Let’s Get Serious About Your Data Privacy

By Brent Collins, Avenir Consulting Group

Cybercrime is continuing to ramp up. Companies are now scooping up massive volumes of data to build new Generative AI models but sometimes failing to adequately protect confidential data and consumer trust. Geopolitical concerns stemming from Russia invading Ukraine and competition with China just add to the growing list of reasons to protect your data. Lack of regulation and enforcement by the US government exacerbates the problem. It's time you learn the basics about data privacy.

When we go online to work, socialize or buy things, we leave a trail of data -- often confidential data about ourselves or our business. We don’t give this much thought since it helps automate the process, drives financial transactions, enables services or features, and supports the security protocols. These are all compelling reasons to share personal information, but what happens when the risk of sharing confidential data gets too high? How will consumer behavior change (human error is how most breaches succeed)? Given the advent of Generative AI and its insatiable demand for data, how strong will demand for data get? How might this lead to an increased demand for blackmarket data that is illegally obtained? Answers are played out in our everyday habits and the emerging business models that influence and capitalize on those habits.

A quick primer: there are three main ways cyber-attacks are used to breach our private data:

1. **Malware** is any program or code that can steal, encrypt or hijack computer functions. It commonly masquerades as a “warning” that convinces users to download software and can penetrate your computer when you navigate hacked websites, or open infected files and emails on a device that lacks anti-malware security.

2. **Phishing** scams are one of the most common ways criminals hack sensitive or confidential information. Phishing involves fraudulent emails that appear to be from a reputable source, with the goal of deceiving recipients into either clicking on a malicious link or downloading an infected attachment, usually to steal PII.

3. **Ransomware** software gains and locks access to an organization’s vital data. Files and systems are locked down and a fee is demanded, commonly in cryptocurrency. Private information can be held hostage, lost, or exposed in the process.

Criminals organize, access hacking tools, and buy/sell illegally obtained Personally Identifiable Information (PII) using the “dark web.” The dark web is part of the internet but consists of hidden sites that can only be found by special browsers and search engines. Here criminals buy personal information like people’s Social Security numbers, birthdates, addresses, and phone numbers. These credentials are then used to steal your identity, which can then be used to apply for credit cards in your name, apply for mortgage loans, and even file your income taxes in the hope of stealing your refund. The advent of cryptocurrency in 2009 provided a major boost because digital currency lets users purchase items anonymously.

Mo Data, Mo Problems

Since the arrival of COVID19, a wave of people, businesses and government services have been compelled to move online, making more private information available and creating excessive demand for cyber security resources. In 2021-2025, data consumption is expected to increase 30% with some Telco's reporting carrying over 60% more data than before the COVID crisis (Global Entertainment and Media Outlook, PWC).

The total volume of data created, captured, copied and consumed was about 64.2 zettabytes in 2020 and around 79 zettabytes in 2021. About 2% of that data is saved and retained. This translates to a compound annual growth rate of 19% for data storage capacity.

Our growing dependence on data obviously leads to more personal information becoming available and exposed. Increased dependence goes hand in hand with potential for disruption.

Let's look at specific indicators of this widening gap between risk and readiness, and reasons why individuals, families, businesses and public entities need to get serious about protecting private data in 2022.

Personally Identifiable Information (PII) is information that can be used to identify or trace an individual, or can be combined with other information that can be linked to that individual, such as:
- Social Security Number
- Date and place of birth
- Mother's maiden name
- Biometric records
- Protected Health Information
- Passport number

Business Identifiable Information (BII) is similar to PII but relates to confidential information that businesses share online.

Protected Health Information (PHI) is:
- a subset of PII requiring additional protection (see HIPPA);
- health information created or received by a healthcare provider, health plan, employer, or business partner identifying an individual.

The Problem: Cyber-attacks are now the fastest growing crime in the U.S. This relates to the massive increase in available data, our relative ineffectiveness at protecting it and the low probability that perpetrators will get caught.

1,862

Total *reported* data beaches in 2021, involving 293,927,708 victims. This is a 68% increase over 2020. (Identity Theft Resource Center)

$6T

Cost of cybercrime worldwide in 2021, up from $3 trillion in 2015. Global cybercrime costs are expected to grow by 15 percent per year, reaching $10.5 trillion by 2025. This makes Cyberattacks the fastest growing crime in the U.S. (Cyber Security Ventures)

> **"Cybercrime represents the greatest transfer of economic wealth in history, puts at risks the trust needed for innovation and investment, and will be more profitable than the global trade of all major illegal drugs combined."**
> **-CyberCrime Ventures**

435% Rate of increase in ransomware in 2020. Malware attacks increased by 358%. (2022 WEF Global Risks Report).

