# **Spy**Cloud

2021 Annual Credential Exposure Report

## **Spy**Cloud

### 2021 Annual Credential Exposure Report

**Overview** 

<u>Trends</u>

Credential Exposure

Password Hashing

Personally Identifiable Information

Top 10 Breaches of 2020

**Government Credential Exposure** 

Your Plan of Action

## **Overview**

Every year, the SpyCloud Credential Exposure Report has examined the data cybercriminals shared over the previous year and explored what it meant for enterprises and consumers.

2020 wasn't a typical year. The global COVID-19 pandemic forced a heightened awareness of ever-present threats, opened countless new doors for criminals, and sent the security community scrambling to catch up and make sense of it all. And yet, many things remained the same. According to our findings, rampant password reuse continues to be a problem, leaving enterprises and their customers at risk of account takeover (ATO).

Criminals wasted no time preying on our collective vulnerability in 2020. As early as March, there was an <u>onslaught of activity</u> that leveraged the coronavirus to manipulate users through various threat types, from phishing campaigns impersonating public health officials to scams promising immunity. And just as the virus fueled criminal activity, it also drove the fundamental shift to remote working. Practically overnight, many businesses had to rapidly change how employees used technology. As a result, initiatives that normally would have spanned years were implemented in haste, leaving weaknesses exposed. Meanwhile, users spent more time online, created more online accounts, and blended more boundaries between work, personal, and family online activity, expanding the criminal playing field.

Throughout this unusual time, SpyCloud's researchers have been embedded in criminal networks, using human intelligence (HUMINT) to recover stolen data before it reaches a broader criminal audience or goes public. Because we collect data early in the breach life cycle, we can help enterprises secure their employee and consumer accounts before the most dangerous forms account takeover begin. As a result of this work, the data we've collected over the last year provides unique insight into breaches and botnet logs that have been released to criminal communities throughout the last year. In 2020 alone, our team recovered **nearly 1.5 billion stolen credentials and operationalized them to protect hundreds of enterprises and over 2 billion consumers** from account takeover and online fraud. As this data gradually becomes accessible to a wider criminal audience, it will be used for increasing numbers of account takeover and follow-on attacks over the months and years to come.

#### About SpyCloud Data

**Truly Actionable Breach Data to Prevent Account Takeover** SpyCloud uses Human Intelligence (HUMINT) to quickly recover breach data, often within days of the breach occurring. Our unique data cleansing and password cracking process reveals compromised credentials faster and with greater match rates. Access to this massive breach database enables enterprises to quickly identify and take action on exposed accounts, preventing those exposures from progressing to account takeovers. SpyCloud safeguards more than 2 billion employee and consumer accounts from account takeover and followon attacks like credit card fraud, phishing, and ransomware. **Learn more at spycloud.com.** 



## 2020 Highlights







465 of the breaches had .GOV emails

The average breach included **5,455,813 RECORDS** 



# **1.5B** TOTAL CREDENTIALS RECOVERED

**193,073** Passwords included pandemic keywords



Of these credentials, **60%** of passwords were reused

**1.6** Passwords contained the number **2020** 

TOTAL PII ASSETS RECOVERED IN 2020 Among these recovered PII assets, we found:

**1.2B** PHONE NUMBERS

**70M** ACCOUNT SECRET

**THE CREDIT CARD NUMBERS** 

## ACROSS SPYCLOUD'S ENTIRE DATABASE

**100+B** TOTAL RECOVERED BREACH ASSETS

TOTAL RECOVERED CREDENTIALS

23+B

Among 57% Pa

.....

Among them, we found a **57%** Password Reuse Rate

••••• 🚺

· · · · · · · [f]

## **Credential Exposure Trends**

#### Retail fraud cut into margins

From ordering household necessities to splurging on retail therapy from online boutiques, the pandemic accelerated ecommerce to <u>triple digit</u> percentage growth. Cybercriminals not only pounced on this sea change in consumerism, they also preyed on our collective vulnerability with COVID-19-related malware distribution scams promising virus cures and low-cost PPE.

