OSINT for evil

by Malcolm Rivera

Three individuals in a conference room intently watch as a television plays a video of someone running on a treadmill, transitioning to a weight bench, then moving to a rowing machine. One of the viewers, Robin, feels a light tap on the shoulder as their attorney asks "Robin, did you hear the question?" Robin turns to face Alex, opposing counsel, and asks "You had people following me? How long has this been going on?" Alex repeats the question "Can you confirm if you are the individual seen in this footage?" Robin's body stiffens, and a quiet "Yes, that is me" escapes their lips. Alex flips through the pages of a document and states, "In your previous deposition – three days after this footage was captured – you stated that your injuries prevented you from attending the gym…" Time blurs under the bright florescent lights, and Robin is eventually informed that they are done for the day. They slowly stand, and just before exiting the room, turn to Alex and say, "I wonder how you would feel if people started following you around". When they return home, Robin navigates to Alex's LinkedIn profile, and begins noting every piece of identifying information they can find:

- Face
- Full name
- Employer
- Professional Email Address
- Education
- Birthday
- Hometown

That should be enough to find them. Robin begins their search with breached data brokers and identifies a unique user handle tied to one of Alex's online profiles. Searching this user handle then reveals breached data taken from a credit reporting agency, which contains a physical address located in Alex's hometown – this must be them. Robin navigates to a popular people search website, pays for a $25 monthly subscription plan, and starts by searching for this address. It is an apartment address, and over 50 associated individuals appear, but they quickly identify the correct profile based on the listed names and birthdates.

Robin begins searching the social media profiles the report has attributed to Alex, and before long, they come across a public account with a matching profile picture. One monitor displays street-view addresses while the other scrolls through the profile's timeline. Two months ago, Alex was tagged in a photo posted by their significant other that shows them napping on a porch, and the paneling of that house looks familiar – "That's right! It's the address on Oak Street!" Early the next morning, before the sun begins to rise, Alex starts to pull out of their driveway when they notice a pair of headlights illuminate nearby. On this silent street, Alex begins their commute by making a right, as does the vehicle behind them. The next block is another right, which the other vehicle makes as well. The driver of this vehicle, Robin, sees Alex's head frantically alternate between looking at the road and their rear-view mirror, and smiles as their previous thought,

"I wonder how you would feel if people started following you around" is answered.

This is an anonymized story that focuses on tools and methodologies that are well within reach of the average individual. While open-sourced intelligence (OSINT) is employed by investigative and law-enforcement personnel to combat bad actors, it is important to remember that any OSINT resource can also be exploited by these same bad actors for malevolent purposes. After reading this paper, you will understand how bad actors obtain your personal data, how you can limit the impact of these bad actors, and how to prevent future bad actors from obtaining your personal data.

WHY IS MY PERSONAL DATA AVAILABLE?

All data available for purchase falls into one of two categories – the data we (knowingly or unknowingly) surrender to the public sphere, and the data that is stolen from us or the services we use. More specifically, the former refers to public records and the content we share online, while the latter refers to breached data (data stolen by criminals).

Regarding public records, while there are jurisdictions that actively protect personal data, the United States is a nation that generally assumes that data held by the government is available to a public records request. The Freedom of Information Act (FOIA) established the standard that "any person has the right to request access to federal agency records or information.[1] " States then modeled their public records legislation after FOIA, including the California Public Records Act which echoes this sentiment with "every person has a right to inspect any public record.[2] " While both acts carve out exemptions for certain record types, they often do not protect our home addresses and other personal data against disclosure. On the contrary, jurisdictions such as Los Angeles County actively advertise the depth and scope of real property data available for purchase.[3] These factors ultimately gave rise to commercial data brokers that not only harvest real property rolls, but also utility records, voter registration, phone numbers, and even social media profiles that are all packaged and sold to anyone with a credit card.[4]

While social media is unlikely to appear on government records, data brokers often identify profiles based on the bevy of data available from these official sources. While updating an online biography to include your hometown, university, or occupation allows friends and family to easily locate your profiles, the same can be said for data brokers and bad actors. It is important to recognize the cost of this convenience is privacy, and not necessarily ours. In 2021, Senators Amy Klobuchar and Lisa Murkowski wrote a letter to the Federal Trade Commission that stated, "people-search sites… often include the names and address of family members. The availability of this data makes it difficult or impossible for [domestic violence] victims to safely relocate with relatives.[5] " When evaluating your digital footprint, you have to account for yourself as well as your relatives and close associates.