95% Share of cybersecurity issues that can be traced to human error. (WEF).

115M Number of stolen debit and credit card numbers posted to the dark web in 2020. (Gemini Advisory)

25% Percentage of 15b exposed credentials collected for analysis by Digital Shadows that were related to financials or banking.

$1,100 Average price for credentials to a bank account with $20,000 balance or more. The price for credentials to a bank account with $3,000 or less ranges from $150 to $300. (The Black Market Report - Armor)

21% Share of identity theft victims who lost more than $20,000.00.
24% reported that they were denied unemployment benefits.
21% lost unemployment benefits because of identity theft.
(ITRC Consumer Aftermath Report)

$30-40 Cost range on the dark web for a "fullz" - a packet of personally identifiable information including a victim's full name, date of birth, Social Security #, phone #, address, mother's maiden name, driver's license #, etc. (The Black Market Report - Armor)

34% Share of data breaches Verizon investigated in 2018 caused by *internal* actor. Outside actors perpetrated 69% of breaches, and 5% involved both.
(2019 Data Breach Investigations Report - Verizon)

.05% Likelihood of detection and prosecution in the U.S. (World Economic Forum's Global Risk Report)

Societal Response.

49% Share of European online users who are aware of domestic data protection and privacy rules, compared to 29% of North American online users. (Statista)

50% Share of private sector businesses already experiencing gaps in basic, technical IT security skills. Compounding the problem: 32% of IT managers and 25% of IT directors are considering quitting their jobs in the next six months – resulting in a wave of issues across HR, management, and IT security. (ThreatConnect survey of more than 500 IT decision makers)

66% Percentage of Industry influencers who cited data security as biggest challenge in moving to the public cloud. 57% expressed the same concern over data privacy in cloud environments. (The Future of the Cloud Study - LogicMonitor)

34% Share of US users who feel their personal data is 'very vulnerable' to compromise. Another 47% feel "somewhat vulnerable" on the issue. Only 2% don't feel their data is vulnerable to compromise. (Statista)

287 Average number of days it takes to detect and contain a data breach.

The Intermediaries

534,465 Number of files containing sensitive data at the average company. More than half of the data (53%) at the average organization is stale; 58% of organizations have at least 1,000 stale user accounts. (2019 Global Data Risk Report - Varonis)

$180 Average cost for each PII record breached. The overall average for all lost or stolen record types is $161, an increase from $146 in 2020. (Cost of Data Breach 2021, IBM)

43% Share of cyberattacks aimed at small businesses. A business falls victim to a ransomware attack about every 14 seconds. (Accenture's Cost of Cybercrime)

$4.24m Average cost of a breach, which rose from $3.86 million. The average cost was $1.07 million higher in breaches where remote work was a factor in causing the breach, compared to those where remote work was not a factor. (IBM)

80% Cost savings when security AI and automation was fully deployed $2.9m) vs not ($6.7m). (Cost of Data Breach 2021, IBM)

44% The portion of organizations that rated complexity as the top barrier to good data security. Based on a survey of 1,200 IT and security executives. The move from single on-premises environments to multiple SaaS, IaaS, and PaaS environments is driving much of the complexity. (2019 Data Threat Report - IDC, Thales)

31% Share of organizations that encrypt data at rest on PCs. Though awareness is high about the need for data encryption, deployment is lacking for primary user cases, like full disk encryption, public cloud workloads, big-data environments, mobile devices, IoT, & containers. (2019 Data Threat Report - IDC, Thales)

10% Share of US companies actively working to comply with 50 or more privacy laws. Some 13% reported working actively on between 6 and 10 data privacy laws at the same time, and 13% on between 11 and 49 laws. (IAPP and TrustArc Report)

$55b The initial overall cost to California companies of complying with the CCPA. This includes legal, operational, technical, and business-related costs like renegotiating contracts and changing data-handling practices. (Standardized Regulatory Impact Assessment - California Office of the Attorney General)

3M The person gap in cyber professionals needed worldwide (WEF)

Regulations

With only a few states regulating how PII data is controlled, and federal regulations are not as comprehensive as GDPR which serves as a benchmark for protecting consumer data. Lack of effective federal laws and regulations increases the likelihood of PII being stolen as time goes on. Statistics reflect a society that is grappling with the growing risks and not positioned to effectively manage the fast-growing problem.

