

2024 Identity Breach Report

Welcome to the GenAI Attack Revolution: How AI and Large, Complex Data Sets Are Amplifying the Risk of Identity Exploitation





Table of Contents

About Constella's Identity Breach Report

Executive Summary

2023 Total Metrics

The Rise of AI and Complex Data Sets

- What are GenAI and LLMs?
- How GenAI is impacting the threat landscape
- What is Synthetic Identity Fraud?
- How Cybercriminals Use AI
- Fighting Back Against Dark Web Tools

Geographic Cyber Warfare and Geopolitical Tactics

- History of Russian Hacking Groups
- How Law Enforcement Uses Geographic Data

Infostealers Continue to Steal the Show

- Botnets and Malware
- KYB and KYE

Conclusion and Recommendations

Appendix

About Constella's Identity Breach Report

Harness the Power of The Largest Identity Risk Data Lake In The World



>1T

Assets

Powered by our data lake of greater than one trillion curated and verified assets

Our data spans

125

different countries

53

& languages



Constella's threat intelligence team continually collects identity records from data breaches and leakages across various sources, including open, surface, social, deep, and dark web platforms. This comprehensive data tracking focuses on identifying breaches related to companies and the specific Personally Identifiable Information (PII) exposed, which poses risks to organizations, their employees, and customers. With over 200 billion breach records and more than 1 trillion curated identity assets spanning 125 countries and 53 languages, Constella provides unparalleled insights to help manage these risks.

The rapid advancement of artificial intelligence (AI) and the vast growth of digital identities have significantly reshaped the cybersecurity landscape. Sophisticated AI tools and the proliferation of digital data are leading to more targeted and complex cyber threats.

This report aims to highlight these emerging trends and offer actionable insights to assist organizations in strengthening their cybersecurity measures. By understanding these evolving threats, organizations can better anticipate risks and implement effective strategies to safeguard their data and systems in an increasingly intricate digital environment.

Executive Summary: Breach Report Key Findings

The transformative potential of artificial intelligence (AI) and large, complex data sets marks the next catalyst for the world's digital transformation journey. Both cybersecurity defenders and adversaries alike are beginning to leverage this combined powerhouse on the frontlines, bringing an unprecedented level of complexity and uncertainty to our ever-evolving web of connectivity.

The exponential growth of digital identities transcends anything we've previously encountered. With 98% of Constella's 2023 breach data comprising Personally Identifiable Information (PII) - an alarming 39% jump from last year - an unprecedented amount of sensitive information is now exposed to the world. Each record boasts up to eight identity attributes that form the cornerstone of a structured dataset that is only beginning to be tapped and tamed.

Every year, Constella analyzes global breach data trends. This year's data makes one thing apparent - it is now more critical than ever to level the playing field by embracing innovative approaches and solutions to not only stop threat actors but also to identify and mitigate cybercrime at its very source.

Among other things, good and bad, GenAI is ushering in a new era of sophistication that allows hyper-targeted attacks using tools like FraudGPT and WormGPT, which lower the bar for hackers' skills and cost, while simultaneously raising the bar for minimum security standards that companies need to stay safe.

The bottom line is that datasets and next-gen innovations - the very same ones threat actors are beginning to implement - present a silver lining for those fighting fraud and cybercrime on the frontlines. By facilitating precise identification of bad actors, big data - turbocharged by AI-driven insights - allows organizations to respond quickly, precisely, and decisively to emerging threats. The GenAI Revolution flows both ways, allowing both hackers and defenders to leverage its power on both sides of the breach equation.

This report provides analysis of the most prominent breach trends from this past year, calls out relevant emerging threats, and offers actionable insights organizations can implement to mitigate these prominent and emerging risks.

Top 3 Data Breach Trends

1

REALISTIC, TARGETED AI SCAMS ARE HERE TO STAY

The quantity and availability of PII, coupled with the rise in FraudGPT attack tools means that phishing has entered a new era, where the Nigerian prince email scam has given way to sophisticated and hyper-personalized scams that look and feel real, and are often undetectable to humans due to how the tools leverage natural language processing (NLP) and Large Language Models (LLMs). In fact, Warren Buffett opined, “When you think about the potential for scamming people ... if I was interested in investing in scamming, it’s gonna be the growth industry of all time and it’s enabled, in a way” by AI. These realistic phishing campaigns appear to come from trusted companies, colleagues, friends, and family, and have alarming click-through rates that highlight their effectiveness. According to research conducted by Harvard Business Review, using AI reduces the costs of implementing phishing attacks by “more than 95% while achieving equal or greater success rates” when compared to those generated by humans, suggesting phishing attacks will drastically increase in quality and quantity over the coming years.

2

UNVEILING THE GEOPOLITICAL FACE OF BREACHES

The landscape of cyber threats has morphed into a battleground of nation-state actors, reminiscent of Cold War geopolitical conflicts. The proliferation of sophisticated breaches, particularly orchestrated by state-sponsored entities, has escalated to unprecedented levels. In 2023, Russian breaches accounted for a staggering six out of our top ten domains compromised, signaling a clear and present danger. Accurate source attribution has become imperative in this new era of cyber warfare. Embracing a global data perspective is no longer a choice but a necessity. Not only does it bolster defenses against identity theft, but it also serves as the linchpin for comprehensive investigations. This approach empowers organizations to pinpoint bad actors with unparalleled precision, thwarting fraudulent activities, and unmasking fake identities.

3

INFOSTEALERS CONTINUE TO STEAL THE SHOW

As we noted in last year’s report, infostealers are yet another frontline weapon in the breach war, along with GPT-created malware and phishing scams. In 2023, we detailed the implications of the proliferation of infostealers and botnet malware and what these threats mean for the security of consumers and companies.

Infostealers are a backdoor to the corporate network of any company. This specific kind of malware is remotely controlled by criminals, and is designed to optimize exfiltration and/or steal credentials, messages, documents, and any other data on an infected device. The expansion of these tools continues as more versions are sold on the dark web, enshrining them as one of the most dangerous cyber threats. Last year’s report noted a 140% increase since 2021, and recent data suggests this trend has continued, with another substantial increase in their use in 2023.

100%
increase plain
text passwords

80%
passwords
re-use

152%
increase in
total breaches

39%
increase in
breaches with PII

4

In 2023, we saw **plain text passwords double** in our ingestion. This provides the urgency of real-time alerts across Identity Theft (IDT) and Account Takeover (ATO) channels and enhances our ability to identify criminals through password correlation. 80% of individuals re-use passwords, including criminals! So, while this trend is alarming, it provides immense insights and collaboration across our structured data.

5

2023 data showed a more than **152% increase in total breaches** - from just over 99K in 2022 to more than 151K in 2023. The proliferation and increased sophistication of ransomware attacks, including utilizing double extortion tactics where attackers not only encrypted data but also exfiltrated sensitive information and threatened to release it publicly if the ransom was not paid.

Additionally, the rise of Ransomware-as-a-Service (RaaS) made it easier for less skilled cybercriminals to carry out attacks, significantly increasing the frequency and scale of breaches. This surge in highly effective ransomware campaigns was the primary driver behind the dramatic rise in total breaches.

6

Another concerning stat was a **39% uptick in breaches with PII** from last year to this year. Reasons include:

→ Increased Sophistication of Cyber Attacks

AI and Machine Learning means more targeted and efficient attacks like spear-phishing and credential stuffing, which are highly effective at extracting PII.

→ Expansion of Digital Services and Remote Work

Continued growth of digital services and the persistence of remote work environments expanded the attack surface for cybercriminals. With more employees accessing corporate networks from various locations and devices, the opportunities for gaining access to PII increased.

→ Supply Chain Vulnerabilities

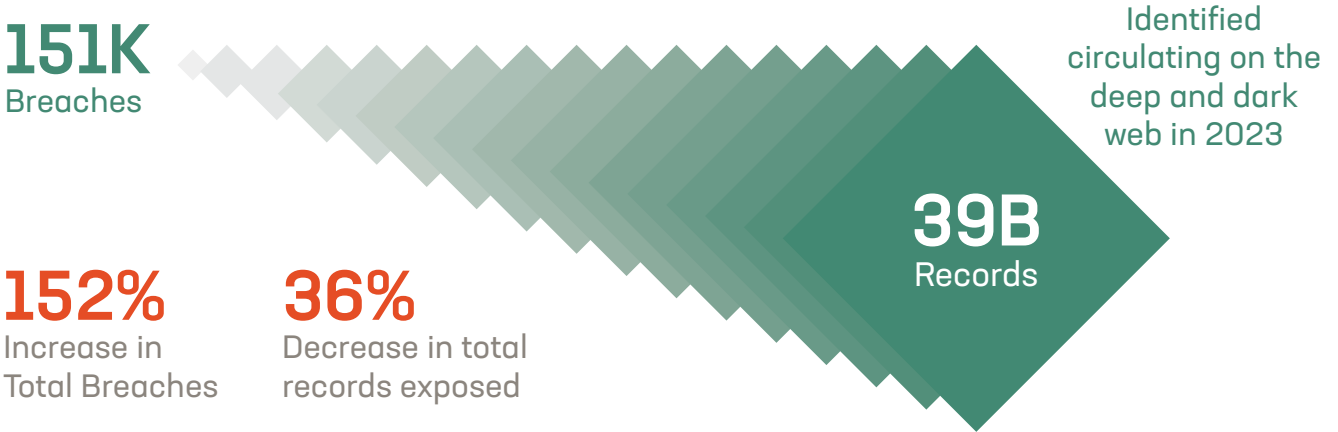
Cybercriminals exploited weaknesses in third-party vendors and service providers to gain access to larger networks, leading to breaches that exposed PII. The interconnected nature of modern supply chains means that a vulnerability in one link can have cascading effects, making it easier for attackers to compromise sensitive data across multiple organizations.

2023 Total Metrics

Most exposed attributes,
sectors, geographies,
and key meta-data

2023 Total Breached Identity Metrics

In 2023, Constella's threat intelligence team identified over 151,000 breaches containing approximately 39 billion total PII records circulating across the deep and dark web.



