protect it?

This doesn't have to have anything to do with being online, and it can be something physical or something abstract.

Do this for one or two things, and discuss it with a pair.

Example –

My old cell phone

Preserving text messages for legal reasons

Store it in a bank vault

Threat model

Threat model – an assessment of potential **risks**, **harms**, and **actors** that can affect you, together with strategies for **mitigating** or **avoiding** them

Transfer – how can risk be shared or shifted?

Avoid – how can risk be sidestepped?

Mitigate – how can harms be reduced?

Accept – what costs can we live with?

WHO	HOW	WHAT
Who is the threat actor?	What are their capabilities?	What harm arises if they succeed?

CAUTION: WRITING THIS INFORMATION DOWN MAY ITSELF CONSTITUTE A THREAT.
HAVE A DATA DESTRUCTION STRATEGY.

(covered in the second half)

State threat actors

Metadata – Data about data, such as who you call/text and when, but not what you say.

Recognition AI – Using Artificial Intelligence tools to match a person in an image or video to a name. Includes **facial recognition**, **number plate recognition**, and **gait analysis**, among others.

Geofencing – Capturing all cell phone numbers in a certain geographic area, which does not require a person to have used their phone

OSINT – open **s**ource **int**elligence, or using publicly-available information to build a case, counter-insurgency strategy, etc.

Stingray – A device that mimics a cell tower but routes all connections through a data capture device controlled by the operator

Parallel construction – the act of using illegal evidence to guide the development of a case using legal evidence

Zero-day – a computer exploit that has not been detected by security researchers and as such has no mitigations

CLOUD Act – a US law that forces US-based companies to surrender data even if it is stored overseas

B. The Google "Geofence" Warrant

1. A "geofence warrant" is "a warrant to obtain cellular phone data generated in a designated geographic area." *In re Information Stored by Google*, No. 21-sc-3217, 2021 WL 6196136, at *2 (D.D.C. Dec. 30, 2021) (citation omitted). The "geofence" in a "geofence warrant" is "the boundary of the area where the criminal activity occurred and is drawn by the government using geolocation coordinates on a map attached to the warrant." *Id*.

On January 13, 2021, the government applied for – and a magistrate judge issued – a geofence warrant directing Google to disclose certain Location History information for devices that connected to its services from a specific geographic area corresponding approximately to the U.S. Capitol Building during specific time windows on January 6, 2021:



1:21-cr-00687-RC Dkt. #59

(USDC D.C.)


PODS
& Storage





11. On the same date, I and other FBI agents viewed the publically-available Etsy webpage for the "KEEP THE IMMIGRANTS, DEPORT THE RACISTS" tee shirt sold by Etsy Store A. On this webpage, I observed that the appearance of the tee shirts for sale on this page appeared to match the style and font of the writing on the tee shirt of the female subject. On the webpage advertising this tee shirt, users are allowed to post comments. These comments were reviewed, and it was found that on March 24, 2020 a user with the display name of "Xx Mv" wrote a 5 star review stating, "Fast shipping, thanks very much!" The publically-available Etsy profile for "Xx Mv" displayed a user location of Philadelphia, PA, and the user name from the Etsy URL for that profile displays as "alleycatlore."

12. Open source searches for user name "alleycatelore" returned a Poshmark user with a display name of "lore-elisabeth." Poshmark is a mobile fashion marketplace built around real-time social experiences. "Posh Parties" are themed, real-time shopping events where people gather to shop, share, and sell fashion right from their phones.

13. Open source Searches for "Lore Elisabeth" in Philadelphia returned a website and a LinkedIn profile for an individual matching the name "Lore Elisabeth" who appears to be employed as a massage therapist with a company that provides massage therapy services.

14. The website for this company has multiple massage therapy videos posted. In these videos, a woman appearing to match the driver's license photograph of BLUMENTHAL is seen. In several of these videos, which appear to be approximately four-years old based on the information visible on the publically-available website, this woman's (again, believed to be BLUMENTHAL) right forearm is visible with the identical tattoo of a "peace sign" in the same location on her right forearm as is observed in the photograph of the female arson subject in Philadelphia on May 30, 2020. Additional tattoos appear around the "peace sign" tattoo observed in the photograph from May 30, 2020 that are not seen in the four-year-old video. However, the unique stylization of the "peace sign" tattoo appears to be identical in location, size, and appearance. A screen shot of this tattoo from one of the videos is below:



2:20-cr-00233-RBS Dkt. #1

(USDC Eastern PA)

Employers / corporate actors

MDM / BYOD – Mobile **D**evice **M**anagement is a way to manage corporate laptops, phones, etc. remotely. It allows for spying, remote wiping, and policy and permissions management. MDM is often required, even for companies with a **B**ring **Y**our **O**wn **D**evice policy.

