Welcome

In so many ways 2020 has become the year of change. Practically every aspect of daily life has been touched by the pandemic, from how we do our work to what we do with our free time. Our data breach research is no exception.

Followers of our research are already aware that breach reporting has taken something of a hit this year. Turbulent times have dominated the news cycle, taking a bite out of the breadth of media attention available to report on data exposure events, while at the same time resources remain stretched at the public service offices that provide information on breach activity. We've said this before but it bears repeating: it is our belief that data breach activity remains relatively unchanged. Instead, it is the information reported about these events that has slowed.

Like so many others, the Cyber Risk Analytics research team is adapting to the change and focusing our efforts on the data most relevant to our customers. That means we're adjusting our priorities, spending more time on new data sources and shifting attention away from less productive ones. One such example is turning our attention to monitoring the many ransomware operator sites that have cropped up this year. We are also in the process of expanding our schema by enriching our incident data with additional impact information. We're looking forward to sharing this research with you in future reports.

For now, this report examines publicly disclosed data breach events first reported between January 1, 2020 and September 30, 2020. As always, it is our sincere hope that you find this report valuable and enjoy reading it as much as we enjoy bringing it to you.

Key Highlights

- There were 2,953 publicly reported breaches in the first three quarters of 2020, a 51% decrease compared to the same time period last year.

- 2020 was already the “worst year on record” by the end of Q2 in terms of the total number of records exposed. The three months of Q3 added an additional 8.3 billion records to the count, bringing the number of records exposed through the end of September to a staggering 36 billion.

- Two breaches in Q3 exposed over 1 billion records each and four breaches exposed over 100 million records. Together these six breaches accounted for approximately 8 billion exposed records, or 22.3% of the records exposed through the end of Q3.

- Malicious actors continue to be the driving force behind the number of breaches occurring, while misconfigured databases and services remain the leading cause behind the number of records exposed.

- In the first three quarter of 2020, 21% of reported breaches involved the use of ransomware. These ransomware-related events contributed to the unusually high number of unknown (11.2%) and miscellaneous (10.4%) data types exposed.

- Following well established trends, the Healthcare sector had the most reported breaches, accounting for 11.5% of the events that could be attributed to a specific economic sector.
In This Issue

FEATURING VIEWPOINTS FROM

Inga Goddijn
Executive Vice President,
Risk Based Security

The Cyber Risk Analytics Research Team
Risk Based Security

WELCOME
Key Highlights .......................................................................................................................... 2
In This Issue ............................................................................................................................... 3
Viewpoints ................................................................................................................................. 4
  Ransomware and Breaches – The Worst of Both Worlds ......................................................... 4
  Hacktivism: A Neglected Threat Looms Over US Election ....................................................... 6
Data Breach Trends in 2020 ....................................................................................................... 10
  Mid Year 2020 At A Glance ..................................................................................................... 10
  What Was Breached in 2020? ............................................................................................... 15
  Who Was Breached in 2020? ............................................................................................... 17
  Where Were Breaches Reported in 2020? ............................................................................ 18
In Closing .................................................................................................................................. 20
  Methodology and Terms ........................................................................................................ 20
  Data Standards and the Use of “Unknown” ........................................................................... 20
About Risk Based Security ......................................................................................................... 21
  About Cyber Risk Analytics ............................................................................................... 21
  No Warranty .......................................................................................................................... 21
Ransomware and Breaches – The Worst of Both Worlds

Inga Goddijn, Executive Vice President, Risk Based Security

Inga found her way to information security after working for twenty years in the insurance industry. During her time managing a multi-million-dollar portfolio of technology and cyber insurance coverages, Inga witnessed first-hand the impact of ineffective security program management and the financial fallout from data breach events. At Risk Based Security, she is responsible for Cyber Risk Analytics and YourCISO. Inga has presented at a variety of industry forums and has led many education sessions throughout the U.S. She currently holds a CIPP/US designation.

I recently commented to a client that, these days, ransomware events are likely to be data breaches as well. The client replied, “we tend to agree with you, but our customers don't always see it that way.” That response took me by surprise and has stuck with me, serving as a reminder there is no singular, uniform definition of a “data breach”. Incident responders, business leaders, lawmakers, regulators, CISOs and cyber insurance providers all have a unique perspective on what constitutes a “data breach”. To that end, it’s worthwhile to take a step back and share our take on the type of incidents that make their way into our “data breach” database.