TYPES OF BREACHES AND EXPOSED RECORDS

97.58%

Breaches with PII

▲ 39% increase

65.40%

Records with email

▲ 4.3% increase

80.69%

Password Combo Breaches

▲ 37% increase

34.60%

Records without email

▼ 4.3% decrease

0.59%

Breaches with only email

▼ 3.9% decrease

0.06%





















Breaches with only email and password

▼ 2.9% decrease

Top 20 Breaches & Leakages

Constella's threat intelligence team compiled a ranking of the breaches and leakages from 2023 that exposed the most records and personal data. These breaches uncovered substantial volumes of personal information. The digital transformation of processes, productivity, and everyday activities increases vulnerabilities by expanding possible sources of breached and stolen data, thereby improving cybercriminals' abilities to execute sophisticated attacks.

BREACHES & LEAKAGES EXPOSING THE GREATEST VOLUME OF RECORDS: 2023

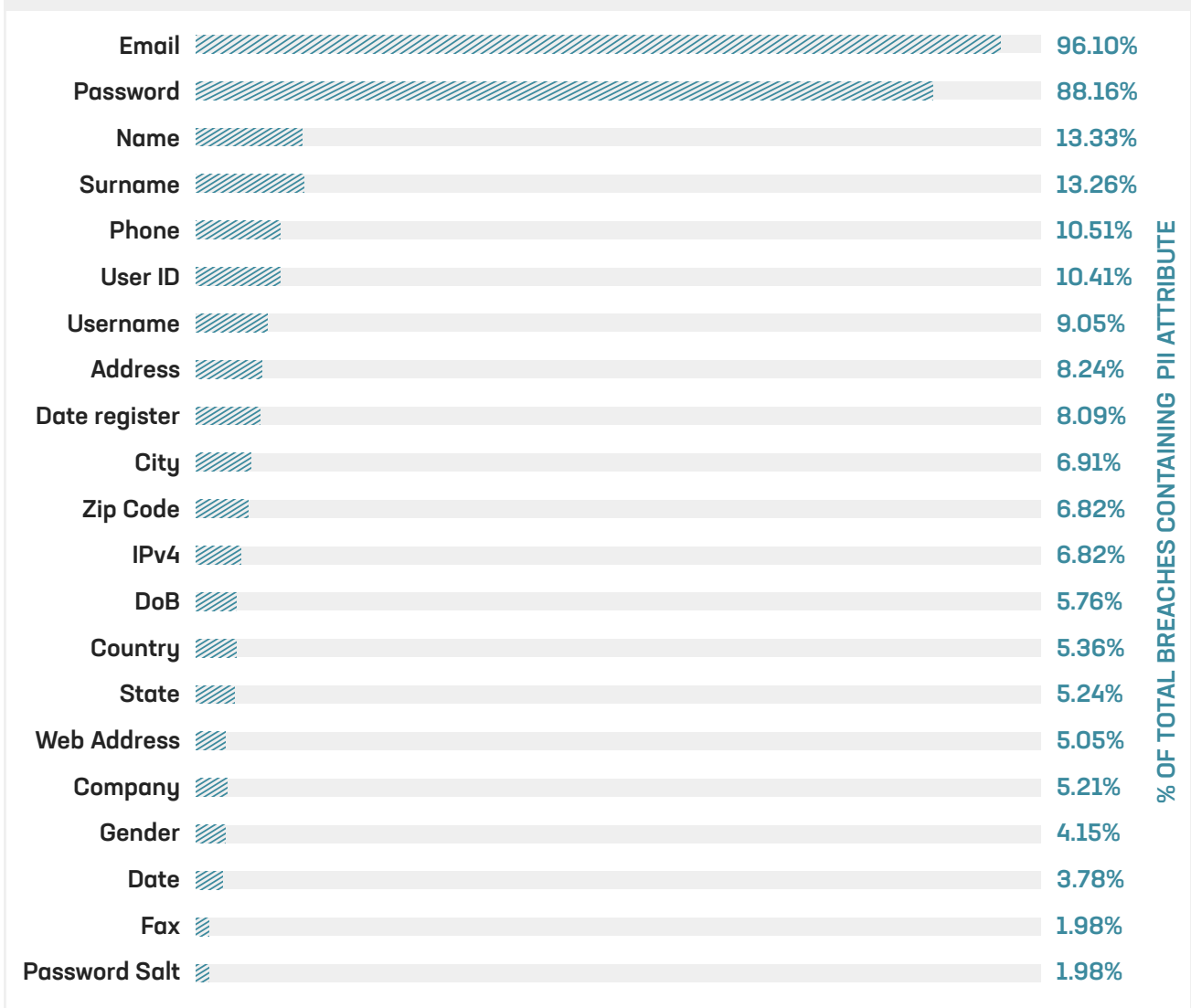
New York-based Real Estate Wealth Network	1.5B		
ICMR	815M		
bvdinfo.com	305M		
deezer.com	245M		
Twitter (200M) accounts	210M		
luxottica.com	206M		
ApexSMS Accounts	67M		
rentomojo.com	64M		
netprospex.com	45M		
Mexico Telecom Database	34M	34M	
eye4fraud.com	29M		
railyatri.in	26M		
vivaair.com	20M		
sunat.gob.pe	15M		
indihome.co.id	14M		
hjedd.com	13M		
yidio.com	12M		
instantcheckmate.com	12M		
paysystem.tech	12M		
lukoil.ru	11M		

Top PII Exposed from Breaches

For 2023, our threat intelligence team identified that emails (96%) and passwords (88%) appear in nearly all breaches and leakages, a significant uptick from 2022 (51.8% and 30.6% respectively). Attributes like names (13%), surnames (13%), and phone numbers (10%) were rarer. Sensitive attributes like address (8%), zip code (7%), date of birth (5%), and company (4%), in addition to sensitive financial information, such as credit card information (<1%) were much less frequently exposed than in 2022.

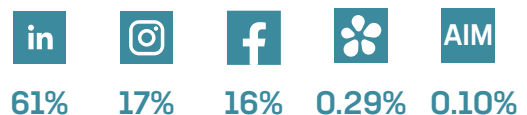
However, given the variety of personal information and details that are revealed and exploited on open sources, as well as the dark web, it is essential to continuously monitor exposed personal data from numerous sources for both personal and corporate security. The increasing number of devices and entry points that can be targeted by malicious actors exacerbates the issue. Consequently, the availability of more exposed personal data provides cybercriminals with valuable tools to carry out attacks.

BREACHES EXPOSING THE GREATEST VOLUME OF RECORDS: 2023



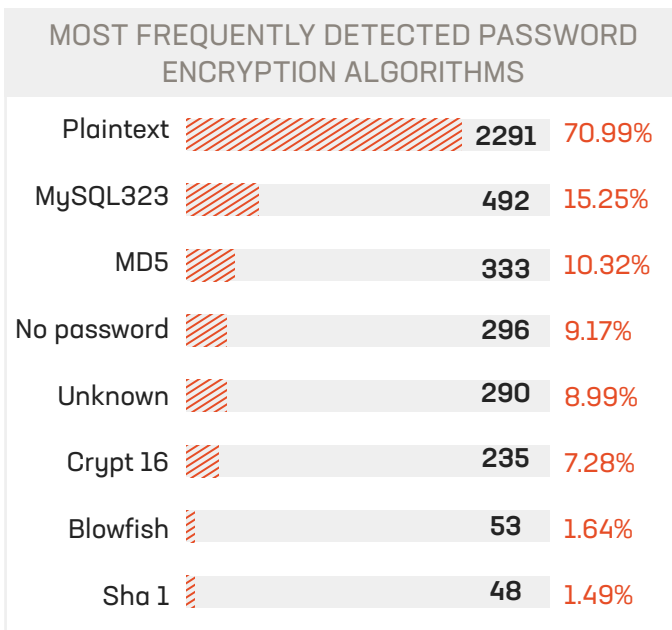
Exposed attributes identified include social media usernames, IDs (user identification numbers specific to a digital platform), and tokens, which can be linked to any identity inside the breach where they are exposed.

From 100% of social media-related attributes



Password Algorithms

The graph below details the most frequently detected password algorithms from the breaches analyzed. Password algorithms determine the complexity and uniqueness of passwords, making them more difficult for cybercriminals to crack using targeted tactics, such as wordlists. A wordlist is an index of possible passwords collected in plain text that can enable threat actors to attempt numerous passwords or variations at scale. Together, passwords in plaintext or encrypted with the MySQL323 or MD5 algorithms were identified in around 88% of breaches.