Dual boxing – Using different computers (“boxes”), phones, accounts, etc. for personal and work reasons.

Contact tracking – Users who grant access to contacts lists give companies unfettered insight into social graphs.

Discovery – The legal right-to-access for personal documents in the course of a legal process. It requires a person to hand over records, including private messages.

Union-busting – Employers can read DMs, emails, etc. at any time, and may use them to disrupt labor organizing without legal accountability.

Private browsing – Browsing the internet without creating a *local* (on your device) history. This does **not** deter MDM or other ways to track your browsing history.

Non-state actors

Online harassment – Digital harassment that can include **brigading**, **doxxing**, **swatting**, **defamation**, **revenge porn**, **deepfakes**, often done with intent to induce self-harm, to make the target lose their job, or to undermine the target's mental and physical safety or prosperity.

Trolling – A form of harassment designed to frustrate the target and cause them to waste time and/or mental energy fighting an unwinnable argument.

Doxxing – The exposure of private personal information in an attempt to antagonize or intimidate a target.

Swatting – Calling in a fake police report in an effort to cause an armed police response.

Ransomware – Software used to take control of a system and hold it hostage in exchange for money or information, often by rendering data unusable.

Phishing – Sending convincingly-created fake messages to a user in an attempt to capture their credentials or data.

SIM Jacking – Convincing a mobile provider's customer service to give an attacker access to the target's cell phone credentials, allowing the intercept of data.

Non-state actors

Defamation – A trolling technique to spread false information about a person in order to injure their reputation or ruin search results for their name.

Revenge porn – The non-consensual sharing of intimate images, often originally created and shared with consent in private contexts, to a broader or public audience, often with the intent to injure the reputation of the target or promote harassment.

Brigading – Harassment of a target through promoting online mobs to send large amounts of hateful or antagonistic messages to a target. Often coordinated on platforms other than where it takes place, making it hard to track or prevent.

Deepfakes – AI-generated images, video, or audio of a target used to defame, mislead, misinform, or harass a target.

He Has 17,700 Bottles of Hand Sanitizer and Nowhere to Sell Them

March 14, 2020



Doug Strickland / The New York Times


U-HAUL
U-HAUL

Comrades

“Weakest link” – Asymmetry in threat models and safety practices can lead to unintentional harm. Messages have a sender and at least one receiver. The other person may be easier to compromise.

Coercion – Threat actors can induce, mislead, or coerce others to give details or access to compromising or protected information.

“Loose lips” – Comrades with different threat models may volunteer information without coercion, out of ignorance, braggadocio, or negligence.

Internecine drama / infighting – Public disagreements provide attackers with exploitable information. Bitterness, jealousy, or anger can lead to intentional or unintentional harm. **Infighting is often more harmful than attacks from the outside.**

Printers – a bonus horror

Microdots – many (if not all) modern printers add invisible secret tracking dots.

Anything printed with a modern printer should be considered traceable. Many consumer-grade printers also require using internet-connected accounts.

A good shredder is an essential activist tool, but most consumer devices won't deter highly motivated attackers. Look for Level P-5 or higher for the best security.

Homework! Catalog your attack surface

Instructions – think about your digital life.

Make a list of your communications tools.

Make a list of your social media accounts

Make a list of your financial/utilities/etc accounts

Make a list of devices you have access to

Do a Google search for your name with a clean browser – what comes up?

For every item in the list, think about what it might mean if someone else had access to it without your knowledge or consent.

In the next section, we'll explore how to manage this.