**OUR DEFINITION OF A BREACH**

Our operational definition of a “data breach” is the **unauthorized access to, or loss of control of, confidential or sensitive information**. This allows our research team to capture and catalog a broad spectrum of data security events, ranging from conventional malicious actions, resulting in data theft, to more nebulous situations such as lost or missing equipment that holds sensitive data. We hold an equally broad view of what constitutes confidential or sensitive information. This includes data routinely subject to regulatory protection, such as financial account details or healthcare information, to sensitive insider information, such as source code, business development plans, and trade secrets. Even information destined to become public can be the source of a breach. Such is the case with several incidents involving unauthorized access to early earnings reports, regulatory investigations in process, and even economic forecasts and meeting notes from central bankers.

**WHY WE VIEW RANSOMWARE AS A BREACH**

Ransomware has evolved far beyond its humble beginnings, with its operators becoming especially pernicious in recent months. It is common to see attackers lurk inside their targets’ systems before launching the encryption process. Even if attackers don’t exfiltrate data in an effort to ratchet up the pressure on their target to pay, the access could allow attackers to peruse files, documents and a host of other sources of confidential information.

Through the first nine months of the year, the research team cataloged over 440 breaches that included a ransomware component to the attack. This includes both the obvious “data breaches”, where information was clearly taken, as well as those events that indicate attackers were inside systems or services and could have accessed sensitive data. These events have had an interesting impact on several statistics, and we’ll highlight those effects throughout the report.
The Data Breach QuickView report is powered by Cyber Risk Analytics

The standard for actionable data breach intelligence, risk ratings and supply chain monitoring.

- Cyber Insurance
- Security & Vulnerability Management
- Vendor Risk Management
- Procurement
- Governance & Management

REQUEST A DEMO
sales@riskbasedsecurity.com

LEARN MORE
Hacktivism: A Neglected Threat Looms Over US Election

The Cyber Risk Analytics Research Team, Risk Based Security

Since 2011, the Cyber Risk Analytics (CRA) research team has provided organizations with the most comprehensive threat intelligence solution on the market, providing details about the number of data breaches, records exposed, industry type, threat vectors, and much more. The CRA research team has analyzed over 47,500 breaches exposing 69 billion records and monitors over 104,000 organizations.

Back in 2016, the news was flooded with debate, discussion and outrage over Russia's alleged meddling in the American presidential election. While many still do not fully agree with the extent or real impact, articles were published that showed that cyber threat actors backed by the Russian government attempted to influence American voters through a variety of actions ranging from social media campaigns, the spread of disinformation, to email inbox hacking, and voter database and election supply chain hacking.

Political leaders and pundits alike expressed concern at the extent of the cyberattacks and many supported investigations into possible Russian cyberattacks to influence the American election. Four years later, Russian interference continues to be an inflammatory topic.

These types of attacks might seem unique to this generation, but in fact influence campaigns and the spread of disinformation as a political tactic have been around for nearly as long as the political systems they seek to undermine. What is new, is that the modern plethora of digital platforms has enabled campaigns to be conducted in near real-time and at a much greater scale.

The potential negative impact is amplified when coupled with the ability to compromise systems and data. With the November elections fast approaching, researchers at Risk Based Security have observed greater circulation of and interest in voter-related information on the dark web. While some of the information may be from public sources, and has not been independently verified for authenticity, it is certainly a signal of renewed interest in this type of data.

Steps to mitigate the risk and secure voter data and electoral systems must be taken to protect the integrity of our democratic process and ensure the legitimacy of the results; if not for this election, all future elections.
Politically motivated hacking, or hacktivism, is not a new occurrence. It can be traced back to the ‘80s and ‘90s, when threat actors realized the reach of the growing internet infrastructure. More recently, governments have dedicated significant resources towards progressing their international goals through hackers.

With so much going on in 2020, there does seem to be a certain alarming sense of complacency among voters, local officials, and the media in combating and discussing the issue prior to the upcoming presidential election. While the pandemic and the campaigns of the current candidates continue to dominate headlines, the cybersecurity industry as a whole appears to be less focused on tackling hacktivism than it did in 2016.

We have published articles and reports detailing the vulnerabilities present in U.S. Electronic Voting Machines, as have others, but with the exception of the excellent work of a handful of researchers including Kim Zetter, the media’s level of interest has been low.

This could be due to general “breach fatigue”, heightened focus on the mail-in voting process, or simply the overwhelming amount of newsworthy events taking place in recent weeks. It makes sense to focus on the issues that are on your doorstep, rather than those that are not as transparent.

