Annual Identity Exposure Report
2022

SpyCloud
Table of Contents

Overview 03
2021 Highlights 04
2021 Identity Exposure Trends 05
  The “New Normal” 05
Credential Exposure on the Rise 07
Stolen Credentials Add Ammunition to Ransomware Attacks 09
Remote Work Plays Into Fraud Schemes 10
Malware, the Riskiest Threat 11
The Growing Bounty of PII 13
Top 12 Notable Breaches of 2021 14
Protecting Your Organization from Digital Identity Exposure 16
What’s Next 16
About SpyCloud 17
Overview

The SpyCloud Identity Exposure Report examines the trends related to exposed data that puts consumers and organizations at risk of online fraud, account takeover, and follow-on cyberattacks including ransomware. Every year, our researchers analyze recaptured credentials and personally identifiable information (PII) that cybercriminals have exploited over the previous year, and we explore the implications of these findings.

2021 carried over some of the same themes we noted in 2020, such as threat actors taking advantage of people working and shopping from home. Our discoveries also reflect trends that escalated last year, including ransomware, identity fraud, and credential exposure.

Our overall observation echoes past findings: the amount of compromised credentials and PII that SpyCloud recaptured has grown significantly year over year. For example, the 1.7 billion exposed credential pairs that we recaptured in 2021 represent a 15% increase from 2020. With consumers relying on digital identities now more than ever, cybercriminals have a lot more opportunities to profit from stolen identity data. Couple that with the higher password reuse rates we observed in 2021, and you can see why leveling the playing field with fraudsters gets more difficult every year.

We’ve entered an era where consumers prefer to conduct business digitally, whether that’s shopping, banking, or applying for services. And, as the outbreak of criminal online activities indicates, cybercriminals have pivoted right along with consumers. The number of data compromises hit a new all-time high in 2021, growing 68% over 2020, according to the annual report from the Identity Theft Resource Center. The report also noted that 2021 was a milestone year “when cybercriminals shifted from mass data accumulation (identity theft) to mass data misuse (identity fraud).”

Because SpyCloud collects data months earlier in the breach lifecycle than what may be publicly reported, we can provide unique insights into how these shifts impact enterprises – and help them secure their employee and customer accounts before the exposed digital identities are used to cause harm.

About SpyCloud Data

SpyCloud’s proprietary engine collects, curates, enriches, and analyzes recaptured data from breaches, malware-infected devices, and other sources from the criminal underground – transforming it into actionable insights that enable enterprises to quickly identify legitimate users vs. potential criminals using stolen information, and take action to prevent account takeover, ransomware, and online fraud.

Learn more at spycloud.com.
2021 Highlights

Last Year, SpyCloud Recaptured...

1.7 BILLION CREDENTIALS

Total Breach Sources
755

Average Breach Size
6,736,241 Records

13.7 BILLION PII ASSETS WERE RECAPTURED

70% of users exposed in 2021 breaches were reusing previously compromised passwords.

- 1.6 BILLION Phone Numbers
- 1.2 BILLION Social Media Handles
- 170 MILLION National IDs & SSNs
- 8.3 MILLION Credit Card Numbers
- 5.9 MILLION Bank Account Numbers
- 1.6 BILLION Phone Numbers
- 1.2 BILLION Social Media Handles

Government Highlights

- 561,000 .gov Credentials Recaptured
- 60% Password Reuse Rate
- #1 Reused Plaintext Password: password

Topical Highlights

- 1.7M ACCOUNTS had '2021' in their password
- 113,467 PASSWORDS were influenced by popular sports
The “New Normal”

The “new normal” became the refrain during the COVID-19 pandemic, as consumers and businesses alike adapted to tremendous disruption. But working and shopping from home wasn’t the only norm in 2021. The digital landscape, exposed more than ever, provided a lush terrain for cybercriminals. As digital identities increased, fraudsters exploited opportunities to their fullest.