Break time!
Recharge your soul

Understanding Protections

Encrypted communications

End-to-end encrypted messaging – messaging that is secured the entire way from sender to receiver, where messages are only readable by the intended recipient

Why it matters: encrypted communications can't be understood even if they are intercepted unless the attacker has physical access to the device – or access to the encryption keys

Signal – encrypted messaging
Caution: Signal exposes your usage of the app to anyone who knows your number

Protonmail – encrypted email
Caution: protonmail is only secure if sending to another protonmail user

Whatsapp, Telegram, Fb Messenger, Lime – not encrypted by default!

Twitter, Instagram, Mastodon, Gmail – not encrypted at all!

Workshop! Download Signal

Instructions – download Signal on your mobile device or computer. Signal is end-to-end encrypted, meaning the service cannot access your content. Signal also retains very little metadata, so state actors cannot exploit it easily.

Disappearing Messages – auto-delete chats after a set window. This “future-proofs” against later attacks.

Note: the time at which disappearing messages are set is visible to anyone with access to the device. Turn them on early and leave them on, otherwise changing the setting can be taken as evidence used against you.

Safe Browsing

Virtual Private Network (VPN) – a way to route your internet traffic through protected servers located worldwide.

Why it matters: VPNs prevent governments or ISPs (Internet Service Providers) from snooping on your web behavior. They can also make you appear to be surfing from another country.

Private browsing – using the internet without maintaining a local history

Why it matters: anyone with access to your computer won't be able to *easily* identify your browsing behavior

Browser fingerprinting – a unique signature that trackers can use to uniquely identify your browser

Why it matters: Even with VPNs and private browsing, your computer probably has a unique and visible signature that can be used to track you. It is very hard to avoid this.

Transport Layer Security (TLS) – A way to encrypt web traffic using special trusted certificates so that you can securely send and receive information.

Why it matters: TLS ensures only intended recipients can read data you send. But this depends on *certificate trust*. It also does not prevent ISPs from seeing what websites you use.

Safe Browsing

Firefox – the last truly independent mainstream web browser on the market. Almost all other browser use Chromium, the engine designed by Google.

Why it matters: Google has made many attempts in the last years that would hurt privacy. The latest attempt is called “Manifest V3”. This is planned to go live in mid-2024. Manifest V3 will seriously inhibit ad blockers. Online ads can track users or install malware. Firefox is one of the only safe alternatives at the moment. The Tor browser is another.

Tracking Parameters – a lot of web tracking is done using tracking parameters, the mess of symbols at the end of a URL.

Why it matters: If you share a link with tracking parameters to a comrade, it’s possible that the site can then connect you and your comrade by connecting your tracking parameters to their browser fingerprint. In this way, it’s possible to build social graphs without access to social media.

Workshop! Download Proton VPN

Instructions – download Proton VPN on your mobile device or computer. It is free of charge and like Signal, Proton collects minimal data. Also, Proton is shielded by Swiss privacy laws.

Note: VPNs do not necessarily auto-connect. Be sure to always double-check whether your VPN is connected before browsing the internet

Mullvad – Mullvad is probably the best service available now. You can pay with cash (by mailing an envelope of cash to Sweden) and the service runs **in RAM**, meaning that if the computers are shut off, all memory of your browsing is deleted.



Your phone

Grayshift – A company that makes **GrayKey**, a tool built to bypass mobile device security features. Little is publicly known about its capabilities.

Why it matters: What little is known is that some older iPhone models running older iOS versions can be easily cracked if the passcode is 4 or 6 characters. **It is necessary to assume that this capability extends to all phone models and operating system versions. Only full-strength, alphanumeric (letters and numbers) passwords should be considered secure.**

iCloud – A service provided by Apple to provide cloud storage and backup for iOS devices. iCloud usage is required for some apps/features.

Why it matters: iCloud can be encrypted but this is not presently the default setting. It is very difficult to keep data locally on a phone. And **thousands** of law enforcement requests for iCloud data have succeeded.

App passwords – Certain apps can require passwords to access certain features/data, e.g. menstrual cycle data. **This password should be different from your phone password.**

Your accounts

Two-Factor Authentication (2FA) – Using a second form of authentication, such as a **T**ime-based **O**ne **T**ime **P**assword (TOTP) or physical key that prevents account access even if an attacker has the account password.

Why it matters: Tons of data breaches! 2FA helps add a layer of protection. However, 2FA over **SMS** (**S**mall **M**essaging **S**ervice, aka texting) is considered insecure due to SIMjacking.