On March 1, the Senate unanimously passed the Strengthening American Cybersecurity Act of 2022, which will require critical infrastructure companies to report significant cyber-incidents and all ransom payments to the Department of Homeland Security's Cybersecurity and Infrastructure Security Agency (CISA). The Act was included in the 2022 omnibus spending bill, which President Biden signed into law on March 15. Here is what companies need to know. Learn more about the Biden administration’s plans here. To understand how federal laws and regulations apply to cybercrimes, see this overview from ICLG.

Current legal, regulatory, and policy frameworks in the US at the state and federal levels of government struggle to address the aggressively growing threats from data breaches. Europe figured out the importance of regulating online privacy and implemented GDPR starting in 2016. In the US, California was the first state to take a stand with CCPA -- only Colorado and Virginia have since passed similar legislation. Other countries like Canada (PIPEDA), Japan (APPI) and Brazil (LGPD) have their own versions, but regulatory frameworks have far to go before they are unified, effective and users can feel safe.

One way to find specific examples is to look at breach disclosures, California law requires a business or state agency to notify any resident whose unencrypted personal information was acquired, or reasonably believed to have been acquired, by an unauthorized person. As we all recognize, the consumers is left to face the consequences, since there is little anyone can do once their PII has been breached. You can search and view a list of sample breach notices here. https://oag.ca.gov/privacy/databreach/list

Building Better Cyber Resilience – What You Can Do

To combat the increasing threat of PII breach, organizations must look at protecting client data as an important aspect of maintaining sustainable operations.

Assess your risk

You can minimize human error which often exposes private or confidential information. A general PII risk assessment should review these common ways PII is exposed:

- Postings of PII on public websites and social networks
- PII sent via email to unauthorized recipients
- Transmitting unsecured emails and unencrypted files containing PII
- Providing hard copies containing PII to individuals without a need to know
- Failing to properly secure documents containing PII when mailing or transporting
- Documents containing PII that can reach anyone other than the intended recipient
- Lost or stolen electronic devices or media storing PII
- Successful network intrusions
- Unauthorized access to computer systems
- Inappropriate disposal of documents containing PII
- PII use by employees for unofficial business
- Unauthorized access to credit card information
- Authorization, access and control processes (including a temporary loss of control)
- Notification and follow up processes
- Health records held by your doctor and 3rd party applications

Standardize Privacy Impact Planning

Projects and programs should involve a data privacy policy which sets out how confidential or private information is being collected, used, and stored within the organization. It can be performed on new or existing processes to quantify risk and readiness. An analysis and plan can be performed at the beginning of a project, new vendor relationship, new software implementation, or in conjunction with an organizational change of some kind. Consultants can be brought in to assess risk and readiness, and suggest plans to improve.

Factors to be assessed include:

- Whose data is being collected? (employees, website visitors, business partners, etc.)
- Why is this data being collected?
- What is the scope or volume of the data being used?
- What controls do the subjects have over their data?
- What risks are associated with this data useful?
- What data privacy controls are being used to mitigate that risk?
- Are controls in compliance with appropriate data privacy regulations?
- How is data classified and tagged?
- Who owns data within the organization and what is the succession plan?
- Who accesses the data and how is access controlled?
- Monitoring & logging detection and prevention?
- Is data being encrypted?
- What data is being shared with 3rd parties and partners, for what reasons?

- Is some data anonymized or masked before being stored or shared?
- For certain sets of data, what is the lifecycle (collection, storing, updating, processing, deleting and transferring)?
- What is the data backup process?
- What is the audit process for detecting when policies are not enforced?
- Is there a sound process disaster and recover plan that responds to incidents?

Regulate and Lead

In our connected society, digital trust facilitates future innovation and prosperity. Trust represents a foundation on which a fair and cohesive society is built. Global leaders in business, government and society must recognize the need to protect the digital security of their client base as they build the data infrastructure required for their business model.

Organizations must establish norms, best-practices and rules of behavior for all stakeholders. Senior executives must remain aware and engaged by continually assessing and adapting resources used to fight cyber-crime. At the organizational level, department leaders must be trained on cybersecurity issues, emerging cyber risks, and key indicators tracking progress. This should extend to board-level commitment to strengthening cyber-resilience.

If we fail to protect private data, trust will continue to degrade, cyber-criminal organizations will grow, and economic potential will be lost. We must rise to the challenge of protecting the data that fuels innovation, drives economic growth, and requires trust to share and utilize.

Brent Collins
Avenir Consulting Group
+1 (714) 906-7113

About: Brent is managing director at Avenir Consulting Group. ACG helps public and private sector clients to organize and deploy sustainable development projects using best industry practices and leading technologies that deliver optimal civic value.
For more information, go to https://avenircg.com/

The following appendix provide more details on the biggest and most notable breaches in recently years, as well as definitions of terminology used.