In response to COVID restrictions, many brick-and-mortar retailers leaned heavily on buy-online-pickup-in-store (BOPIS) programs with great success. BOPIS saw a <u>62%</u> year-over-year increase in activity from February 24 to March 21, 2020 alone. But for every good intention there is someone out there eager to exploit it; in recent years, BOPIS-related fraud has jumped <u>250%</u>. The rise of BOPIS provides the perfect cover for criminals to monetize stolen accounts.

Over the last year, SpyCloud recovered nearly 1.5 billion credentials from the criminal underground – data that bad actors are actively using to take over users' accounts and commit online fraud. Meanwhile, password reuse rates have not declined; SpyCloud observed a **60% password reuse rate** for users with more than one password exposed in the last year, matching last year's rate exactly.



#### Remote life blurred personal and professional boundaries

When the world shifted into lockdown mode, criminals were more than ready. The rest of us,

not so much. Practically overnight, people were forced to work, learn, shop, socialize and more, online. This opened up a whole <u>new set of security challenges</u> for IT teams, many of whom lacked the experience, protocols and technologies to enable a remote workforce securely. For employees, the sudden shift to remote life has introduced new accounts to keep track of, and blurred the boundaries between work and personal browsing. SpyCloud spotted evidence of remote work's effects on security hygiene in a surprising place: botnet logs. Devices infected with credential-stealing malware can capture users' every online move and send the data to attackers, who often share those logs with other criminals. Bad actors can use the stolen data to spoof victims' devices, answer account security questions, <u>bypass</u> <u>multi-factor authentication</u>, and steal their identities, putting these users at exceptionally high risk of hard-to-detect account takeover and online fraud.

Last year, we noticed an uptick in the overlap between personal and corporate data collected in botnet logs, showing that people are increasingly using personal devices for work and corporate devices for play. This is bad news for corporate IT teams, who can monitor the security of employees' work-managed devices but have no visibility into personal systems. If an employee logs into corporate resources using an infected device, attackers can easily access enterprise resources while evading detection.

#### "Superbreaches" give old data a facelift

In November, 23,600 hacked databases were leaked from a defunct "data breach index" called Cit0day, a popular service for leaked data (names, emails, usernames, addresses, and plaintext passwords) on the dark web. Much like the <u>Collection Combolists</u> that went public in 2019, Cit0day represents a compilation of many older breaches packaged together into a single "superbreach," significant not because it exposes new data, but rather because of how much easier it makes it for criminals to use that stolen information for credential stuffing attacks.

The Cit0day leak included as many as 226 million usernames and passwords, although affected users have had a hard time finding enough information about how they were exposed to do anything about it. Several services could tell you that your credentials were compromised somewhere within the Cit0day breach, but couldn't tell you which of the 23,600 databases your credentials were found in and, by extension, which exposed passwords you needed to change. At SpyCloud, we matched the collection against our own database to understand the original sources of the breach data, deduplicate them, and make the data more actionable for our customers. As expected, much of the data was already included in our database. For the purposes of this report, we've counted them as a single breach.

This trend of repackaging old data into massive combolists and releasing them as newsworthy superbreaches will certainly continue, as the <u>COMB combolist</u> of early 2021 already demonstrated. For enterprises, the Cit0day breach and others like it serve as a reminder that stolen data sticks around and remains useful to cybercriminals for many years after the original breach.

#### Credential stuffing is the new data breach

Considering 2020's unusual circumstances, it was perhaps inevitable that businesses would face heightened security threats. Zoom, Nintendo, Activision, The North Face and other brands made headlines as hundreds of thousands of consumers' accounts landed in the hands of bad actors, exposing sensitive information such as purchase history, billing and shipping addresses, names, birthdays, telephone numbers, rewards point balances, and email addresses. However, despite the way many media outlets described the attacks, these weren't traditional data breaches. Instead, they represent a growing trend of credential stuffing at scale being categorized as data breaches in mainstream media.

For enterprises, the distinction between whether consumer accounts have been accessed due to a breach of internal resources or via credential stuffing is critical. A breach results from a company's failure to protect its assets and often has regulatory implications, whereas consumer account takeover is typically the result of consumers' bad password hygiene. For consumers and media outlets, that distinction is becoming less important.