Building off the seismic shift to digital transactions we saw in 2020, the digital revolution continued its momentum in 2021, even as in-person interaction returned. In the first quarter of 2021, for example, online banking comprised 96% of financial institutions’ activity, and accounted for 93% of all fraud attempts. The costs of fighting fraud rose accordingly, and for every $1 of fraud, U.S. financial services spent $4 in 2021, compared to $3.64 in 2020 (and $3.25 in 2019).
Ecommerce traffic also continued its upward trajectory, spiking 51% in the first half of 2021 (compared to the same period the previous year) – and the sector reported a 140% increase in the 2021 volume of fraud attacks. Likewise, fighting fraud got more expensive for ecommerce merchants, rising from $3.36 in 2020 to $3.60 for every $1 of fraud.

These trends indicate we’re moving forward in the digital age. For consumers, this means a higher reliance on online identities. And they’re struggling to keep up with their expanding number of passwords (as many as 100, according to some research). The SpyCloud data reflects this pain point, with more users opting for shortcuts like reused passwords. We discovered a 64% password reuse rate for users with more than one password exposed in the last year. This is a 4-point jump from last year’s report, which showed a 60% reuse rate.

Our data also tells us what to expect in the fraud landscape in 2022 and beyond. Reused passwords have been the leading vector in cyberattacks in the last few years. Users’ growing propensity to recycle their passwords, especially as they spend more time online, will further improve the cybercriminals’ odds of successful attacks.

The amount of PII freely available on the criminal underground is also fueling new fraud activity. In 2021, SpyCloud recaptured 13,789,875,451 PII assets from the criminal underground – sources including data breaches and bot logs from malware-infected devices. Cybercriminals use this stolen PII to create synthetic identities – and the broad types of PII we discovered is proof that they can draw from a vast and constantly growing collection of data to perpetrate identity fraud.

Passwords Culled Straight from Headlines

We know that cybercriminals take advantage of current events to wreak havoc. But users are just as inspired by headlines as threat actors. We were curious what was on people’s minds, so we searched our 2021 recaptured credentials for the year’s most popular words in pop culture and news. We expected, of course, to find a fair number of keywords from popular 2020 passwords including covid, coronavirus, and mask. As we anticipated, “2021” was the most popular thematic keyword, just as “2020” was the year before.

Some of the other top pop culture keywords we noted – like loki, falcon, and wanda/wandavision – may be a bit more surprising. That is, unless you’re a Marvel enthusiast, which means you know that the media franchise had a blockbuster year. Marvel released numerous TV shows and movies in 2021, and the superheroes made their way not only into the fans’ hearts but also into their passwords. We surmised that some of the other top choices may be due to consumers being starved for entertainment, with their longing reflected in popular password keywords like atlantabraves/braves, britneyspears/britney/freebritney, and dune.
Credential Exposure on the Rise

In the cybersecurity world, the conversation often revolves around the increasing sophistication and evolution of threat actors’ tactics. That may be true, but it doesn’t take much prowess to log into an account with a stolen password. Cybercriminals don’t need complicated techniques when they have your login credentials — and 89% of web application breaches involve credential abuse. That’s why compromised credentials are threat actors’ most treasured type of data.

SpyCloud researchers recaptured more than 1.7 billion exposed credential pairs – combinations of email address and password or username and password – from 755 breach sources over the course of 2021. This degree of exposure represents a nearly 15% increase from the previous year’s 1.48 billion.

Each of those exposed credential pairs put your customer or employee accounts at risk. Worse yet, given the astonishing password reuse rate, one set of exposed credentials is likely giving cybercriminals access to more than one account. And the credential exposure we observed in 2021 is only the tip of the iceberg. To date, our database includes more than 25 billion total recaptured passwords.

As we noted earlier, we found a 64% password reuse rate for users with more than one password exposed in the last year. But worse, for users we can tie to breach exposures in 2021 and prior years (with the same email address or username exposed), 70% were still reusing the same exposed passwords. For example, if “password123” was exposed for bob.smith@example.com in 2018 and then we collected the same password associated with bob.smith@example.com in 2021, then we consider that reuse in this 70% statistic.

Across our entire database for breaches we’ve recaptured since 2016, we’ve calculated that 60% of users have reused passwords, up from 57% last year.
Why are we seeing an increase in password reuse, despite widespread education on cyber hygiene from businesses’ cybersecurity awareness programs, vendors, and the media? Still only 20% of people rely on password managers, and 66% of Americans are not required to use a password manager at work. With more than 100 logins to remember and users admitting to relying mostly on memory, the unfortunate result is an increase in reuse. According to our analysis, more than 82% of reused passwords are an exact match (not even a number or special character added).