But if hacktivism is “old news” to the cyber world, it certainly doesn't appear to discourage hackers. In fact they are profiting from the misdirection, and probably would prefer to keep it that way.
THE USUAL SUSPECTS

Our researchers observed numerous databases claiming to consist of American voter information being shared and discussed at a growing rate on both Russian speaking and English speaking hacking forums.

There is little to indicate these databases signal recent breach activity. Since voter information is public record in some states, the databases may just simply consist of scrapes of public sources. It is also possible the data is recycled or extracted from previous breaches. Federal officials have gone so far as to issue a warning that these databases are being circulated to further fear, uncertainty and cast doubt about the upcoming election.

While our researchers have not analyzed these databases, the recent increase in shares or discussions of these databases points to a timely interest:
• **Michigan** – Database of over 7 million voters including PI and voting information.

• **North Carolina** – Database of over 8 million voters including PI and voting information.

• **Washington** – Database of 5 million voters including PI and voting information.

• **Florida** – Multiple databases including PI and voting information, with voting history.

• **Other US States** – Database containing 66 files of individuals from various states with personal information and political views; private database of 186 million voters with personal information and political party affiliation offered for sale.

It is an uncomfortable fact that these purported voter databases involve **pivotal states like Florida, Michigan and North Carolina**.

While much of this data might have been collated from older or publicly accessible sources, the potential dangers are still very real. The increased attention and cooperation between hackers points to a growing interest and overall risk. They would most likely prefer for us to think that hacking isn't a real issue, given the current climate, but circulating these types of databases can leave voters feeling vulnerable and feed mistrust of voter systems. Additionally, just because the data may be older or publicly sourced, it can still have value for attempting voter fraud or target voters using highly personalized campaigns as detailed in the 2016 election.

**ARE WE OUT OF TIME?**

While hacktivism during a year with major political events seems inevitable, critical safeguards have likely not kept up with threat actor advancements or ambitions.

Research published at the end of last year found that experts were easily able to breach the voting machines that are being used for the 2020 election. The same was found in the 2018 midterm elections. Naturally, the outdated systems and teams of local election volunteers are no match for the nation-backed threat actors. Similarly, voters may not know their information is public record and that they are at an increased risk for targeting. More informed voters on how their information is potentially being abused could help combat the risk in the future.

The good news is that there are resources available. The Cybersecurity & Infrastructure Security Agency has implemented numerous programs for protecting election infrastructure. However, despite being made freely available, local election boards have struggled to make use of the services.

The U.S. election is only days away and with new reports that state-sponsored hacktivism is once again on the rise, it seems imperative that more needs to be done. An increase in the reporting from the media as well as discussion from the cybersecurity industry over the heightened risks can help create more pressure necessary to move quickly.

With days until the election and already more than 2 million votes cast, it is likely too late in this election cycle for a radical shift in mentality toward security, but we can certainly still change the narrative for all the ones to come.
Data Breach Trends in 2020

The quagmire that formed in the breach landscape this Spring has continued through the third quarter of the year. Breach disclosures continue to be well below the high water mark established just last year despite other research indicating the number of attacks are on the rise.

How do we square these two competing views into the digital threat landscape? As noted in our previous 2020 quarterly reports, one factor contributing to the decline in publicly disclosed breaches is simply less data breach media coverage.

Another factor that we believe is playing a role is the pivot by malicious actors to more lucrative ransomware attacks. While many of these attacks are now clearly breach events, the nature of the data compromised can give some victim organizations a reprieve from reporting the incident to regulators and the public. After all, while the compromised data may be sensitive to the target organization, unless it contains a sufficient amount of personal data to trigger a notification obligation the event can go unreported.

Mid Year 2020 At A Glance

The number of publicly disclosed data compromise events through September 30th is approximately 50% below 2019 activity and 25% below 2018. While we do anticipate the gap will close somewhat as additional information is surfaced from lagging sources, it is doubtful that 2020 will reach the level of reporting seen last year. In fact, 2020 is on track for the slowest pace of breach disclosure in the past five years.

The number of records compromised so far this year, however, stands in stark contrast to the number of breaches reported. **With over 36 billion records exposed, this year has already seen twice the number of records exposed than in all of 2019.** This is in spite of the fact that at this point in 2019, only 23.6% of reported breaches did not include record count information, while a sizable 43.6% of 2020’s breaches omit the number of records exposed, suggesting that the number of records exposed year to date could be much higher than currently reported.
Despite initial indications that breach reporting was slowing by February, the first three months of the year showed some signs it would be a temporary glitch with the general trend toward increasing disclosures. As the pandemic wore on, social upheaval gripped communities, and workplace disruption continued, the number of reported breaches languished.