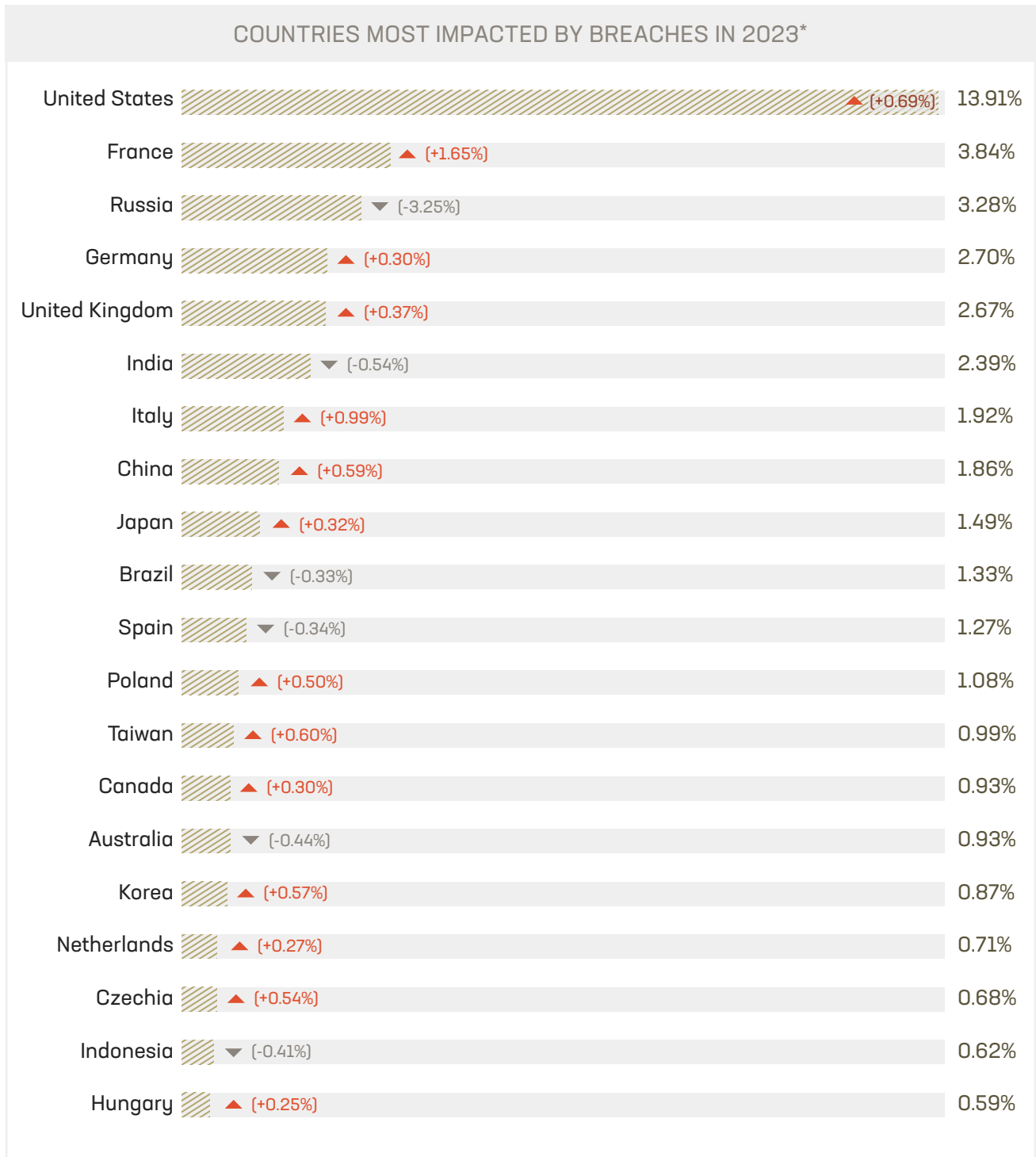
Note: Password encryption algorithms including wordpress, cisco-asa(md5), crc-96(zip), peoplesoft, bcrypt, sha-256, mysql5.x, snefru-256, juniper-netscreen/ssg(screenos), md5-crypt, drupal>v7.x, sha-512, double-sha-1, minecraft(authme-reloaded), dnssec(nsec3), sha-512-crypt, phpbb, md4, snefru-128 and joomla were identified in less than 1% of breaches.

70.99 %
of breaches are plain text or encrypted with weak encryption algorithms

Geographic Distribution

This chart indicates the most affected countries based on the location of the impacted companies and the number of cyber incidents detected in each country. Breaches are classified by the country of origin of the breach.*

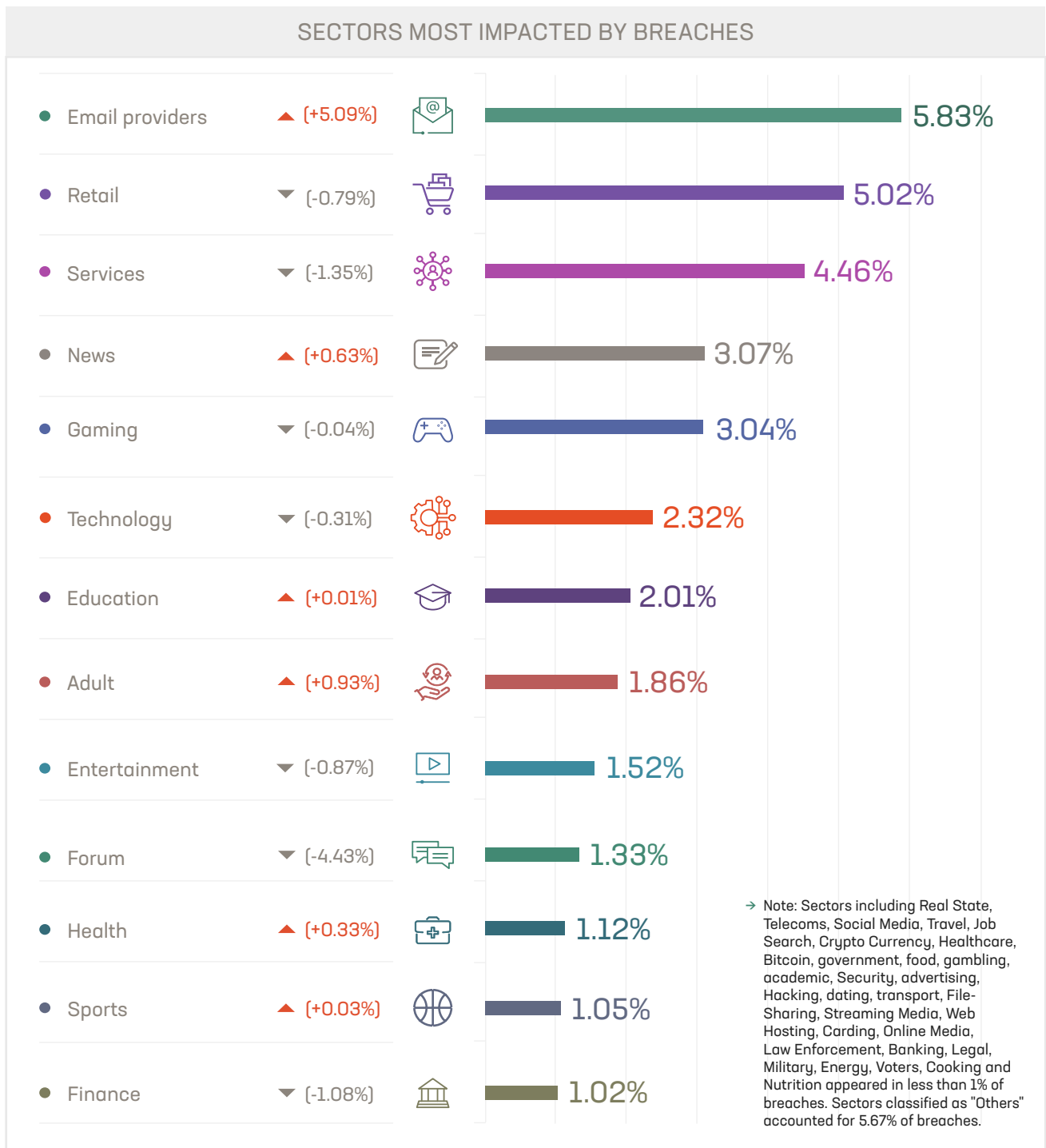
The United States still tops the list of breaches by country this year at 13.9%, a slight increase from 2022. France moved up 4 places to 2nd with a 1.65% increase from 2022, while Russia rounded out the top 3 but saw a relatively significant decrease of 3.25% from the previous year. Germany, Great Britain, and India shifted slightly in order but remained in the middle of the top ten. Australia dropped from 9 to 15, while Italy rose from 14 to 7. Mexico and China also saw dramatic, positive shifts in their numbers, dropping out of the top 20 altogether.



Most Impacted Sectors

Our 2023 top industries impacted by breaches include a new No. 1 - Email Providers, which accounted for nearly 6% of all breaches in 2023, and which wasn't even on the list in 2022. By breaching email providers, attackers were able to intercept, manipulate, or harvest data from many users. Additionally, email accounts often serve as gateways to other online services, making them prime targets for attackers. The ubiquity of email in both personal and business contexts, combined with its potential to unlock access to other systems and services, made email providers a particularly attractive and lucrative target for cybercriminals in 2023.

Other newcomers to the top industries affected were Adult (1.86%) and Health (1.12%). Retail and Service industries fared a little better this year, while Gaming and Technology were nearly identical to last year.