Many organizations, as well as individual consumers, think they’re protected if they rely on a dark web monitoring service. Unfortunately, it can take a year or longer before compromised credentials make it to the dark web. By the time you receive that notification, the damage may be done.

When cybercriminals harvest fresh credentials, they hold that data close, sharing it only within a select group of trusted associates. This tightly knit network seeks to monetize the credentials as quickly as possible, either launching high-value, highly targeted attacks, or perhaps selling them to malware operators. To avoid detection, humans drive these targeted attacks (instead of bots). And because you or your employee or customer aren’t likely to be aware yet of the credential compromise, the attackers are all but assured access.

Only after many months of exploitation will they finally leak the credentials on a mass scale — selling them at bargain prices as commodity combo lists. In the meantime, the threat to your organization is massive.

To help enterprises keep up with the cybercriminals, SpyCloud quickly recaptures fresh data from the criminal underground by using human intelligence researchers, supplemented by technology. We work fast to cleanse the data, add context, determine if our customers’ information was involved, and provide fast insights and automated remediation. This speed to action enables organizations to respond when they’re at the highest risk and minimize the impact.
A Closer Look at Government Risk

The government sector continued to have a tough time in 2021. Many U.S. government agencies were still reeling from the massive SolarWinds attack when the president issued an executive order in May aimed at strengthening national cybersecurity, with a special focus on the software supply chain. A bipartisan report released in August by the U.S. Senate’s Homeland Security and Governmental Affairs Committee didn’t deliver any better news — concluding that the eight agencies reviewed had “significant cybersecurity weaknesses” and earned an average grade of C- for their overall information security maturity.

Threat actors also continued to pummel government agencies with malware and ransomware. The Cybersecurity Infrastructure Security Agency (CISA) reported that the evolution of ransomware tactics in 2021 demonstrated the “ransomware threat actors’ growing technological sophistication and an increased ransomware threat to organizations globally.” CISA also noted that authorities in several countries observed an escalation in sophisticated, high-impact ransomware incidents against critical infrastructure organizations, and 14 out of 16 critical infrastructure sectors were hit in the U.S.

In light of these developments, SpyCloud researchers wanted to learn how government agencies fared in data breaches in 2021. We found 611 breaches containing .gov email addresses, which is 81% of the overall total breach sources recaptured by SpyCloud. In total, we identified 561,753 credential pairs (email addresses with plaintext passwords) from government agencies around the world.

These exposed credentials give threat actors a potential foothold inside of those agencies — and this risk is compounded given the high password reuse rate.

Last year’s reuse rate was 60% among .gov users with more than one password in our database where at least one of those passwords was collected in 2021.

Stolen Credentials Add Ammunition to Ransomware Attacks

Ransomware, too, became the “new normal” and a big headline in 2021. In a SpyCloud survey of IT security professionals, 72% reported that their organization was affected by ransomware in the previous 12 months. Nearly one-fifth said they experienced six or more ransomware incidents during that time.

The respondents also identified weak or exposed credentials as the second riskiest entry point, following phishing emails. Not surprisingly, 79% agreed that news coverage of ransomware attacks significantly elevated their organization’s concern about customers and employees using weak or stolen credentials.

Ransomware operators commonly use compromised credentials to gain initial entry into an organization and to escalate privileges. And while it may only take 20 minutes for them to execute an attack, the median dwell time for ransomware is five days.
The earlier in the lifecycle that you can disrupt ransomware, the more successful you’ll be in preventing threat actors from lurking inside your network, undetected.

Last year's Colonial Pipeline ransomware attack was an eye-opening example of how a single password can not only bring down a company and cost millions, but also create deep ripple effects. The fact that researchers found the exposed password on the dark web further drives home the point that all it takes is one instance of poor employee cyber hygiene like password reuse to shake a company to the core.

As noted earlier, compromised credentials pose the highest risk early in their lifecycle, and this holds true for ransomware. An entire class of threat actors, called initial access brokers, specialize in selling access into target organizations to ransomware gangs. This access often comes in the form of stolen logins, with the access brokers using scripts to discover vulnerabilities in backend servers and scraping databases that contain credential pairs.