The transformation of credential stuffing's media image begs the question: exactly how much responsibility do enterprises share for users' password choices? Only time can tell whether changes in public perception will influence the way regulatory authorities handle consumer account takeovers. At a minimum, monitoring consumer logins for weak and stolen credentials serves as reputation mitigation, helping enterprises avoid ending up in the news for the wrong reasons.

#### All eyes on the supply chain

2020 ended with the revelation of the largest supply chain attack we've ever seen, affecting over 17,000 enterprises and government agencies. Attackers used SolarWinds' update servers to deliver a trojan that FireEye researchers have dubbed SUNBURST, providing attackers with an entry point into the networks of major customers like the U.S. Department of Homeland Security, U.S. Treasury Department, and Microsoft. Some have estimated that recovery costs for affected customers will surpass <u>\$100 billion</u>.

Poor password security played an important role in the attack. For starters, SolarWinds' update servers were secured using a password format our data shows is all too common: solarwinds123. (A recent SpyCloud analysis of Fortune 1000 employee data revealed that 6 of the 10 most popular Aerospace & Defense sector passwords include company names.) In addition, compromised credentials from multiple employees enabled attackers to access network resources and extend their foothold. Attackers also bypassed popular multi-factor authentication software, serving as a reminder that while MFA provides an important layer of account protection, it's not foolproof.

The SolarWinds attack was by far the largest supply chain attack of 2020, but it certainly <u>wasn't the only one</u>. Going forward, we expect criminals to continue to target the supply chain, and compromised credentials will surely come into play again.



## **2020 Credential Exposure in Review**

Criminals don't need to use sophisticated technologies to breach firewalls or other security measures intended to protect the enterprise. They just need your password.



Account takeover represents a significant problem for enterprises and consumers, and the needle is moving in the wrong direction. In 2019, ATO was the top fraud method for financial institutions, with a <u>72%</u> year-over-year increase in financial account takeovers alone. In 2020, with COVID-19 disrupting our world, the yearover-year growth was startling — <u>nearly 300%</u>. With more stolen credentials available to criminals all the time and no change in consumers' bad password habits, the problem is only going to get worse.

**Over the course of 2020, SpyCloud researchers recovered 1,486,416,779 stolen credentials from 854 breach sources.** Each of these credentials could be used to gain illegitimate access to consumer and employee accounts.

Worse, these credentials represent only a fraction of the problem: SpyCloud's database contains over 23 billion passwords alone that have been recovered from data breaches and botnet logs, with billions added every year. While fresh credentials are the most valuable to criminals and pose the greatest threat to potential victims, stolen passwords stick around for many years.



#### **3%** MORE BREACH SOURCES IN 2020

Over the last year, SpyCloud recovered data from over 33% more breach sources than in 2019: 854 sources, versus 2019's 640 sources. Stolen credentials are most dangerous to individuals and enterprises in the 18 to 24 months after a breach occurs, while they're likely to be limited to a small circle of criminals and monetized using creative, targeted methods. When SpyCloud researchers recover fresh data, we quickly cleanse that data, add context, crack passwords so we can determine if our

customers' information was contained in the breach, and get the data into their hands so exposed passwords can be reset before they are exploited.



We focus on speed at every step so enterprises can outpace cybercriminals, as well as automated remediation so there is no heavy lifting to keep users safe from ATO and online fraud.

#### 2020 Recovered Breaches by Total Records



Over the last year, SpyCloud recovered data from 33% more breach sources than in 2019: 854 sources, versus 2019's 640 sources.

For this total, we've counted the Cit0day breach we described earlier as a single breach, even though it contains thousands of separate databases that are individually tracked within our database. Rather than ingesting the Cit0day data as one large combolist, making it difficult for affected users to figure out what sites were breached and when, our team cracked hashed passwords and mapped the data to individual breaches, including many that were already in our database.

Counting the Cit0day databases as a single breach, we found an average breach size of 5,455,813 records across the data we collected in 2020. Only 200 of these breaches contained between 1 million and 50 million records; only 16 contained more than 50 million records. Factoring the individual Cit0day breaches into the calculation, which includes thousands of breaches with fewer than 500,000 records, that overall average drops to just 235,573.