APPENDIX

A) Biggest PII Breaches

CAM4 **Date:** March 2020**. Impact:** 10.88 billion records.
Adult video streaming website had its server breached, exposing over 10 billion records.

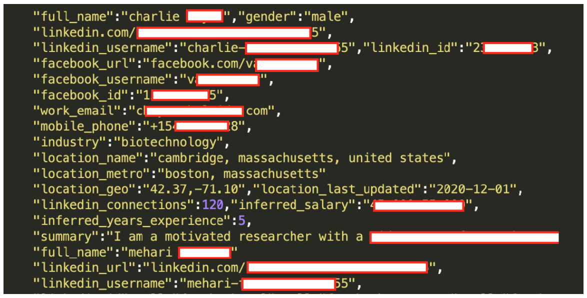
Yahoo **Date**: October 2017**. Impact**: 3 billion accounts
The breach was first reported by Yahoo while in negotiations to sell itself to Verizon, on December 14, 2016. Yahoo forced affected users to change passwords and security questions.

Aadhaar **Date:** March 2018**. Impact:** 1.1 billion people
In March of 2018, it became public that the personal information of more than a billion Indian citizens stored in the world's largest biometric database could be bought online. The type of information exposed included the photographs, thumbprints, retina scans and other identifying details of nearly every Indian citizen.

First American Financial Corp. **Date**: May 2019. **Impact**: 885 million users
In May 2019, First American Financial Corporation reportedly leaked 885 million users' sensitive records that date back more than 16 years, including bank account records, social security numbers, wire transactions, and other mortgage paperwork.

Verifications.io **Date: February 2019. Impact: 763 million users**
In February 2019, email address validation service verifications.io exposed 763 million unique email addresses in a MongoDB instance that was left open with no password. Many records also included names, phone numbers, IP addresses, dates of birth and genders.

LinkedIn **Date**: June 2021.
Impact: 700 million users
Data associated with 700 million LinkedIn users was posted for sale in a Dark Web forum on June 2021. This exposure impacted 92% of the total LinkedIn user base of 756 million users.



```
"full_name":"charlie [redacted]","gender":"male",
"linkedin.com/[redacted]5",
"linkedin_username":"charlie-[redacted]5","linkedin_id":"2[redacted]3",
"facebook_url":"facebook.com/v[redacted]",
"facebook_username":"v[redacted]",
"facebook_id":"1[redacted]5",
"work_email":"c[redacted]com",
"mobile_phone":"+15[redacted]8",
"industry":"biotechnology",
"location_name":"cambridge, massachusetts, united states",
"location_metro":"boston, massachusetts"
"location_geo":"42.37,-71.10","location_last_updated":"2020-12-01",
"linkedin_connections":120,"inferred_salary":"[redacted]",
"inferred_years_experience":5,
"summary":"I am a motivated researcher with a [redacted]
"full_name":"mehari [redacted]"
"linkedin_url":"linkedin.com/[redacted]",
"linkedin_username":"mehari-[redacted]55",
```

Hackers published a sample containing 1 million records to confirm the legitimacy of the LinkedIn breach. (9to5mac.com)

Facebook **Date:** April 2019.
Impact: 533 million users
In 2019, the UpGuard Cyber Risk team revealed two third-party Facebook app datasets had been exposed to the public Internet. This massive database was then leaked on the dark web for free in 2021, adding a new wave of exposure to the data originally breached in 2019.

Yahoo **Date:** 2014. **Impact:** 500 million accounts
Yahoo believed that a "state-sponsored actor" was behind this initial cyberattack in 2014. The stolen data included personal information such as names, email addresses, phone numbers, hashed passwords, birth dates, and security questions and answers, some of which were unencrypted. Yahoo become aware of this breach back in 2014, taking a few initial remedial actions but failed to investigate further. It was two years later that Yahoo publicly disclosed the breach after a stolen database from the company went up for sale on the black market.

Starwood (Marriott) **Date**: November 2018. **Impact**: 500 million guests
In November 2018, Marriott International announced that hackers had stolen data of ~500 million Starwood hotel customers. The attackers gained unauthorized access to the Starwood system back in 2014 and remained in the system after Marriott acquired Starwood in 2016. However, the discovery was not made until 2018. According to the NY Times, the breach is attributed to China's Ministry of State Security, seeking to gather data on US citizens. If true, this would be the largest known breach of personal data conducted by a nation-state.