Password managers – A tool that stores and generates unique, secure, and un-rememberable passwords.

Why it matters: Reused passwords are the #1 security risk by a long shot. Billions of passwords are available for sale on the dark web. Easily remembered passwords are easily cracked. Strong passwords should be **high entropy** (using lots of different characters) and unique to each account. Password managers let you manage many passwords while remembering only one passphrase (e.g. correct-horse-battery-staple)

Workshop! Download FreeOTP

Instructions – download FreeOTP on your mobile device or computer. It is free of charge and doesn't share any data over the internet.

Note: when you set up 2FA you will be given backup codes. DO NOT LOSE THESE. The downside of 2FA on the phone is that if your phone is lost or damaged, your account may become inaccessible. You can store these codes in a password manager.

Other authenticators are available, for instance similar tools from Google and Microsoft. But I'm trying to stay away from big tech in this workshop.



Your accounts

Map your accounts! What accounts do you have and who can access them? Consider:

- Social media
- Banking
- Utilities, mobile service, insurance, etc.
- Online entertainment (gaming, Netflix, etc)
- Government accounts
- Dating apps
- Fitness apps, delivery apps
- Travel
- Email
- “Find my phone” and location sharing

Refuge Tech Safety Digital Breakup Kit



Your identity

Digital Footprint – The traces you leave online through your activity, including blogs, social media posts, images, videos, comments, and so on.

Why it matters: Embarrassing photos, opinions you no longer hold, arguments that you regret... all of this can be used against you. And archive tools like the **Wayback Machine** can preserve this content even if you delete it.

Right to be forgotten (RTBF) – The General Data Protection Regulation (GDPR) gives you the right to be forgotten. Unfortunately, few services are fully compliant.

Why it matters: There are many tools that can help you e.g. mass delete social media posts, submit deletion requests, or remove your data from data brokers. Periodically reducing your online footprint helps improve digital safety.

Your identity

Impersonation – Few controls exist to prevent a person from impersonating another online, a common form of harassment.

Why it matters: You generally hold the copyright to any writings, images, videos, etc. you put online. **But copyright takedown requests will expose your personal data to the impersonator.** A lawyer can help prevent this, for a fee.

Impressum – German law mandates an imprint or Impressum on every website (and technically, every identifiable social media profile).

Why it matters: This is a huge privacy risk. However, impressum services can provide mail forwarding and identity shielding.

Workshop! Explore your digital footprint

Instructions – open Cover Your Tracks from the EFF. Run a test to see how unique your browser fingerprint is.

Note: this is only an approximation based on a small subset of data. Across the larger internet the reality may be more or less extreme.

Bonus Exercise – open a private browsing tab in your browser (incognito mode). Do a Google search for your name. Check the Images tab. This is optional – you may want to do this in private.

Anti-harassment and anti-doxxing

General rules for avoiding doxing:

- Do not post images with identifiable backgrounds near your home or neighborhood, or in real-time (wait a few days before sharing a vacation selfie).
- Use delete-me services to remove your information from data brokers.
- Close unused social media accounts.
- Search social media for any posts you've made that mention your location, address, school, or other details that can be used to identify you.
- Change profile pictures to something other than an image of you.

General rules for demo safety:

- Don't interact with any strangers with cameras or voice recorders claiming to be media.
- **In fact, do not speak to media for anything unless you have been trained to do so.**
- Do not film protests. Faces are not the only identifiable details. Shoes, tattoos, t-shirts, hats, and other details can be equally identifying.
- Leave your phone at home.
- If you have access to sensitive data, do not attend risky demos.

Anti-harassment and anti-doxxing

Online harassment sucks. But we can protect ourselves from it. A layered approach works best.

Body – Are you physically safe? Is your physical location compromised? Are people trying to do bodily harm?

- Stay with friends, family, or in a hotel.
- Order delivery and don't go out.
- Lock doors, close blinds, and arm security systems.
- Flee only if necessary.
- **Do not post your whereabouts.**



Anti-harassment and anti-doxxing

Online harassment sucks. But we can protect ourselves from it. A layered approach works best.

Mind – Are you mentally ok? Do you have ideation of self harm? Are you drinking water, taking your meds, and eating? Are you sleeping?