#### **Password Reuse**

Despite years of advice about the importance of strong passwords, people inevitably end up reusing or recycling the same credentials for multiple sites. Outdated password complexity requirements have complicated the issue by providing people with a false sense of security when they recycle a favorite password with a few simple changes, like adding a 1 or ! at the end (an especially when prompted to change passwords every 90 days by corporate IT).

Unfortunately, reused passwords provide little protection. When login details from one site are exposed in a data breach, cybercriminals can access any other accounts that are protected by the same credentials.

To understand the prevalence of password reuse in the data SpyCloud collected last year, we looked at the habits of users who had at least two passwords exposed during 2020. We found that **60% of these individuals had reused at least one password across more than one account.** This figure exactly matches the rate of password reuse in our 2019 data,\* when calculated the same way, and it's close to the password reuse rate for users with more than one password across the entire SpyCloud database (57%). Interestingly, our database-wide rate of password reuse is not far off from the percentage of people willing to admit to it in previous years' studies (<u>66% in 2019</u>).

This analysis of user data across multiple breaches might lead you to wonder — if my data has been exposed in 1 data breach, am I more likely to appear in multiple breaches? By looking at the number of times email addresses appear across breaches, **we** estimate that the average person, if exposed once, will be included in 8-10 breaches, 3-4 of which could be during a given year.

\* Sharp-eyed readers may remember that last year's report listed the password reuse rate as 28%, not 60%. That figure was lower because the calculation included users who only had one exposed password in 2019. Since one password doesn't give us enough information to identify password reuse, we decided it would be more accurate to stick to users with two or more exposures.



#### **Popular Passwords of 2020**

Memorable passwords may seem unique to users, but they often aren't. Among the millions of passwords SpyCloud recovered from breaches in 2020 alone, "123456789" was found over 3.6 million times; "password" was found 1.2 million times, "qwertyuiop" 343,504 times, "princess" 108,216 times, and "monkey" 103,921 times. Unless these passwords are banned and password complexity requirements put in place, some users will always select easy-to-remember passwords.

#### **Need A New Password? Check the Headlines!**

After such a turbulent year, we wondered if users might have taken some inspiration from 2020 trends and events when creating their passwords. We checked last year's recovered credentials for some popular words to find out how often they had appeared. Sure enough, we found these keywords embedded within over 2 million passwords. Since this list isn't limited to complete passwords, our imaginations have run wild wondering how people have used these terms in context and what sentiments they may have been expressing ('sourdough4ever' or 'sourdoughsadness'?).



123456 123456789 12345678 passwd 12345 111111 password 123123 1234567890 1234567 qwerty 000000 1234 123321 654321 picture1 666666 abc123 qwertyuiop 11111111 987654321 Ghjghjghj9 qwerty123 zxcvbnm 1q2w3e4r 12345678910 pass123 123123123 1q2w3e4r5t 123456789aB 112233 5201314 123qwe a123456 121212 555555 admin iloveyou defaultPassw0rd 123456abc 7777777 password1 test1 senha 0000000 1q2w3e 1qaz2wsx Aa123456. asdfghjkl 123456a 8888888 Brasil Password 0123456789 password123 999999 88888888 123654 aaaaaa 123456789a 222222 159753 abcd1234 qwe123 147258369 dragon 123abc 1111 qazwsx 0987654321 princess 777777 1234qwer monkey 102030 a123456789 12341234 Aa123456 789456 pokemon asdasd qwerty1 azerty 147258 sunshine asdfgh 0000 qwer1234 1111111 11111 minecraft 789456123 1314520 159357 football michael 7758521 11223344 [censored]you

THE KEYWORD 2020 APPEARED IN 1,638,383 PASSWORDS



THE KEYWORD **POLICE** APPEARED IN 111,544 PASSWORDS

## VOTE DEMOCRAT CORONAVIRUS REPUBLICAN TIGER KING BLM COVID 2020 PANDEMIC MASK BIDEN POLICE PRESIDENT SOURDOUGH

#### **Hashing Algorithms**

Industry standards call for enterprises to hash stored passwords, which should theoretically slow down criminals who gain access to the hashes. SpyCloud research has shown that both the strength of the hashing algorithm and complexity of the passwords help determine whether the hashes could take criminals years to crack, minutes, or just seconds.