Adult Friend Finder **Date:** October 2016. **Impact:** 412.2 million accounts
In October 2016, hackers collected 20 years of data on six databases that included names, email addresses and passwords. Most passwords were protected only by the weak SHA-1 hashing algorithm, which meant that 99% of them had been cracked by the next month.

See more at UpGuard: 62 Biggest Breaches Ever.

B) Notable PII Breaches

Parler
January: News of the conservative social media app, Parler, having its data scraped by a hacker came to light after Amazon Web Services removed the platform from its servers. The 70TB of leaked information includes 99.9% of posts, messages, and video data containing EXIF data — metadata of date, time, and location. Parler's Verified Citizens, or users who had verified their identity by uploading their driver's license or other government-issued photo ID, were also exposed.

Facebook, Instagram and LinkedIn
January: A Chinese social media management company, Socialarks, suffered a data leak through an unsecured database that exposed account details and Personally Identifiable Information (PII) of at least 214 million social media users from Facebook and Instagram, and LinkedIn. The exposed information for each platform varies but includes user's names, phone numbers, email addresses, profile links, usernames, profile pictures, profile description, follower and engagement logistics, location, Messenger ID, website link, job profile, LinkedIn profile link, connected social media account login names and company name.

Pixlr

January: A database containing 1.9 million user records belonging to Pixlr, a free online photo-editing application, was leaked by a hacker. The database was stolen at the same time as the attack on 123RF, which exposed over 83 million user records. The leaked records include email addresses, usernames, hashed passwords, user's country, whether they signed up for the newsletter, and other sensitive information.

MeetMindful

January: The dating platform, MeetMindful.com, was hacked by a well known-hacker and had its user's account details and personal information posted for free in a hacker forum. The leaked details of more than 2.28 million users registered included names, email addresses, location details, dating preferences, marital status, birth dates, IP addresses, Bcrypt-hashed account passwords, Facebook user IDs and Facebook authentication tokens.

Bonobos

January: Customer data was stolen from the men's clothing retailer, Bonobos, was found for free in a hacker forum after a cybercriminal downloaded the company's backup cloud data. The exposed database contains order information for over 7 million customers, including addresses, phone numbers, and account information for 1.8 million registered customers, and 3.5 million partial credit card records.

VIPGames

January: VIPGames.com, a free gaming platform, exposed over 23 million records for more than 66,000 desktop and mobile users due to a cloud misconfiguration. The leaked user records include usernames, emails, IP addresses, hashed passwords, Facebook, Twitter and Google IDs, bets and data on players who were banned from the platform.

U.S. Cellular

January: Through a targeted attack on retail employees of U.S. Cellular, the fourth-largest wireless carrier in the U.S., hackers were able to scam employees into downloading malicious software onto company computers. Once downloaded, the software granted remote access to the company devices and to the customer relationship management (CRM) software containing account records for 4.9 million customers. While viewing a customers' account in the CRM, the hacker had access to names, addresses, PINs, cell phone numbers, service plans, and billing/ usage statements.

"Compilation of Many Breaches" (COMB) -

February: A database containing more than 3.2 billion unique pairs of cleartext emails and passwords belonging to past leaks from Netflix, LinkedIn, Exploit.in, Bitcoin, Yahoo, and more were discovered online. This is the largest compilation of data from multiple breaches, which is where the name "Compilation of Many Breaches" or COMB comes from. The searchable and well-organized database was leaked to a popular hacking forum, giving hackers access to account credentials, including approximately 200 million Gmail addresses and 450 million Yahoo email addresses, and more.

Nebraska Medicine
February: A malware attack allowed a hacker to access and copy files containing the personal and medical information of 219,000 patients of Nebraska Medicine. The health network notified affected individuals that the accessed information includes names, addresses, dates of birth, medical record numbers, health insurance information, physician notes, laboratory results, imaging, diagnosis information, treatment information, and/or prescription information, and a limited number of Social Security numbers and driver's license numbers.

California DMV
February: The California Department of Motor Vehicles (DMV) alerted drivers they suffered a data breach after billing contractor, Automatic Funds Transfer Services, was hit by a ransomware attack. The attack exposed drivers' personal information from the last 20 months of California vehicle registration records, including names, addresses, license plate numbers and vehicle identification numbers (VINs).

Kroger / Accellion
February: A third-party data breach at cloud solutions company, Accellion, allowed hackers to steal human resources data and pharmacy records belonging to the supermarket giant, Kroger. The records disclosed could include names, email addresses, phone numbers, home addresses, dates of birth, Social Security numbers as well as information on health insurance, prescriptions and medical history.