Remote Work Plays Into Fraud Schemes

After scrambling to quickly transition to a remote workplace in 2020, organizations had months to adapt to the new cybersecurity environment. Despite that adjustment period, IT and security teams continued to struggle with this “new normal” last year. Only 43% of security pros surveyed in April 2021 were confident that their organization was prepared to respond to a data breach caused by their remote workforce. Nearly half also believed the breaches they experienced were the result of remote work.

By August, 50% of cybersecurity professionals felt increased job stress due to remote worker support. With their poor security habits, employees didn't make life easy for already thinly stretched security teams.

A common habit, for example, was to log into a corporate virtual private network (VPN) from a home computer and save the VPN password in the browser. Lacking robust security, the family computer would get infected with malware like RedLine Stealer (more on this below), enabling the threat actor to harvest the VPN login and then access corporate assets. SpyCloud researchers continued to find evidence of employees conducting work and personal activities on the same device, with our data showing corporate logins mixed in with personal ones.

Even if your employees only log into work accounts from a corporate device, their poor security hygiene still puts them – and your company – at risk. Let’s say they’re saving corporate passwords in the browser of their work device, then syncing that browser to a personal device for convenience. If their personal device is infected by malware, the cybercriminal will exfiltrate all their data, including the synced corporate passwords. This type of problem will get harder to solve, especially if security and convenience remain at odds in your workplace.
Malware, the Riskiest Threat

When looking at digital identity exposure trends, you can’t recognize the full picture of your risk if the data sources don’t contain information exposed through malware infections.

The role of malware is commonly overlooked in identity risk solutions, yet malware is responsible for the fraud that’s hardest to detect, and thus poses the highest exposure severity for both consumers and enterprises.

Individuals may download malware on their local device by clicking on a malicious link or downloading an executable file that masquerades as something benign, like a free game, app, or invoice. Cybercriminals can help themselves to a vast variety of data on the infected machine because the malware executes actions such as creating a backdoor into the system and logging keyboard strokes. After establishing a command-and-control (C2) connection with the threat actor’s server, the compromised device transmits logs with details ranging from login credentials and browser history to geolocation, installed software, autofill info, and web session cookies.

Malware infections create an extreme risk of online fraud and identity theft in several ways:

- The adversary can use the freshly harvested credentials immediately for account takeover (ATO). The success rate for ATO is much higher than for credential stuffing or other types of password attacks since the bad actor has the exact password for specific websites. Also, even if the victim changes passwords, the bad actor will get the updated information for as long as the device remains infected.

- The siphoned data includes specific details that establish what’s known as a browser or device fingerprint (a combination of operating system, IP address, browser type, system fonts, browser extensions, bookmarks, and other details). Since companies often use browser fingerprints behind the scenes to seamlessly authenticate customers, cybercriminals can successfully impersonate the consumer without raising any red flags; the fake user will look indistinguishable from the authentic one. And neither the company nor the user has any idea about the exposure until they’re defrauded.

- Many websites offer users the option to save their session (as a browser cookie) for convenience — it allows them to skip the login or shortcut the authentication process next time. Because cookies are part of the malware-siphoned data, the criminal can use them to bypass logins altogether.

- The PII that the malware collects can be used to impersonate a victim completely, fooling not only security and anti-fraud solutions but humans as well. Threat actors use PII for crafting social engineering scams or creating synthetic identities to apply for loans or credit cards, for example.
Cybercriminals also monetize the malware-collected data by selling the logs on the criminal underground. We regularly see advertisements on popular underground forums from people looking to either buy or sell logs with specific companies’ accounts. In November and December, we tracked prices for these logs as low as $130.

And finally, malware is extremely dangerous when a threat actor is targeting specific victims. We noted earlier that exposed credentials are the most valuable when they’re freshly harvested and that cybercriminals closely guard those logins at first to launch targeted attacks, such as ATOs of highly valuable targets. Malware is an effective tactic for this purpose because it’s difficult to detect and it captures passwords even after they’ve been reset. Undiscovered bad actors using accurate logins can exploit a victim and remain stealthy for a long time with a high rate of success.