When SpyCloud recovers data from closed criminal communities, we often receive passwords in a hashed format — but it's only a matter of time before criminals start sharing plaintext credentials. We crack password hashes in-house to operationalize the data for enterprise security teams, helping our customers identify exposed users and reset their passwords before account takeover attempts begin.

Companies that are serious about protecting their employees, consumers, sensitive corporate data, and PII must modernize their password hashing efforts. Only the strongest hashing functions stand a chance against savvy cybercriminals. We recommend that organizations follow <u>NIST guidelines</u> for authentication as they make decisions about how to store authentication secrets.

On the other hand, even the strongest hashing algorithm means little when users make <u>weak or common password choices</u>. Older breaches, which are more likely to have been hashed using now-outdated algorithms, can help criminals launch association attacks against harder to crack breaches and confirm whether users are still recycling old passwords.

Each year, we see a slightly different distribution of hashing algorithms in the data we've collected. Year after year, one trend is obvious: Far too many credentials are available to cybercriminals that have been hashed using weak, outdated methods that are easy to crack.



#### Personally Identifiable Information (PII)

It might be cliche, but personally identifiable information (PII) is a goldmine for cybercriminals. Whether lost, stolen, or exposed, PII is how identity thieves perpetrate crimes. Sometimes all it takes is one or two pieces of information to compromise a person's identity.









#### Total PII Recovered in 2020: **4.6B Assets**

Total PII in SpyCloud's Database: **30.9B Assets** 

Total Breach Assets in SpyCloud's Database: **115B** 

Bad actors can use PII to apply for loans or lines of credit, make purchases with our credit cards, steal our tax refunds, drain financial accounts, and more. They can use it to create synthetic identities, or to bypass multi-factor authentication and take over existing accounts. Eventually, PII is bought-and-sold as a commodity on the dark web. Full packages of PII (known as "fullz") give criminals everything they need to commit identity fraud — typically name, national ID number, date of birth, and specific account credentials — and can be sold for \$8-10 according to our own research. When financial information is included, the criminal can command a 10x+ higher price.

Looking at some examples of our PII collection over the last few years in the charts on the following page illustrates how much the types and quantities of exposed PII can shift from year to year, depending on what specific data breaches emerge. Unfortunately for enterprises and consumers, the net result is still that billions of PII assets are exposed in breaches every year - and previous years' stolen data continues to provide bad actors with value as time goes on.

Over the last year, SpyCloud has discovered an eye-opening amount of exposed PII data.

#### PII Data Recovered by SpyCloud in 2020



#### 27M ISP 3 **19M** D National ID **6M** Age PE $\odot$ 10**M** Geolocation **5**M Company Name **4**M Estimated Income \$ ---**3M** Home Value **1M** Bank Number **1M** Credit Card Number

#### PII Data Recovered by SpyCloud, 2017-2020



## Top 10 Breaches of 2020

Of the countless breaches that occur every day, the ones that make headlines are those associated with popular brands, federal agencies and, of course, celebrities. Next to the attack on the U.S. government, the compromise of Twitter accounts held by several high profile stars probably captured the most attention in 2020.

Because SpyCloud's focus is on recovering data early in the breach timeline, we often can't talk about specific breaches until they're old news. Publicly disclosing our most exciting findings could jeopardize ongoing investigations or interfere with our researchers' undercover work. For example, our researchers' access to criminal communities means we often acquire stolen data before a victim organization is even aware they have been breached. In these cases, we practice responsible disclosure and support the efforts of the affected organizations and authorities to bring the responsible parties to justice.

When we make data from these breaches available to our customers to help them protect vulnerable accounts, we categorize them as "sensitive sources." We also classify potentially-controversial breaches (such as dating sites) as sensitive sources, particularly when they don't validate email addresses and could be used to tarnish an employee's reputation.

Excluding sensitive breach sources and combolists, here are the largest 10 breaches of 2020 based on the total number of breach records.



## **Government Credential Exposure**

The ongoing digitization of all aspects of modern life did not exclude the government. In Q1 alone, there were 17 million leaked government records: a 278% increase compared with Q1 of 2019. And it didn't let up. The pandemic-related shuttering of federal office buildings sent an already high-risk workforce home with access to some of the most accurate and sensitive databases available.