T-Mobile
February: An undisclosed number of T-Mobile customers were affected by SIM swap attacks, or SIM hijacking, where scammers take control of and switch phone numbers over to a SIM card they own using social engineering. With access to customer phone numbers, scammers receive messages and calls which allows them to log into the victims' bank accounts to steal money, change account passwords, and even locking the victims out of their own accounts that use two-factor authentication. The attack also exposed customer information including names, addresses, email addresses, account numbers, social security numbers (SSNs), account personal identification numbers (PIN), account security questions and answers, date of birth, plan information, and the number of lines subscribed to their accounts.

Microsoft Exchange
March: Cybercriminals have targeted four security flaws in Microsoft Exchange Server email software. The attackers used the bugs on the Exchange servers to access email accounts of at least 30,000 organizations across the United States, including small businesses, towns, cities and local governments. The cyberattack gives the hackers total remote control over affected systems, allowing for potential data theft and further compromise. Microsoft has released security patches for these bugs and urges customers to apply the updates as soon as possible.

SITA

March: The global IT company, SITA, which supports 90% of the world's airlines confirmed it fell victim to a cyberattack, exposing the PII belonging to an undisclosed number of airline passengers. The stolen information includes names, traveler's service card numbers, and status level.

MultiCare

March: A third-party ransomware attack exposed the personal information of over 200,000 patients, providers and staff of MultiCare Health System, a non-profit health care organization. The attack allowed access to personal information including names, insurance policy numbers, Social Security numbers, dates of birth, bank account numbers, and more.

California State Controller's Office (SCO)

March: A phishing attack targeting the California State Controller's Office (SCO) Unclaimed Property Division led to an employee clicking on a malicious link, logging into a fake website, and granting a hacker access to their email account. The criminal had access to the account for 24 hours, allowing permission to view Personally Identifying Information (PII) contained in Unclaimed Property Holder Reports and to send more phishing emails to the hacked SCO employee's contacts. The number of employees affected and the types of personal information impacted is undisclosed.

Hobby Lobby

March: A database containing records of over 300,000 customers of the arts and crafts chain store, Hobby Lobby, was exposed after the company suffered a cloud-bucket misconfiguration. The disclosed information included customer names, phone numbers, physical and email addresses, and the last four digits of their payment card, as well as the source code for the company's app.

Cancer Treatment Centers of America

March: The Cancer Treatment Centers of America sent out notifications to 104,808 patients, alerting them a compromised email account led to medical information being accessed by an unknown third-party. The compromised account contained patient names, health insurance information, medical record numbers, CTCA account numbers, and limited medical information.

Facebook

April: The personal data of 533 million Facebook users from 106 countries has been posted online for free in a low-level hacking forum. The data was scraped in a vulnerability that the company patched in 2019, and includes users' phone numbers, full names, location, email address, and biographical information.

LinkedIn

April: Over 500 million LinkedIn user profiles were discovered on the Dark Web. The hackers shared two million of these LinkedIn records for only $2 total to prove the legitimacy of the information in the stolen data. The LinkedIn account users' data was scrapped or imported from the website into a database, and includes names, LinkedIn account IDs, email addresses, phone numbers, gender, LinkedIn profile links, connected social media profile links, professional titles, and other work-related personal data.

ClubHouse
April: A database containing 1.3 million scraped Clubhouse user records were leaked for free on a popular hacker forum. The leaked database from the audio chat social network includes user ID, name, photo URL, username, Twitter handle, Instagram handle, number of followers, number of people followed by the user, and account creation date – all of which the company claims is public information.

ParkMobile
April: A third-party software vulnerability is responsible for exposing 21 million customer records belonging to ParkMobile, a contactless payment parking app. The stolen data includes email addresses, phone numbers, license plate numbers, hashed passwords and mailing addresses.

GEICO
April: The auto insurance company Government Employees Insurance Company, known as GEICO, filed a data breach notice announcing information gathered from other sources was used to "obtain unauthorized access to your driver's license number through the online sales system on our website." The total normal of insured drivers affected has not been disclosed but the hackers had accessed between January 21 and March 1. Driver's licenses contain Personally Identifiable Information (PII) such as name, address and date of birth.

Reverb
April: A database containing the personal details of over 5.6 million users of the popular music instruments online marketplace, Reverb, was discovered after it was leaked into the Dark Web. The database contained full names, email addresses, postal addresses, phone numbers, listing/order count, PayPal account email, IP address, and more.

Experian
April: An independent security researcher uncovered a data leak caused by an unsecured Experian application programming interface (API) while researching student loan vendors online. The tool, used by Experian and many other lending sites, allowed anyone to easily access the private credit scores of tens of millions of Americans by supplying their name, date of birth, and mailing address.