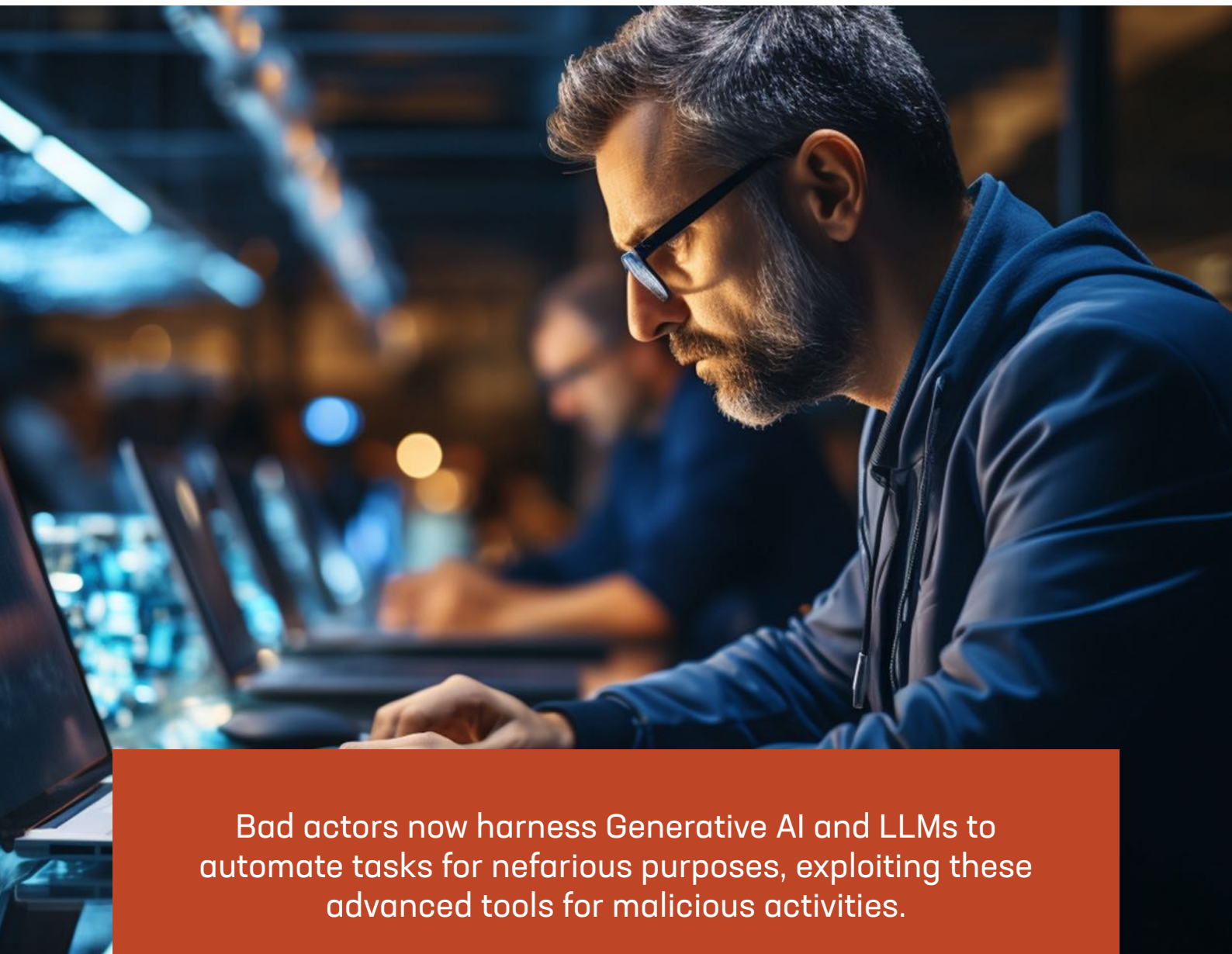


The Rise of AI and Complex Data Sets

What are GenAI and LLMs?

While the term “artificial intelligence” or “AI” has been around for a long time, it's only in the last few years that the term “Generative” has been added to the lexicon. Generative Artificial Intelligence (GenAI) refers to a specific kind of AI that uses the power of Machine Learning (ML) to create new ideas and content from user inquiries. The knowledge that GenAI pulls from is derived from Large Language Models (LLMs) - think ChatGPT - that gather and store massive amounts of information that the model uses to learn language patterns effectively.

Humans are learning how to harness these tools to automate tasks, write more effectively, and make their jobs easier within every industry - including bad actors who now have a tool that does the same things, but for nefarious purposes.



Bad actors now harness Generative AI and LLMs to automate tasks for nefarious purposes, exploiting these advanced tools for malicious activities.

How GenAI Impacts the Threat Landscape

The current threat landscape in cybersecurity is increasingly complex and multifaceted, with a notable rise in the exploitation of personally identifiable information (PII). Cybercriminals are leveraging advanced AI/ML-based techniques to infiltrate networks, employing methods such as phishing, ransomware, and social engineering to gain access to sensitive data. The ubiquity of Internet of Things (IoT) devices and the widespread adoption of remote work have further expanded the attack surface, providing more entry points for malicious actors. As organizations collect and store vast amounts of PII, the potential for data breaches escalates, posing significant risks to both individuals and businesses.

Bad actors range from individual hackers to well-organized cybercrime syndicates and nation-state actors. These perpetrators are highly motivated by financial gain, political objectives, or personal vendettas. They often operate on the dark web, where they can buy, sell, and trade stolen PII with relative anonymity. The consequences of PII theft are severe, including identity theft, financial fraud, and reputational damage. Cybercriminals exploit this information to launch targeted attacks, such as spear-phishing campaigns, which can further compromise sensitive systems and data.

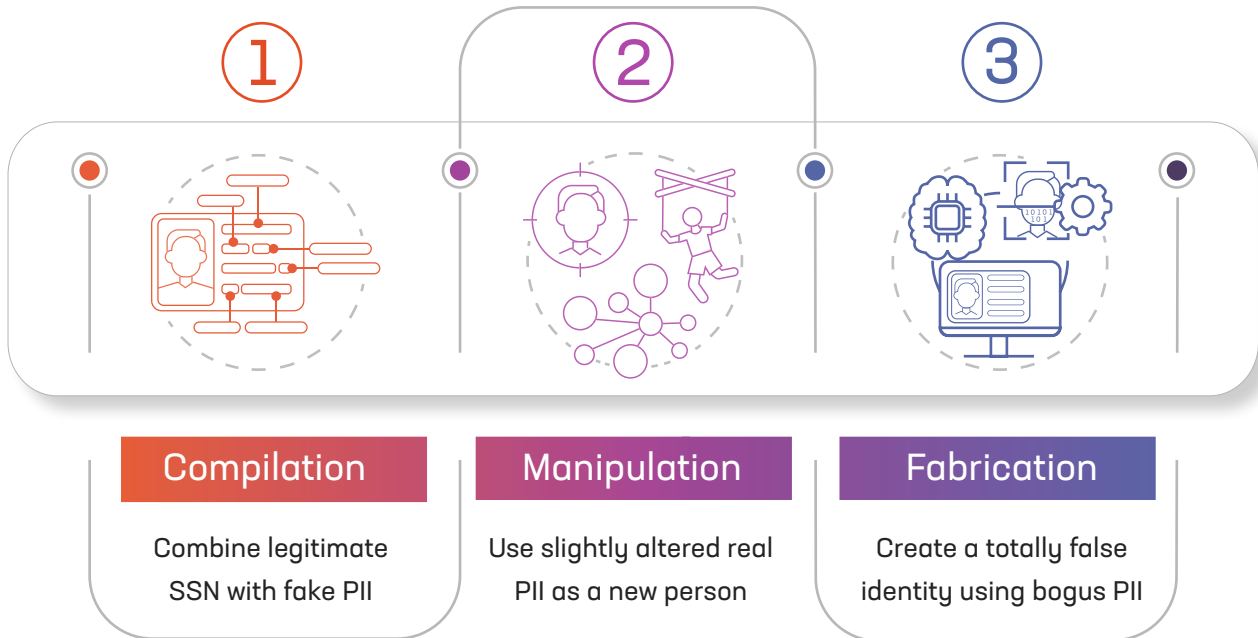


Emerging threat

Synthetic and Fake Identities: Current internal (current/onboarding employees) and external (supply chain, contractors, B2B clients) vetting processes are point-in-time and lack the depth needed to uncover these fraudulent identities as well as risky/criminal behavior.

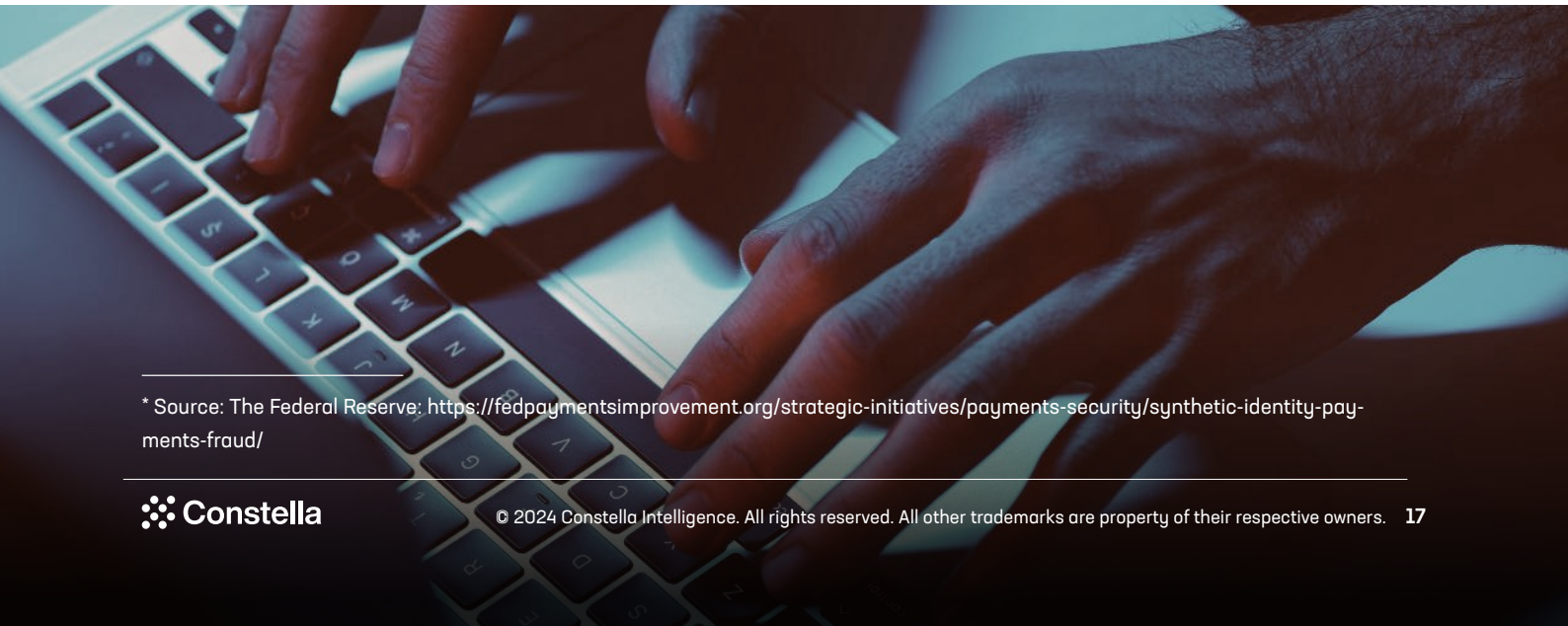
What is synthetic identity fraud?

Synthetic identity fraud (SIF) is defined as “the use of a combination of personally identifiable information (PII) to fabricate a person or entity in order to commit a dishonest act for personal or financial gain”.* Criminals create fictitious identities by combining elements of real and fabricated PII, making it a pervasive and challenging form of fraud. There are three ways synthetic identities can be developed:



These synthetic identities are then used for various fraudulent activities, such as securing fraudulent bank loans, evading law enforcement, or obtaining cash through illicit means. What exacerbates the severity of this type of fraud is its stealthy nature and prolonged execution. Hackers often target individuals with low-activity credit files, exploiting the fact that irregularities in credit reporting may not be immediately detected. The ramifications typically surface when victims attempt significant financial transactions or apply for new accounts, uncovering discrepancies that can be challenging and time-consuming to rectify.