- Log off and give a friend your phone.
- Call the International Helpline Berlin (030-44 01 06 07) or similar.
- Surround yourself with friends and family 24/7.



Anti-harassment and anti-doxxing

Online harassment sucks. But we can protect ourselves from it. A layered approach works best.

Integrity – Is your integrity or reputation at stake? Are people lying about you?

- Do not argue, you cannot win.
- Work with friends to prepare a statement to post in a few days.
- Assess how valuable the community of harassers is to you (in- or out-community).
- **Lock your social media accounts.**



Anti-harassment and anti-doxxing

Online harassment sucks. But we can protect ourselves from it. A layered approach works best.

Finances – Are you financially injured? Are you at risk of losing a job or opportunity?

- Contact a lawyer
- Ask a friend to monitor your account and make and document real-time archives of defamatory content
- Only if necessary, speak to your boss(es) and inform them you are subject to a harassment campaign



Anti-harassment and anti-doxxing

Online harassment sucks. But we can protect ourselves from it. A layered approach works best.

Network – Are friends, family, comrades, or coworkers likely to be targeted by this campaign?

- Contact people who might also be vulnerable and inform them of the campaign.
- Instruct them not to respond, either in your defense or in their own.
- Ask them to lock their accounts.



What doesn’t work: burner phones

Burner phone – Anonymous phones (e.g. prepaid) that aren’t connected to your name or number.

Why it doesn’t work: Burner phones are a Hollywood trope, not something practical in reality. They are **very** hard to use. You have to consider metadata, location data, usage patterns, temporal data, financial data, purchase data, etc.

To use a burner properly, you must:

- Purchase it somewhere with no direct connection to you, often in another city
- Pay for it with cash
- Never turn it on at home
- Never call anyone on a non-burner number
- Create usage patterns away from your home for it, with no trace to your presence there
- Never log into a service under your name with it

What doesn't work: secret code words and euphemism

Code words – Innocuous words or phrases designed to cover up or add plausible deniability to an act or intention.

Why it doesn't work: Everyone sees through this and they can make it worse because it shows evidence of an intent to hide behavior.

79. Other messages show what appears to be Morris trying to sell a fully automatic "AK74" to UNSUB 1. During this chat the term "in Minecraft" is used. In my training and experience, the term "in Minecraft" is used in chatrooms to protect the author from potential prosecution when making threats, statements of violence, or statements regarding other illegal activity:

1:22-mj-00022-AJ Dkt #1

(USDC N.H.)

What doesn’t work: dazzle camouflage

CV Dazzle – Makeup, clothing, or accessories designed to fool AI.

Why it doesn’t work: Even if it does fool AI recognition algorithms, it doesn’t fool humans. Like euphemisms, it may convey evidence of intent. And it’s not really effective against AI – algorithms have been re-trained around this. All AI spoofing techniques can be trained around.



Pinterest

Digital Safety Advocacy

Privacy is under attack

eIDAS (EU) – An update to the existing eIDAS regulation is about to come up for a first vote in less than 2 weeks. All indications suggest it will pass.

Why to be alarmed: The new regulation puts a backdoor in TLS, forcing browsers to trust certificates issued by any EU government. This means e.g. the Hungarian government can spy on your internet traffic, even if you use a VPN.

Learn more

KOSA (US) – The Kids Online Safety Act is a proposed law in the US that will require platforms to do age verification via real IDs and filter “inappropriate” content.

Why to be alarmed: The proposed regulation has no clear definition of “inappropriate” content and will give states broad reach to decide what to censor. LGBTQ+ communities, mental health communities, and others are at serious risk.

Learn more

Further reading

- *Hate Crimes in Cyberspace*, Danielle Keats Citron
- *Silicon Values*, Jillian York
- *Twitter and Tear Gas*, Zeynep Tufecki
- *Crash Override: How Gamergate (Nearly) Destroyed my Life*, Zoe Quinn

The Electronic Frontier Foundation (EFF)

Refuge Tech Safety

Meta in Myanmar (blog series), Erin Kissane



https://emilygorcenski.com/digital-safety/Digital%20Safety%20for%20Activists%20and%20Organizers.pdf

This work is marked with CC0 1.0 Universal

Images copyright of their respective holders used without permission for educational purposes.

Thanks!