CaptureRX
May: CaptureRx, a healthcare system IT company, exposed almost 2 million patient records belonging to over 100 hospitals and healthcare organizations after it was targeted by a

ransomware attack. The sensitive medical information involved in the cyberattack includes names, birthdates and prescription details.

Bailey & Galyen
May: A cyberattack targeting the law offices of Bailey & Galyen exposed the personal information of an undisclosed number of clients and employees. The PII included clients' names, dates of birth, driver's license or personal identification card numbers, Social Security Numbers, payment account numbers, payment card information, biometric data including but not limited to medical information and history, medical diagnosis and treatment information, health insurance information, and other personal information.

Volkswagen & Audi
June: A third-party marketing services supplier disclosed the personal information of 3.3 million customers of Volkswagen and its Audi subsidiary. The exposed data includes their name, mailing address, email address and phone numbers. The data may also include information about a vehicle that has been purchased, leased or inquired about, including vehicle identification numbers, makes, models, years, colors and trim packages.

CVS Health
June: A third-party vendor accidentally posted an unsecured database containing more than a billion search records of CVS Health customers. The 204 GB leaked database was not password protected and included visitor and session IDs, device information, configuration data, as well as multiple records for medications, including COVID-19 vaccines and CVS products.

Carter's
June: The personal and shipping information of over 410,000 customers of the baby clothing retailer, Carter's, were exposed due to a third-party data breach with the company's online purchases software. The information disclosed in the data leak includes names, email addresses, billing addresses, phone numbers, purchasing details, and shipping tracking IDs and links.

Wegmans
June: The U.S. supermarket chain, Wegmans Food Markets, notified an undisclosed number of customers that their data was exposed after two of its cloud-based databases were misconfigured and made publicly accessible online. The personal information in the databases included customer names, addresses, phone numbers, birth dates, Shoppers Club numbers, email addresses, and hashed passwords to Wegmans.com accounts.

Forefront Dermatology
July: U.S. healthcare provider, Forefront Dermatology, announced unauthorized access to its IT systems exposed the personal data and medical records of up to 2.4 million patients. The data exposed included patient names, addresses, dates of birth, patient account numbers, health insurance plan member ID numbers, healthcare provider names, and/or medical and clinical treatment information among other sensitive data.

OneMoreLead
August: An marketing company, OneMoreLead, has exposed the personal records of 126 million individuals through an unsecured database posted online. The database containing names, job titles, email addresses, work email addresses, home device IP address, home address, work address, personal phone number, work phone number, and employer.

SeniorAdvisor
August: Cybersecurity researchers found an unsecured database containing over 3 million personal records of members belonging to a senior living review site, SeniorAdvisor. The database was not password protected and allowed access to information including names, emails, phone numbers and dates contacted.

UNM Health
August: An unauthorized third party gained access to the personal and medical data of over 637,000 patients of UNM Health. The information gathered by the third party includes patient names, addresses, dates of birth, medical record numbers, patient identification numbers, health insurance information, and some clinical information related to the healthcare services provided by UNM Health.

Microsoft Power Apps
August: A misconfiguration within Microsoft Power Apps, a Microsoft product, exposed at least 38 million records. The data leaks impacted American Airlines, Microsoft, J.B. Hunt and governments of Indiana, Maryland and New York City. The disclosed data includes COVID-19 vaccination statuses, social security numbers and email addresses.

GetHealth, FitBit and Apple
September: An unsecured database belonging to GetHealth, a health and wellness data app, exposed over 61 million records of Apple and Fitbit users' data related to fitness trackers and wearables. The database included names, display names, dates of birth, weight, height, genders and geolocations, most were from Fitbit devices and Apple Healthkit.

Neiman Marcus
October: Neiman Marcus announced a data breach that occurred in May 2020. The intrusion was only detected in September 2021 and included the exposure and potential theft of over 3.1 million payment cards belonging to customers, although most are believed to be invalid or expired.

Panasonic
November: The Japanese tech giant revealed a cyberattack had taken place -- a data breach occurring from June 22 to November 3, with discovery on November 11 -- and admitted that information had been accessed on a file server.

Robinhood:

November: Robinhood disclosed a data breach impacting roughly five million users of the trading app. Email addresses, names, phone numbers, and more were accessed via a customer support system.

C) Definitions

Anonymization

The process of data anonymization consists of using AI to identify and alter personal information in a database so that it cannot be identified. Companies can store and transfer data more securely if private elements of the data files have been replaced with pseudonyms, ID Tags, or can be masked. Anonymization can be permanent or reversable using a security key.