While the use of synthetic identities is most commonly used in financial crimes such as obtaining fraudulent credit cards, bank accounts, or loans, Constella’s data shows an uptick in concern around Know Your Employee (KYE) and even third-party risk management (TPRM). Both employees and third-party vendors often have access to sensitive company and client data. Using appropriate measures to validate both staff and contractors helps lower breach risk. Our report data shows a 40% increase in breaches that include PII - a huge jump that underscores the added exposure that companies face in protecting critical data.



* Source: The Federal Reserve: <https://fedpaymentsimprovement.org/strategic-initiatives/payments-security/synthetic-identity-payments-fraud/>

How cybercriminals are currently using AI to develop more sophisticated attacks

Expanding on the ways cybercriminals are leveraging AI for sophisticated attacks, there are three prominent methods worth exploring in detail.

First, AI is being used to automate the process of creating highly convincing phishing attacks. Traditional phishing relied on generic templates that could be easily recognized by vigilant users or spam filters. However, with AI's launch and ever-increasing usability, cybercriminals can generate phishing emails that are personalized using stolen data and mimic the writing style and familiar language of the target's contacts. This level of customization makes these emails significantly more difficult to identify as duplicitous, increasing the likelihood of successful infiltration into systems or extraction of sensitive information.

Next, AI-generated malware is becoming increasingly popular in use. Malicious AI algorithms can autonomously evolve and adapt to security measures, making them more resilient to traditional cybersecurity defenses. These AI-driven malware variants can learn from their interactions with a target system, adjusting their tactics in real-time to avoid detection and maximize the damage they inflict. This adaptability poses a serious challenge to cybersecurity experts who must continually update their defenses to keep pace with evolving threats.

Cybercriminals are leveraging AI to automate sophisticated phishing attacks, develop adaptable AI-generated malware, and create realistic deepfakes, escalating the threat landscape and necessitating advanced cybersecurity measures.

Last, dark web AI makes it easy to invent deepfake content for malicious purposes. Deepfakes are AI-generated images, videos, or audio recordings that convincingly depict individuals saying or doing things they never actually did. By scraping publicly available data, including social media profiles, existing interviews, or podcasts, cybercriminals can create realistic simulations of corporate executives, political figures, or celebrities making statements that could damage their reputation or manipulate public opinion. These sophisticated manipulations can have profound consequences, including financial fraud, reputational damage, or even political destabilization.

The intersection of AI and cybercrime represents a growing threat landscape that demands proactive measures from organizations and individuals alike. As AI continues to advance, so too will the capabilities of cybercriminals, necessitating ongoing vigilance and innovative cybersecurity strategies to mitigate these emerging risks effectively.

HOW AI IS ENHANCING CYBERCRIMINAL TACTICS

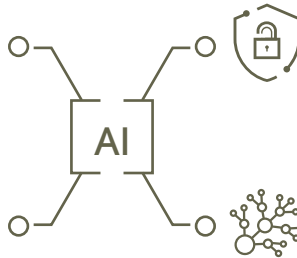
Automated Phishing Attacks

AI automates the creation of highly convincing phishing attacks, making them harder to detect.



Deepfakes and Dark Web AI

AI makes it easy to create realistic deepfakes for malicious purposes, causing potential financial and reputational damage.



AI-Generated Malware

AI malware adapts and evolves, learning from interactions with targets to avoid detection.



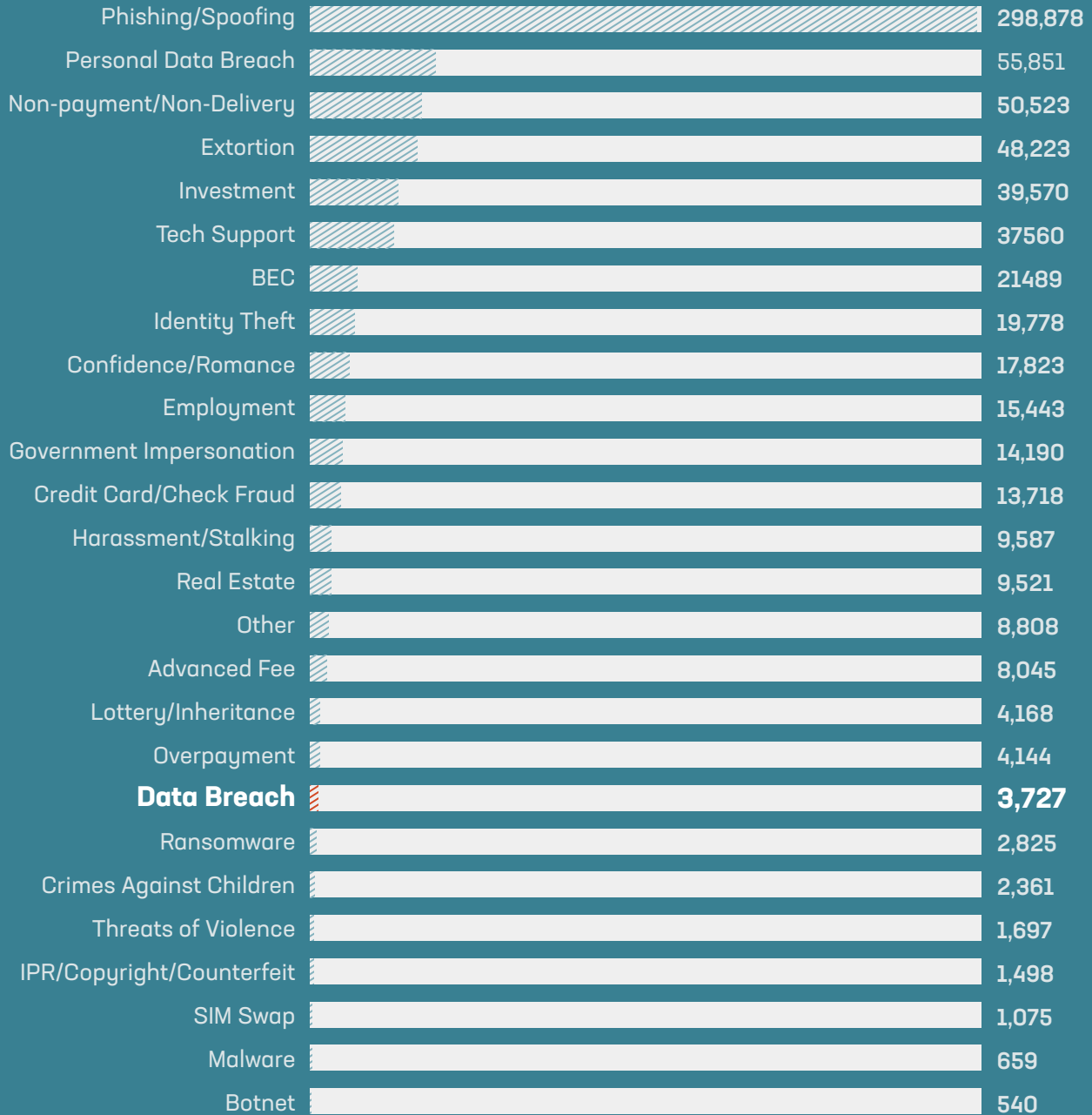
Intersection of AI and Cybercrime

The synergy of AI and cybercrime demands innovative cybersecurity strategies to mitigate emerging risks.