SpyCloud data includes records from malware-infected devices, including compromised web session cookies. In 2021, we noticed a surge in infostealer (information-stealing malware) logs being distributed and shared on various forums and chat groups. In particular, RedLine Stealer accounted for more than 50% of all infections that we analyzed, followed closely by Racoon, Vidar, and a handful of other malware families. To help enterprises reduce fraud risks and protect their corporate resources, we have sorted and parsed hundreds of thousands of post-infection bot logs resulting in hundreds of millions of stolen credential records over the last 12 months.

By recapturing data and surfacing it to our customers quickly, we’re also enabling them to devalue it faster and contain the potential damage. For example, customers can check their users against a continuously updated feed of compromised session cookies, protecting accounts before fraudsters can impersonate users with stolen browser fingerprints.

### How It Works: A Closer Look at RedLine Stealer Malware

Over the course of 2021, SpyCloud researchers noted that RedLine Stealer was one of the most widely used infostealers for Windows. Currently, the malware is available for purchase in the underground for around $800, or as a malware-as-a-service subscription for $200 a month.

RedLine Stealer is typically distributed through phishing campaigns, malicious search results for free or cracked software, and comment links in online videos. This malware steals cryptocurrency wallets and data from installed software like Discord and Steam, as well as from browsers. Once the stolen data – stored passwords, cookies, credit cards, browser autofill details, etc. – is at the cybercriminals’ fingertips, they can easily search and export it through a simple web interface.

In the last few years, there has been a rise in anti-detect browsers, which allow web surfers to create separate browsing environments with different browser fingerprints. These browsers are useful for legitimate purposes, such as managing multiple social media accounts or browsing privately. But in the hands of bad actors using malware like RedLine Stealer, these browsers serve to impersonate users. The fraudsters simply log into the browser with the malware-stolen credentials and bypass multi-factor authentication while the session or device cookies are still fresh.

RedLine Stealer is an especially tricky malware to fight – security researchers found it can masquerade as legitimate downloads for software, such as Windows updates.
The Growing Bounty of PII

Stolen PII is the fraudsters’ bread and butter. They need it to perpetrate fraud by creating synthetic identities or to impersonate authentic consumers and open new accounts. Every year, the vast quantity of PII available on the criminal underground grows more expansive. And we’ve heard from customers throughout this last year that because so many new accounts were created as people shifted to online transactions, distinguishing legitimate new customers from fraudulent ones often became overwhelming.

The sheer variety of exposed PII is astounding. In addition to the more common types of data such as names, dates of birth, and national identification numbers or driver’s licenses, we saw everything from vehicle makes and models, number of children, and smoker status, to marital status, estimated income, job title, and even Reddit handle. In other words, just about everything a criminal would want for stealing an identity.

In 2021, SpyCloud recaptured 13.8 billion PII assets, a 200% increase from 4.6 billion the year before. This brings the total in our database to 44.7 billion pieces of PII.

The highest numbers of PII assets were in the categories that criminals need for creating synthetic identities, opening new accounts, and perpetrating other forms of online fraud. These categories included full name, email address, phone number, date of birth, IP address, Social Security or other national identity number, and driver’s license, as well as credit card and bank numbers.

Unfortunately, the PII exposed in data breaches every year fuels new criminal activity as attackers also use it for phishing and social engineering schemes, leading to new data breaches – and fresh batches of compromised PII.
Top 12 Notable Breaches of 2021

The data breaches that make headlines typically impact large numbers of consumers or involve government agencies, large enterprises, and popular brands. Capturing a lot of attention last year were brands like Facebook (now Meta), LinkedIn, T-Mobile, and Android, where each of the associated data leaks affected hundreds of millions of consumers. But flying under the radar every year are countless other breaches, too.

SpyCloud’s focus is to recapture the data as early as possible, but we don’t disclose some breaches externally until they’re old news in order to support responsible disclosure. We also classify as sensitive any data obtained from potentially controversial sites, such as dating services – particularly when those breaches don’t validate employee email addresses – to avoid tarnishing employees’ reputation.

Excluding sensitive breach sources and combo lists, here are some of the most notable breaches of 2021.

**JANUARY**

ShinyHunters Strikes Again
Records Leaked: **129.4 Million**

After a busy 2020, ShinyHunters returned for an encore. This time, the leak comprised 10 databases from sources such as the popular image-editing cloud app Pixlr, India’s largest cryptocurrency exchange BuyUcoin, and B2B music-streaming site tunedglobal.com.