Finally, regarding data breaches, this is unfortunately one area that we have little control over beyond utilizing strong and unique passwords. As of 2018, cybercrime was estimated to be a $1.5 trillion industry, with "Data trading" accounting for $160 billion of that total.[6] With figures this high, it becomes apparent why bad actors would be so interested in our data. Once a platform's vulnerability is identified, and its data has been harvested, it can then be sold on the dark web for fractions of a penny per account.[7] After the point of sale, the owner of this data may keep it for personal use, or even disseminate it to the surface web – in the latter scenario, this stolen data can then be harvested by commercial data brokers. In the story this paper opened with, Robin utilized a paid database to locate additional information about Alex; while these are often useful tools for individuals and security professionals to assess risk and build correc-

tive plans, any OSINT resource designed for good can be twisted for nefarious purposes.

WHAT SHOULD I DO TO PROTECT MYSELF?

At this point, it is normal to feel some anxiety about your privacy. The good news is that there is one change you can make to assert control – don't share personal data with anyone unless you absolutely have to. Deed your property under an anonymous trust, make your social media profiles private, and do not reuse login credentials across different platforms - these are three small changes that limit dissemination of your data.

Being privacy-oriented is a habitual practice, and not one that is always going to be the easiest option. However, if there is one actionable item you apply after reading this paper, it should be to block/remove non-essential cookies/trackers on the platforms you visit. One of Meta's recent ad campaigns has pushed the narrative that "Good ideas deserve to be found", and that the company is able to connect you with relevant businesses without sharing any of your identifying information.[8] The issue with this assertion is that research has shown that 99.98% of Americans can be re-identified based on their recorded demographic attributes.[9] Make no mistake when reading a company's pledge to consumer privacy, because the only one in a position to value it are the consumers themselves. To quantify exactly how much consumer privacy is valued, we need not look any further than Amazon's "Ad Verification" program, which tracks the ads you view across different browsers and applications in exchange for $2 per month.[10] Protecting your privacy means assigning it great value than what data brokers are willing to pay for it, and taking measures such as deeding your home to a trust and rejecting cookies are the first steps toward making it more difficult for data brokers to package your data for sale.

HOW CAN I REMOVE DATA AVAILABLE ONLINE?

The names, nature, and number of data brokers are ever-changing, but the methods used to combat them are generally the same. In 2018, California voters passed the California Consumer Privacy Act (CCPA), which gave consumers the right to know the personal data platforms collect, the right to delete this collected data, and the right to opt-out of the sale of their data.[11] In response to this, data brokers created avenues for consumers to place these requests – some less burdensome than others.[12] The CCPA, and subsequent legislation passed in Nevada (2019), Virginia (2021), Colorado (2021), Utah (2022), and Connecticut (2022), has given consumers the ability to reclaim ownership of their data from brokers who are all too eager to disseminate it for a fee.[13] At the time of writing, this means roughly 19% of the US population can cite legislation when exercising their privacy rights. While there are platforms that have adopted a blanket approach when it comes to privacy, it would be unwise to assume that all data brokers will adopt this behavior until a large majority of states enact their own privacy acts as well. Until such a point, any privacy-oriented individual should avoid adding to the already robust data stockpile data brokers have collected by keeping their personal data exactly what it should be – personal.

1 https://foia.state.gov/learn/foia.aspx
2 https://leginfo.legislature.ca.gov/faces/codes_displaySection.xhtml?lawCode=GOV&sectionNum=6253.
3 https://assessor.lacounty.gov/Real-Estate-Toolkit/data-sales.
4 Michael Bazzell, Extreme Privacy: What It Takes to Disappear, 4th ed., 2022, Introduction.
5 https://twitter.com/millsrodrigo/status/1367521239049392136
6 https://www.bromium.com/press-release/hyper-connected-web-of-profit-emerges-as-global-cybercriminal-revenues-hit-1-5-trillion-annually/
7 chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://portal.ct.gov/-/media/DCF/HumanTrafficking/Schools/Language-in-the-life.pdf
8 https://about.meta.com/supportsmallbusiness/personalized-ads/
9 https://www.nature.com/articles/s41467-019-10933-3
10 https://www.entrepreneur.com/business-news/amazon-shopper-panel-will-pay-2-a-month-for-your-data/440535
11 https://oag.ca.gov/privacy/ccpa
12 While many brokers allow request to be placed online, many others require more time-consuming methods such as physical mail
13 https://www.womblebonddickinson.com/us/insights/alerts/long-conn-connecticut-joins-growing-list-privacy-rights-states