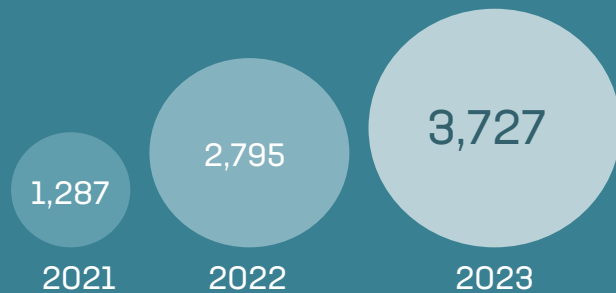


Prevalent types of attacks

2023 Crime types by complaint count*



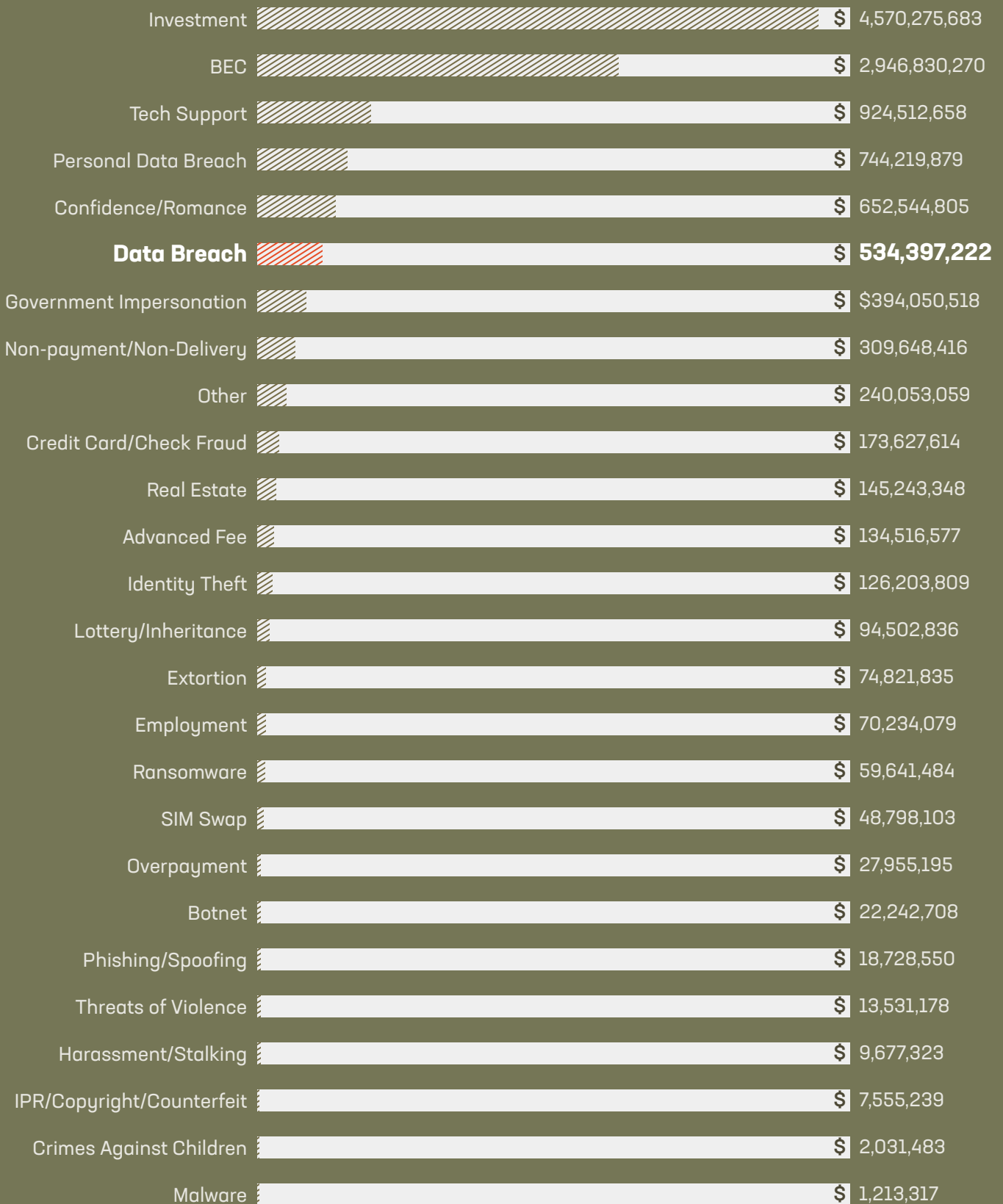
What's more is that data breaches have tripled in just 2 years - from 1,287 in 2021 to 3,727 in 2023.



* Source: Federal Bureau of Investigation: Internet Crime Report 2023 Report: https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf

2023 Crime types by complaint loss

According to the 2023 FBI Cyber Crime Report, data breaches ranked 19/26 in volume of cyber crimes reported but an astounding 6/26 in dollar value of loss - \$534M. *



* Source: Federal Bureau of Investigation: Internet Crime Report 2023 Report: https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf

Why dark web ChatGPT-like programs are causing upheaval among even traditional hackers

Dark web LLM programs, like FraudGPT or WormGPT, are causing a sea change in the hacker community by democratizing advanced due to their ability to democratize advanced cybercrime techniques. These AI-powered tools offer sophisticated functionalities, such as crafting convincing phishing emails, generating malicious code, and automating social engineering attacks with unprecedented efficiency and accuracy. This automation reduces the need for extensive technical expertise, enabling even novice cybercriminals to execute complex attacks that previously required skilled hackers. The accessibility and ease of use of these programs are disrupting the cybercrime landscape, intensifying competition and creating new challenges for traditional hackers who now face a surge in automated, AI-fueled threats.

Emerging threat

The dark web is teeming with new and sophisticated hacker tools that pose significant threats to individuals and organizations alike including:



→ AI-Driven Cyber Attacks

The use of AI in cyber attacks is expected to increase, making it easier for criminals to automate and scale their operations. This includes AI-generated phishing emails, chatbots facilitating ransomware payments, and AI-enhanced malware that can evade detection systems.



→ Deepfake Social Engineering

Deepfake technology, which can manipulate audio and video to create convincing but false representations, is being used for sophisticated social engineering attacks. This technology can be used to impersonate individuals, such as company executives, to gain unauthorized access or manipulate financial transactions.



→ Infostealers and Malware Loaders

Infostealers continue to be a significant threat, used to collect sensitive data from victims. Additionally, "loader" malware services are evolving, allowing threat actors to deploy multiple types of malware, including ransomware and banking trojans, on infected systems.



→ State-Sponsored Attacks and Hacktivism

Geopolitical tensions are driving an increase politically motivated rather than state-sponsored cyber attacks, targeting critical infrastructure and sensitive data. Hacktivist groups are also leverage tensions, often aligning with state interests to disrupt operations and steal information.



→ Ransomware and Extortion

The use of ransomware remains prevalent, with an increasing number of attacks targeting supply chains and exploiting zero-day vulnerabilities. The trend is towards more targeted and destructive attacks, often combined with data theft to increase leverage over victims.

Fighting Back Against Dark Web Tools

2023 FBI stats

880,418
COMPLAINTS

▲ 10%

\$12.5Billion
POTENTIAL LOSSES

▲ 22%

In 2023, the FBI's Internet Crime Complaint Center (IC3) received a record number of complaints from the American public: 880,418 complaints were registered, with potential losses exceeding \$12.5 billion. This is a nearly 10% increase in complaints received, and it represents a 22% increase in losses suffered, compared to 2022.*

Expanding on efforts by organizations and law enforcement to combat dark web tools and cybercrime, significant strides have been made to mitigate the escalating threat landscape highlighted by the 2023 FBI statistics. Law enforcement agencies now recognize the urgency of coordinating actions with private organizations, national and local entities to leverage specialized task forces and technological advancements to disrupt cybercriminal operations wherever possible.

Law enforcement agencies now recognize the urgency of coordinating actions with private organizations.

Established in 2018, the Recovery Asset Team (RAT) streamlines communications with financial institutions and FBI field offices to facilitate the freezing of funds for victims. The RAT functions as a liaison between law enforcement and financial institutions supporting statistical and investigative analysis. In 2023, IC3's RAT initiated the Financial Fraud Kill Chain (FFKC) on 3,008 incidents, with potential losses of \$758.05 million. A monetary hold was placed on \$538.39 million, representing a success rate of 71%.

Moreover, law enforcement agencies have ramped up efforts to dismantle dark web marketplaces and arrest key operatives behind illicit activities. Task forces of cybercrime specialists work across jurisdictions to gather intelligence, conduct undercover operations, and execute strategic takedowns of criminal networks. These operations are often complemented by international partnerships aimed at extraditing cybercriminals to face justice in their respective countries. By cutting the infrastructure and supply chains of dark web tools, law enforcement works to create a hostile environment for cybercriminals, deterring future illicit activities and safeguarding global digital ecosystems.

The concerted efforts of organizations and law enforcement agencies demonstrate a proactive approach in countering the evolving threats posed by dark web tools and cybercrime. Through innovative technologies, collaborative partnerships, and targeted interventions, stakeholders continue to strengthen defenses, mitigate losses, and uphold the integrity of online commerce and communication. As cyber threats persist, ongoing vigilance and adaptive strategies remain paramount to staying ahead of adversaries and preserving public trust in digital platforms.

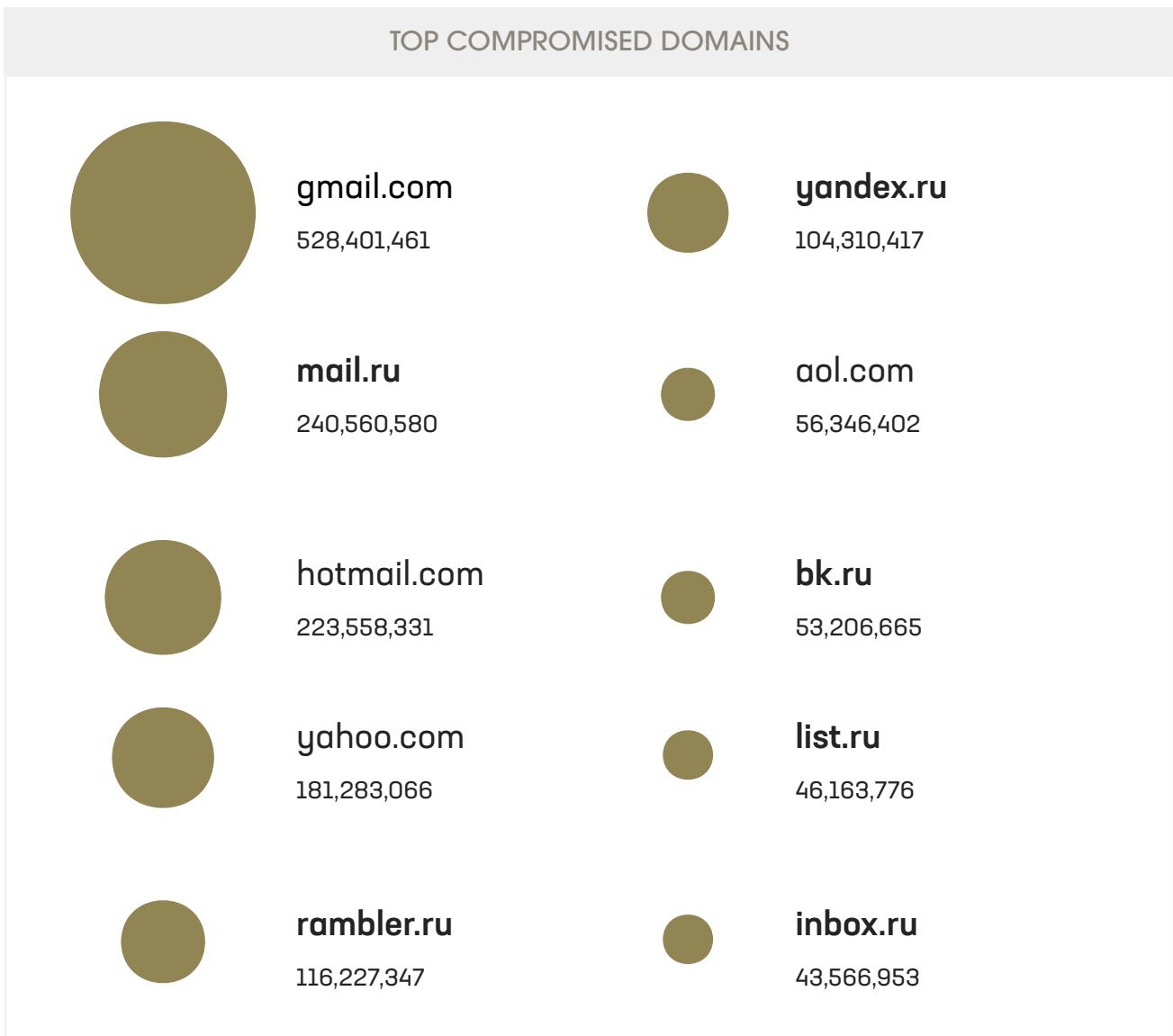
* Source: Federal Bureau of Investigation: Internet Crime Report 2023 Report: https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf

Geographic Cyber Warfare and Geopolitical Tactics

Constella’s data shows a sharp jump in the number of country-specific webmail compromise, particularly in Russia: Six of the top ten domains originated from a .ru domain, and that domain altogether accounts for more than 35% of the total compromise data gathered from the top 50 domain breaches in 2023.