In the United States, cyberattacks have been a cause for concern for years; not only has the frequency of breaches increased, but so have the global and economic implications. In 2018, the United States was the country most severely affected by cybercrime in terms of financial damage: industry experts estimate that the U.S. government faced costs of over <u>\$13.7 billion USD</u> as a result of cyberattacks. 2020 culminated in the revelation that attackers had infiltrated at least 17,000 U.S. government and private networks in the SolarWinds supply chain attack, an unprecedented campaign that will take years to fully unravel.

Given the spotlight on government security, we combed through the breach data SpyCloud researchers recovered in 2020 to find breach records containing .gov email addresses. Within our 2020 data, we were able to identify 269,690 plaintext government credentials within a total of 465 breaches, all of which provide potential avenues for bad actors to access government resources.

		J
ſ		J
ſ		Ĵ
٦		







**Password reuse rate** for .gov emails with two or more passwords collected in 2020

## Top 10 Passwords Associated with Government Emails

	PASSWORD	COUNT
1	Abcd1234	3,488
2	password	1,645
3	aaron431	771
4	EvoPassword	702
5	pvtagent25	577
6	g_czechout	493
7	N0=Acc3s	480
8	123456	424
9	*NO-PASSWORD*	423
10	952013	384

## After the Breach: What Happens to Stolen Credentials?

Contrary to what you may have heard elsewhere, the first step to monetizing stolen data is not to sell it on the dark web. That's actually the last step. What happens first is the highest effort, most profitable activities. Once maximum value has been extracted from the data, only then is it packaged up for sale on the dark web.

#### With stolen data, criminals can:



Distribute ransomware and other malware



Drain financial accounts, crypto wallets, or loyalty point balances



Exploit victims' work accounts for data theft and business email compromise



SIM-swap victims to bypass MFA



Make fraudulent purchases



Sell or trade account access to other criminals



Create synthetic identities

## Your Plan of Action

Each year has its cybersecurity themes — in the past, criminals have taken advantage of natural disasters, election cycles, and economic turmoil. 2020 was a pandemic year, and it stands to reason that the trends of 2020 will continue to disrupt our lives in new, accelerating ways.

Coupled with high rates of password reuse, the 1.5 billion exposed credentials SpyCloud identified in 2020 represent significant account takeover risks for both consumers and enterprises.

Attackers actively test stolen credentials against different accounts to exploit bad password habits and gain access to

corporate systems and data. Even worse, stolen PII and account data make it easy for criminals to craft highly targeted, creative attacks that cause great harm and are difficult to detect.

Enterprises must be able to trust the identities of the employees, consumers, and suppliers logging into their networks — and safeguard the corporate assets and IP behind those logins. The answer is to build early detection and remediation of exposed credentials into their cybersecurity strategy, and the best method, simply put, is to use SpyCloud.



#### **Consumer ATO Prevention**

Protect your users from account takeover fraud and unauthorized purchases.

Learn More  $\rightarrow$ 



#### **Employee ATO Prevention**

Protect your organization from breaches and BEC due to password reuse.

Learn More  $\rightarrow$ 

## The SpyCloud Difference

Building a security program around technologies that proactively leverage data acquired through Human Intelligence (HUMINT) tradecraft very early in the breach timeline is a critical path to success. SpyCloud's solutions, backed by the world's largest repository of recovered stolen credentials and PII, enables enterprises to stay ahead of account takeover by detecting and automatically resetting compromised passwords early, before criminals have a chance to use them.

Our customers continue to tell us their ability to prevent account takeover hinges both on access to relevant data (including the most plaintext passwords in the industry) and in being able to make that data operationally actionable through automation.



#### VIP Guardian

Protect your highest-risk executives from targeted account takeover.

Learn More  $\rightarrow$ 

#### <u>See Your Account Takeover Risk</u> →

Discover how many breach records we have associated with your email address and your domain as a whole. Once you know, you can take action.

#### **Active Directory Guardian**

Automatically detect and reset exposed Windows accounts.

Learn More  $\rightarrow$ 



#### **Third Party Insight**

Monitor third party exposures and share data to aid in remediation.

Learn More  $\rightarrow$ 

## **Spy**Cloud

Learn more at spycloud.com