Followers of our QuickView reports will quickly recognize the long standing trends reflected here. Unauthorized access to networks, systems and services, shown as “Hack” in the charts, is responsible for 64% of reported breaches. These include events where attackers were able to gain access to data prior to launching ransomware. Even with 1.6 billion records compromised via hacking, sensitive information accessible online due publishing errors, programming flaws or database misconfigurations, shown as "Web" in the charts, is the breach type responsible for exposing the most records this year.
Not surprisingly, with the majority of breaches attributable to "Hacking", 77.5% of events originate outside of the victim organization. Of the approximately 17% of breaches originating from within the organization, the overwhelming majority, 67%, are from errors. These range from the misconfigured databases that expose hundreds of millions of records to lost devices with minimal risk resulting from the error.

Figure 6: Number of breaches by attack vector, reported by Q3 2020
As noted in the outset of the report, our working definition of a "data breach" is unauthorized access to or loss of control of confidential or sensitive information. With such a large scope, it can be useful to qualify the confidentiality impact. This is measured in terms of ‘confirmed’, meaning there is affirmative evidence data was exposed such as when it appears for sale or is posted to a ransomware operators website, versus "potential", meaning the data was accessible but there is uncertainty as to whether it was actually accessed. This is unfortunately a common scenario to see when malicious actors gain access to employees’ email accounts with little evidence to rule out access to the contents of the at risk emails.

![Pie Chart](image)

**Figure 7:** Number of breaches by known confidentiality impact, reported by Q3 2020

- **Confirmed:** 1,901
- **Potential:** 1,051
- **Unknown:** 1
Breach severity is calculated based on a number of factors including the number and type of records exposed, parties impacted, and follow on events such as lawsuits arising from the event. In the 2020 Mid Year Report, we noted that breach severity was increasing. In Q1 the average severity was 4.75, while at the time the Mid Year report was published, the Q2 average severity stood at 5.47. As data has developed, the average severity for Q2 has since inched up to 5.54. A similar trend appears to be unfolding for Q3. At the time of this report, the average severity for the current quarter is once again higher than the previous quarter. As information about Q3 events unfolds we anticipate Q3 will ultimately adjust upward as well.

Figure 8: Severity distribution of breaches in 2020 Q2 and Q3
What Was Breached in 2020?

Names, along with access credentials in the form of email addresses and passwords, continue to be the most exposed data types of the year. “Miscellaneous” has been gaining ground, along with “Unknown”, due in large part to the number of ransomware-related events.

![Data types exposed in breaches reported by Q3 2020](image)

As mentioned earlier in the report, victim organizations are understandably reluctant to publicize the type of data exposed unless required to do so. When no confirmation is provided as to the type of data compromised, the data type “Unknown” is applied. Internal, non-personally identifiable information is generally categorized as “Miscellaneous” or in some instances, Intellectual Property. That said, it is worth noting that some forms of personal information such as educational records, passport details, or metadata associated with user accounts can also be classified as “Miscellaneous.”
Fifteen breaches reported during Q3 contributed to the 2020 running total of 51 breaches exposing 10 million or more records. Included in that 15 were four breaches exposing over 100 million records and two breaches exposing over 1 billion records. The largest breach of Q3, attributed to an open Elasticsearch server, exposed approximately 6 billion records. As damaging as that may sound, the 6.4 TB of data included multiple interactions with the same clients. In all, approximately 700,000 individuals had data exposed in the event.
Who Was Breached in 2020?

After two consecutive quarter with the Healthcare sector and the Information sector nearly tied for the most data breaches, Healthcare has definitively taken the top spot. There was a wide range of events taking place across all types of healthcare service providers. Some events reflected the unique workings of healthcare professionals, such as when a Dictaphone containing patient diagnoses went missing from an office, or when a high-profile patient's medical record was accessed on multiple occasions by curious staff. Other events follow a more common arc, with ransomware and employee email account hacking making frequent appearances.

Figure 10: Number of reported Q3 2020 breaches experienced according to economic sectors
Differences appear when diving deeper into the various business groupings that make up each economic sector. For the organizations that could be classified, software services, hosting and websites under the Information sector collectively accounted for approximately 11% of the breaches reported through September. Banks and other financial services companies accounted for approximately 8.4% of breaches, and the two retail groups combined and the three manufacturing groups combined nearly tied for third place with approximately 7% each of the reported breaches. None of the four business groups that make up the Healthcare sector accounted for more than 7% of the total. However, when the four are combined, the Healthcare sector as a whole experienced the most breach events.