This finding indicates an escalation of continued cyber warfare on a global scale which, when coupled with the prevalence of breaches originating from Russia and other nation states, underscores a pattern that aligns with broader geopolitical tensions. As in years past, cyber operations often serve as an extension of traditional statecraft, providing a means to exert influence and achieve objectives using digital conflict to complement or replace direct military conflict. It’s worth noting that Russia, although an active player, is not the sole player in the escalating arena of cyber warfare.

Ongoing events illustrate that other geopolitical hacking groups are also actively contribute to global digital conflicts, such as the following recent cyber attack on Spain: The Holy League is a large hacker alliance with 70+ affiliated groups that has declared its intent to target global alliances and countries like NATO, Europe, Ukraine, and Israel. The Holy League includes groups like UserSec, NoName057(16), CyberArmy Of Russia, and LulzSec, and has already attacked Spanish infrastructure and government websites.



* Source: Defence Intelligence Agency: https://www.dia.mil/Portals/110/Images/News/Military_Powers_Publications/Russia_Military_Power_Report_2017.pdf

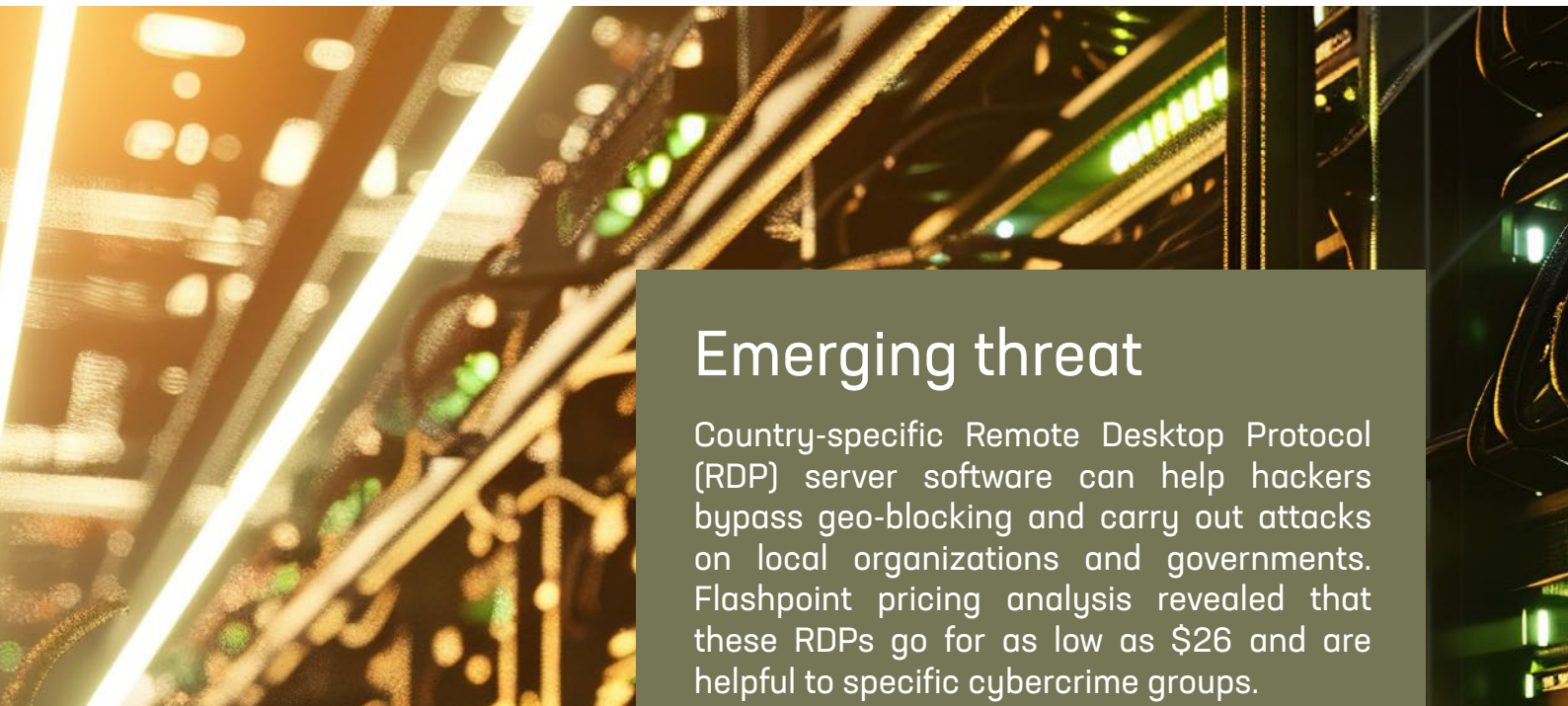
History of Russian hacking groups and the relationship with Russian government

From Cozy Bear and Fancy Bear to Sandworm and Turla, many hacker groups originate and/or operate in Russia, including some that are believed to be linked to the Russian government or other pro-Russian groups. In 2017, the Defense Intelligence Agency explained that Russia used "Information Countermeasures" as "strategically decisive and critically important to control its domestic populace and influence adversary states."* Within 'Information Countermeasures' is a sub-category called "Informational-Technical" that encompasses activities relating to defense, attack, and exploitation of networks. Plainly put, Russia sees hacking activities as inherently useful to influencing adversary states.

How law enforcement uses geographic data to fight cybercrime

Hiding their identity and location is essential for cyber criminals, but advances in GIS, as well as crime-mapping and geographic profiling software are useful. While cyber crime takes place in the cloud, cyber criminals are in a physical location and the vast amount of data now available to law enforcement means that it is possible to utilize geographic profiling to track hackers down to capture and prosecute them.

Law enforcement agencies at the local, state, and federal levels might use geographic data to identify patterns and connections between cybercrime incidents, pinpointing hotspots of activity and potential sources. By analyzing IP addresses, server locations, and other digital footprints, investigators can create detailed maps that highlight areas of interest. This geographic data, when combined with other intelligence sources, enables law enforcement to not only identify and apprehend cyber criminals but also predict and prevent future attacks by understanding the geographic distribution of cyber threats.



Emerging threat

Country-specific Remote Desktop Protocol (RDP) server software can help hackers bypass geo-blocking and carry out attacks on local organizations and governments. Flashpoint pricing analysis revealed that these RDPs go for as low as \$26 and are helpful to specific cybercrime groups.

* Source: Defence Intelligence Agency: https://www.dia.mil/Portals/110/Images/News/Military_Powers_Publications/Russia_Military_Power_Report_2017.pdf

Infostealers Continue to Steal the Show

As we noted in last year's report, infostealers continue to serve as a frontline weapon in the breach war, along with GPT-created malware and phishing scams. The 2023 report detailed the implications of the proliferation of infostealers and botnet malware and what these threats mean for the security of consumers and companies.

Infostealers are a backdoor to the corporate network of any company. This specific kind of malware is remotely controlled by criminals, and is designed to optimize exfiltration, steal credentials, messages, documents, and any other data on an infected device. The proliferation of these tools continues as more versions are sold on the dark web, enshrining them as one of the most dangerous cyber threats. While last year's report noted a 140% increase since 2021, recent data suggests this trend has continued with a substantial rise in the use of these tools in 2023.

The FBI reports that between 2022 and 2023, loss from botnet activity increased by nearly 25%, the highest percentage of increase among types of crime. This rise in crime and dollars lost is consistent with our prediction from last year's report that this type of malicious activity would continue to be a prominent and prevalent issue.

Botnet Malware

Botnet malware is a sophisticated and pervasive threat in the cybersecurity landscape, consisting of various programs designed to hijack multiple computers, connecting them to a network controlled by cybercriminals. Torpig Mebroot is one such notorious example. Torpig is known for its ability to steal sensitive information, such as banking credentials and personal data, by intercepting traffic and logging keystrokes. It operates as a part of the Mebroot rootkit, which is adept at evading detection and removal by embedding itself deeply within the infected system's boot sector. Another significant program is SocGhosh, a malware campaign that utilizes fake software updates to trick users into downloading malicious payloads. SocGhosh is particularly insidious because it often spreads through compromised legitimate websites, making it difficult for users to recognize the threat.

In addition to these, many iterations of Remote Access Trojans (RATs) form a substantial part of botnet malware arsenals. RATs like DarkComet, njRAT, and NanoCore provide attackers with extensive control over the infected machines, allowing them to access files, activate webcams, record keystrokes, and execute commands remotely. These capabilities make RATs highly versatile tools for cyber espionage, data theft, and even launching further attacks from compromised systems. The adaptability and stealth of RATs enable them to be customized for specific targets and continuously updated to evade security measures. Collectively, these programs illustrate the diverse and evolving nature of botnet malware, posing significant challenges to cybersecurity defenses worldwide.

What about Constella's data? Within our data lake, we saw over 500 million total devices infected in 2023 alone, and nearly 2 billion infostealer records inserted in the same timeframe.