**FEBRUARY**

TicketCounter
Records Leaked: **5,531,555**

The European e-ticketing platform’s user database was stolen from an unsecured server used for staging. Affected PII included names, email addresses, hashed passwords, and phone numbers.

**MARCH**

Park Mobile
Records Leaked: **26,090,473**

The popular mobile parking app was allegedly breached, with user data including passwords and PII made available for sale underground.

**APRIL**

BigBasket
Records Leaked: **20,774,436**

In another 2021 leak from the actor ShinyHunters, the compromised database contained IP addresses, credential pairs, and other data. BigBasket, an Indian online grocery delivery service, had disclosed in 2020 that it suffered a data breach.
Facebook wasn't breached, but scraped user profiles — with details like Facebook IDs, dates of birth, emails, and phone numbers — from more than 100 countries were posted on a hacking forum and have been made available since then on several others.

A vendor left the automaker’s customer data unsecured on the web. Compromised records, spanning five years, included PII and sensitive data such as loan eligibility.

A hacker stole a database containing names, addresses, phone numbers, and email addresses from this UK weapons marketplace and published it on the dark web. The link to the download has since been shared on a public website.

Hacktivists breached the domain registrar and exposed a decade’s worth of email addresses, passwords, usernames, and other personal information. The group shared the entire 180GB dataset on underground forums.

A part of the scheduling platform’s database was breached, exposing subscribers’ PII such as names, email addresses, and phone numbers. The data was being traded on an underground forum.
Protecting Your Organization from Digital Identity Exposure

The 2021 findings from our researchers prove that, every year, risk from digital identity exposure grows by leaps and bounds. The 1.7 billion exposed credentials, the 64% password reuse rate, the 13.8 billion recaptured PII records — all these data points and others show an increase from previous years. This trend hasn’t changed, even as the fraud themes vary from one year to the next.

Compromised credentials, coupled with the high password reuse rate, create a significant risk for your organization. The best way to safeguard your company, customers, and employees is by protecting users from themselves. This can be as simple as monitoring for exposed credentials and resetting them as quickly as possible after a data breach or malware infection.

It’s not impossible to break bad cyber hygiene habits, but it’s certainly an uphill battle. A more effective way to minimize your exposure is by using data from the criminal underground against the adversaries. This data adds an advanced layer of protection by helping you understand your riskiest users (such as those with stolen device fingerprints due to a malware infection), recognize legitimate customers from fake ones, and take corrective action based on your fraud mitigation policies.

What’s Next

SpyCloud’s annual research confirmed that digital identity exposure is a growing problem for consumers and enterprises. The intertwining of personal and work lives, along with the expanding digital footprint, will continue to accelerate the rates of online fraud. Yet the nature of criminal activity is constantly shifting, making it much tougher to separate legitimate customers and activity from fraudulent ones. This reality will endure for the modern enterprise.

Fraud prevention solutions are numerous, but even the most sophisticated ones don’t always keep up with the fast pace of change in the digital fraud landscape. Find a partner who keeps you ahead of the threat actors — not in perpetual reaction mode.

SpyCloud’s unique combination of human intelligence, technology, and breadth of recaptured data allows you to proactively stop fraud before it occurs. We give you the confidence to make risk mitigation decisions quickly and at scale — and protect your enterprise from account takeover, ransomware, and online fraud.
About SpyCloud

SpyCloud transforms recaptured data to protect businesses from cyberattacks. Our products leverage a proprietary engine that collects, curates, enriches, and analyzes data from the criminal underground, driving action so enterprises can proactively prevent account takeover and ransomware, and protect their business and consumers from online fraud. Our unique data from breaches, malware-infected devices, and other underground sources also powers many popular dark web monitoring and identity theft protection offerings.

SpyCloud customers include half of the 10 largest global enterprises, midsize companies, and government agencies around the world. Headquartered in Austin, Texas, SpyCloud is home to over 150 cybersecurity experts who aim to make the internet a safer place.

Enterprise Protection

Prevent account takeover that can lead to ransomware.

Learn More

Consumer Protection

Combat account takeover and online fraud.

Learn More

Investigations

Unmask criminals attempting to harm your business.

Learn More

Data Partnerships

Enhance your solution with SpyCloud’s data.

Learn More

Learn more at spyccloud.com
SpyCloud