Figure 11: Distribution of breaches by business group and business subgroup, reported by Q3 2020
Where Were Breaches Reported in 2020?

In the 2019 Q3 Data Breach QuickView Report, there were 1,692 breaches reported taking place in the United States and 591 outside of the US. For the same nine months of 2018, there were 1,846 breaches reported taking place in the US and 708 outside of the US. In all, this trend has been remarkably stable in recent years.
A broad spectrum of economic sectors are represented in the breaches reported in the United States. Not surprisingly, the Information and Healthcare sectors reigned supreme taking first place as the sectors with the most reported breaches in 46% of the states. The Information sector accounted for the most reported breaches in 12 states, with the Healthcare sector taking the top spot in 11 states.
In Closing

Early in the year, we predicted that 2020 breach trends would level out and we'd see the year unfold along the same patterns we observed in the prior two years. By the time of our Mid Year report, the data had us questioning that assumption and wondering if indeed we had arrived at a new normal. Although we were surprised by the decline in breach disclosures and equally surprised by the extraordinary amount of data exposed online, at the close of the Mid Year report we were still holding on to the notion that 2020 would find its way back to something more familiar. Clearly that is not in the cards for this year.

That said, we do remain confident that slow reporting and a reduction in the amount of information surfaced about breach events should not be taken as an indication less breaches are occurring. As the saying goes, absence of evidence is not evidence of absence.

Be sure to tune in for the Year End Data Breach QuickView Report in January to see how the dust settles on this most unusual year.

Methodology and Terms

Risk Based Security’s research methods include automated processes coupled with traditional human research and analysis. Our proprietary applications crawl the Internet 24x7 to capture and aggregate potential data breaches for our researchers to analyze. In addition, the research team manually verifies news feeds, blogs, and other sources looking for new data breaches as well as new information on previously disclosed incidents.

The database also includes information obtained through Freedom of Information Act (FOIA) requests, seeking breach notification documentation from various state and federal agencies in the United States. The research team extends our heartfelt thanks to the individuals and agencies that assist with fulfilling our requests for information.

Data Standards and the Use of “Unknown”

In order for any data point to be associated with a breach entry, Risk Based Security requires a high degree of confidence in the accuracy of the information reported as well as the ability to reference a public source for the information. In short, the research team does not guess at the facts. For this reason, the term “Unknown” is used when the item cannot be verified in accordance with our data validation requirements. This can occur when the breached organization cannot be identified but leaked data is confirmed to be valid or when the breached organization is unwilling or unable to provide sufficient clarity to the data point.
About Risk Based Security

Risk Based Security (RBS) provides detailed information and analysis on Vulnerability Intelligence, Vendor Risk Ratings, and Data Breaches. Our products, Cyber Risk Analytics (CRA), VulnDB and YourCISO, provide organizations access to the most comprehensive threat intelligence knowledge bases available, including advanced search capabilities, access to raw data via API, and email alerting to assist organizations in taking the right actions in a timely manner.

For more information, visit www.riskbasedsecurity.com or call +1 RBS-RISK.

About Cyber Risk Analytics

Cyber Risk Analytics (CRA) provides actionable threat intelligence about organizations that have experienced a data breach or leaked credentials.

Along with our PreBreach Risk Ratings, this provides a deep dive into the metrics driving cyber exposures, as well as understanding the digital hygiene of an organization and predicting the likelihood of a future data breach.

The integration of PreBreach ratings into security and underwriting processes, vendor management programs, and risk management tools allows organizations to avoid costly risk assessments, while enabling businesses to act quickly and appropriately to proactively protect its most critical information assets.

NO WARRANTY

Risk Based Security, Inc. makes this report available on an "As-is" basis and offers no warranty as to its accuracy, completeness or that it includes all the latest data breaches. The information contained in this report is general in nature and should not be used to address specific security issues. Opinions and conclusions presented reflect judgment at the time of publication and are subject to change without notice. Any use of the information contained in this report is solely at the risk of the user. Risk Based Security, Inc. assumes no responsibility for errors, omissions, or damages resulting from the use of or reliance on the information herein. If you have specific security concerns please contact Risk Based Security, Inc. for more detailed data loss analysis and security consulting services.

*Maps by freevectormaps.com and mapchart.net