KYB and KYE

The intersection of cybercrime with Advanced Machine Learning, fraud, Know Your Employee (KYE) and Know Your Business (KYB) initiatives highlights a critical battleground in the ongoing struggle against illicit activities in both corporate and financial sectors. In particular, KYE and KYB frameworks were originally designed to enhance due diligence processes, ensuring organizations have a thorough understanding of the individuals they employ or conduct business with. However, these initiatives have increasingly become targets for cybercriminals aiming to exploit vulnerabilities in identity verification and data management systems.

Cybercriminals often seek to infiltrate existing protocols to gain access to sensitive personal information or corporate credentials. By leveraging sophisticated phishing attacks, malware, or social engineering tactics, malicious actors attempt to manipulate or compromise these verification processes. For instance, falsified documents or synthetic or fake identities may be used to bypass stringent background checks, enabling criminals to conduct illicit transactions under false pretenses. Similarly, compromised employee verification processes can result in insider threats or unauthorized access to confidential company data, posing significant risks to organizational security and integrity.

To mitigate these risks, organizations are increasingly adopting advanced technologies such as biometric authentication, AI-driven identity verification systems, and blockchain-based data management solutions. These technologies enhance the accuracy and reliability of KYE and KYB processes by reducing human error and enhancing fraud detection capabilities. Furthermore, stringent cybersecurity measures, including encryption protocols, secure data storage practices, and regular security audits, are critical in safeguarding sensitive information from unauthorized access or manipulation.

By integrating robust cybersecurity frameworks with AML, KYE and KYB initiatives, businesses can better protect themselves against evolving cyber threats while maintaining compliance with regulatory requirements aimed at combating financial crime and ensuring data privacy. This proactive approach not only strengthens organizational resilience but also enhances trust and transparency in customer relationships and employee management practices.

Emerging threat

Advanced technologies and techniques enable fraudsters to create highly convincing synthetic or false identities, forged documents, or even AI-generated video and audio clips that fool traditional verification systems. Such techniques pose a significant challenge to the integrity of AML, KYE and KYB processes, potentially allowing criminals to bypass stringent checks and gain unauthorized access to sensitive information or financial resources.



Conclusion and Recommendations

The evolving nature of future attacks. What's next on the cybersecurity frontlines?

1

GENERATIVE AI WILL CONTINUE TO EVOLVE AND BE UTILIZED BY HACKERS AND ALSO BY COMPANIES PROTECTING THEIR DATA.

Recommendation: Security teams should embrace AI in fighting breaches and attacks, staying informed on new AI-enabled Tactics, Techniques, and Procedures (TTPs) used by malicious actors. AI enhances threat detection by analyzing data rapidly to detect anomalies and predict vulnerabilities, empowering proactive defense strengthening. AI-powered tools automate routine security tasks, freeing human teams to focus on sophisticated threats. Solutions like ScamGPT from Constella exemplify this approach, leveraging AI and a vast data lake to identify hyper-targeted scams and educate staff about evolving threats, fortifying defenses against sophisticated attacks.

2

GEOGRAPHY IS AN IMPORTANT DATA POINT, BUT GLOBAL INFLUENCES WILL CONSTANTLY CHANGE.

Recommendation : Understanding the geopolitical context allows security teams to anticipate, identify, and prioritize potential threats that may be targeted based on the organization's location, industry, or political stance. By staying informed about global events and collaborating with law enforcement and intelligence agencies, security teams can access timely and relevant threat intelligence, enabling them to proactively defend against attacks and build resilience against both localized and global cyber threats.

3

INFESTEALERS AND BOTNETS ARE NOT GOING ANYWHERE IN THE THREAT LANDSCAPE.

Recommendation: Implement a multi-layered security approach that combines advanced threat detection technologies with robust employee training and awareness programs. This strategy should include the deployment of comprehensive endpoint protection solutions that can identify and mitigate malware in real-time, regular software updates and patch management to close vulnerabilities, and network segmentation to limit the spread of infections. Additionally, prioritize employee education to recognize phishing and other social engineering tactics. Foster a culture of security awareness and leverage cutting-edge defense mechanisms to reduce the risk and impact of this pervasive malware.



4

KYE AND KYB PROGRAMS HAVE BECOME A MAJOR POINT OF VULNERABILITY THAT HACKERS ARE INCREASINGLY LEVERAGING TO GAIN ACCESS TO AND EXPLOIT PII.

Recommendation:As technology evolves, organizations must prioritize robust AI-driven fraud detection tools and enhanced verification methods to combat the growing risk of identity fraud. This includes advanced authentication mechanisms like identity risk scoring and comprehensive vetting beyond basic checks. Regular audits and updates of KYE and KYB processes are essential to adapt to evolving threats. In addition, the adoption of advanced approaches such as Constella's Advanced KYE and KYB solutions, which utilize AI profiling paired with the world's largest data lake, can help to further streamline compliance, mitigate risks, and detect synthetic identities and fraud effectively -- safeguarding data integrity and protecting against emerging threats.

5

WHEN IT COMES TO PASSWORDS AND OTHER SECURITY MEASURES, HUMANS REMAIN THE WEAKEST LINK IN CYBERSECURITY, WITH 80% OF INDIVIDUALS RE-USING PW'S, INCLUDING CRIMINALS.

Recommendation: Despite advances in technology, individuals often undermine security protocols through poor practices such as using weak or easily guessable passwords or reusing passwords across multiple accounts. Hackers, too, sometimes fail to secure their own systems against counterattacks. Therefore, continuous education and awareness programs are essential to reinforce the importance of robust security practices among all users, emphasizing that cybersecurity is as much about human vigilance and behavior as it is about technological defenses.

Appendix



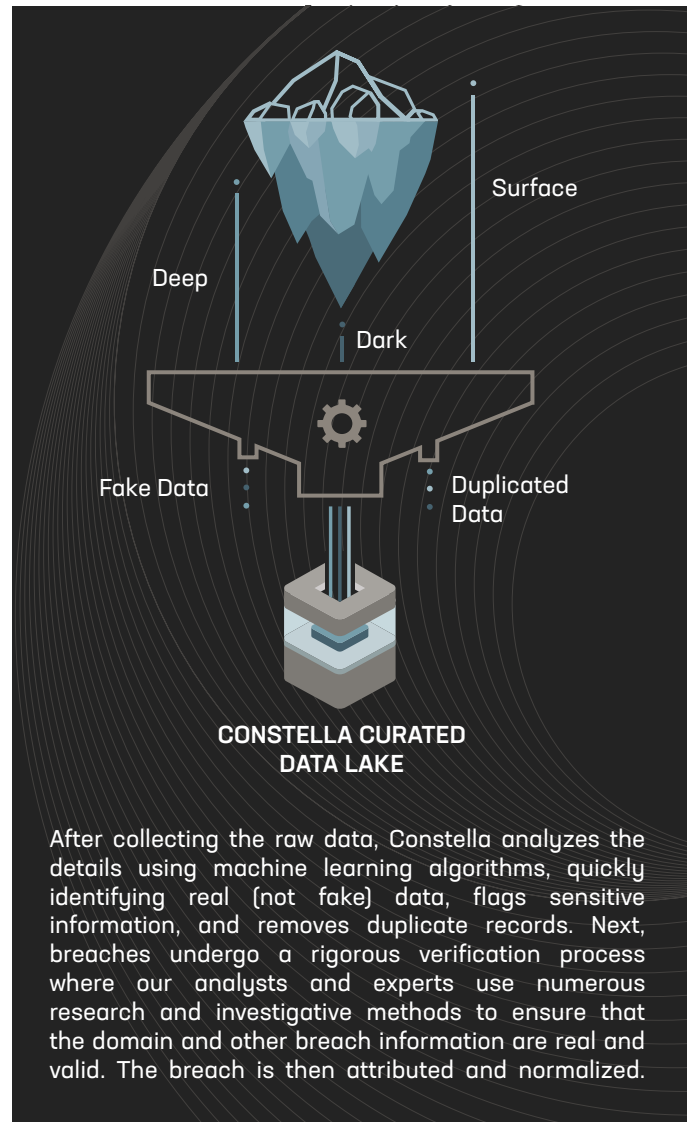
Data Sources

Constella has monitored the TTPs of threat actors closely and developed this report based on breaches and leakages identified in 2023. In addition to the known breaches and leakages reported in the media, Constella detects information found in data dumps posted in the open, but often transient, sources in the deep and dark web. Constella's automated crawlers and subject matter experts use a variety of sources to authenticate and verify the data, including:

- **Underground communities and forums**
- **Black markets**
- **The deep web**
- **The dark web**

Constella analyzes, verifies, cleans, and attributes the data to further understand the severity of risks facing consumers and companies. Constella then alerts the impacted parties to mitigate risks. We assess the severity of risk based on multiple factors, including:

- **Sensitivity of information**
- **Authenticity of the data**
- **Number of individuals impacted**
- **Age of each type of sensitive identity attribute exposed**



Data Verification/Methodology

While the number of accumulated raw identity records provides insight into the exposure of activity of identity-based data, it is not the best indicator of overall risk.

This is because not all of the data gathered is authentic or unique. After collecting the raw data, Constella analyzes the details using machine learning algorithms, quickly identifying real (not fake) data, flags sensitive information, and removes duplicate records.

Next, breaches undergo a verification process where our analysts and experts use numerous research and investigative methods to ensure that the domain and other breach information are real and valid. The breach is then attributed and normalized.

After a breach is verified, the Constella platform calculates a risk score based on several variables, including types of attributes, date, and password strength.

About Constella Intelligence

Constella.ai is the global leader in AI-driven identity risk and deep and dark web intelligence for such applications as identity theft, insider risk, Know Your Employee (KYE), Know Your Business (KYB), and deep OSINT investigations. With the world's largest breach database, containing over one trillion data attributes in 125+ countries and over 53 languages, Constella empowers leading organizations across the globe to monitor and secure critical data through unparalleled visibility and actionable insights. Ready for a secure future? Reach out to Constella today and stay one step ahead of digital threats.

Stay ahead of identity threats with AI-driven insights and unmatched intelligence.



www.constella.ai

contact email : constella@constellaintelligence.com

© 2024 Constella Intelligence. All rights reserved. All other trademarks are property of their respective owners.