Credit monitoring

Credit report monitoring is the monitoring of one's credit activity and credit history in order to detect suspicious, unauthorized activity. Credit monitoring is the act of checking credit reports, and it may also refer to a service that an individual subscribes to in order to monitor such activity. All consumers are advised to monitor their credit to identify and dispute unauthorized activity, such as activity that might indicate identity fraud has occurred. After a data breach, victims whose information was exposed are often offered free credit monitoring services for a certain period of time.

Cybercriminal

A cybercriminal is someone who uses computers to illegally gain access to data that does not belong to them for the purpose of causing harm.

Data breach

A data breach is an incident in which an unauthorized person hacks into a company or other institution's stored data. Hackers breach this sensitive, protected, or confidential data in order to view, steal, and share with or sell it to others. Data breaches may expose records including personally identifiable information (such as Social Security Numbers, dates of birth, email addresses), personal health information, trade secrets, intellectual property or other types of data.

Data security

Data security means protecting data from the unwanted access and actions of unauthorized users.

Encryption

Encryption is one of the most effective way of data security. Files that are encrypted must be decrypted by a secret key or password in order to be read.

Identity fraud

Identity fraud (also known as ID fraud or ID theft) refers to types of crime in which someone wrongfully uses another person's personal data fraudulently or deceptively. Identity fraud is typically used for economic gain by the fraudster.

Identity theft monitoring

Identity theft monitoring (also known as identity theft protection, identity protection, and similar terms), unlike credit card monitoring, refers specifically to the paid subscription services of a company hired to safeguard a consumer from identity theft. ID theft protection services help monitor accounts, place fraud alerts or freezes on your credit reports or remove your name from marketing mailing lists. Many people find it to be a convenient service and worth the expense, and it is often offered for free by breached organizations to affected consumers. It's important to note, however, that most of the services ID theft monitoring companies offer can be done by a consumer on her own for free.

Malware
A combination of the terms "malicious" + "software," malware refers to computer programs that are intended to damage or disable computers and computer systems. Malware can also be used to attack an individual's computer in order to gain unauthorized access to a computer's files. Malware is frequently used to hack into large computer systems.

Personally identifiable information
Personally identifiable information (PII) is a legal term used in U.S. privacy law and information security. PII is information that can be used to identify, contact, or locate a single person, or to identify an individual in context. PII is collected by the companies we have relationships with and by the websites we visit. A good website's privacy policies should specifically address how PII about users is gathered, and lawmakers work to protect our PII. PII is extremely valuable to hackers and those who intend to commit crimes. PII is accessed through data breaches, and a profitable black market exists where PII is collected, shared, and resold. PII can be used to commit identity fraud and other criminal acts.

Phishing
Phishing is a form of fraud in which a victim is tricked into providing sensitive personal information by a criminal posing as a legitimate companies. Victims are often reached by an official-looking email that directs them to click on a link or open an attachment. A common tactic to in phishing emails is asking consumers to click on a (fraudulent) link to confirm account information or other sensitive data. Consumers may also receive phishing phone calls or text messages.

Skimming
Skimming is a type of fraud in which a criminal gains access to the numbers on a legitimate credit card and transfers them onto a duplicate card, which can then be used to illegally make unauthorized charges against the original account (known as "card cloning"). The skimmer does this without the knowledge of the original card holder. In order to "skim" or capture the card information, thieves covertly attach card readers to ATMs, gas pumps, and other places people swipe their credit and debit cards. These readers capture information from a card that is swiped and stored for the criminal to use or sell to others.

References and Resources

IBM's annual Cost of a Data Breach Report, featuring research by the Ponemon Institute, offers insights from 537 real breaches to help you understand cyber risk in a changing world. Now in its 17th year, this report has become a leading benchmark tool, offering IT, risk management and security leaders a lens into factors that can increase or help mitigate the cost of data breaches.

https://fortune.com/2021/10/06/data-breach-2021-2020-total-hacks/

https://www.airiodion.com/privacy-impact-assessment/

https://www.zdnet.com/article/the-biggest-data-breaches-of-2021/

https://www.goanywhere.com/blog/the-10-biggest-data-security-breaches-of-2021

https://www.spirion.com/blog/looking-back-at-the-data-breaches-of-2021/

https://constellaintelligence.com/2021-identity-breach-report/

WebInterpret

Internet use and the pandemic. Pew Research

Social media use in 2021 (Pew)

Covid internet usage and more (Statista)

https://www.upguard.com/blog/biggest-data-breaches