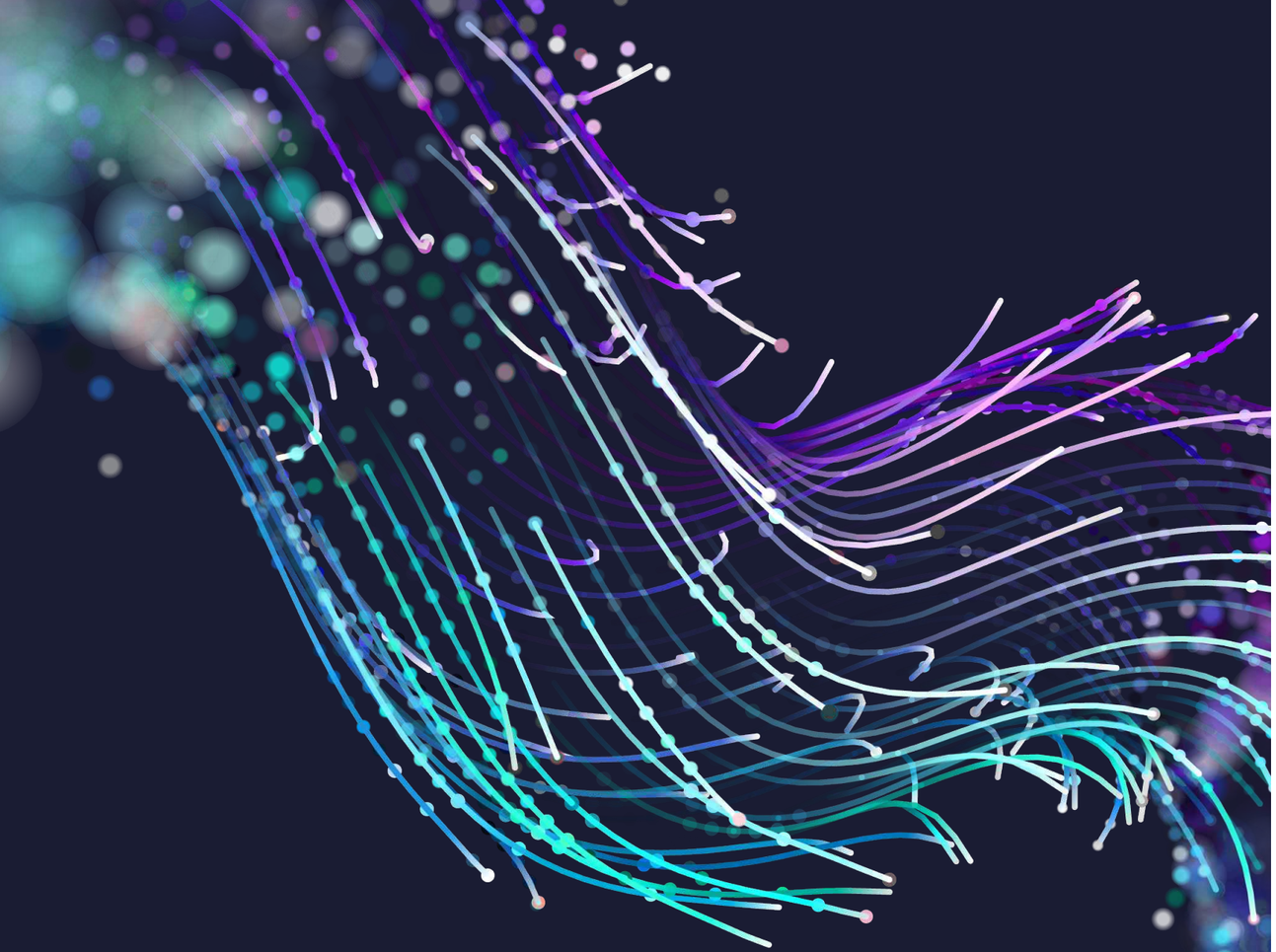

FLASHPOINT

The State of Data Breach Intelligence

2022 Midyear Edition

In This Issue

Welcome
to the State of Data Breach Intelligence

Welcome to the tin (or aluminum, if you prefer) edition of Flashpoint's State of Data Breach Intelligence Report! We've reached a milestone, as this is the tenth straight year for publication (formally titled the Data Breach QuickView).

Over the decade of producing this report, we have seen significant shifts in the type of data that has been exposed and the attack methods that were used to gain illicit access to it. Slowly but steadily, lawmakers have expanded the boundaries of what constitutes sensitive data. Changes aside, one observation has remained consistent over the years. It is the organizations that can adapt to the threat landscape and have learned how to bend their processes, to address the most relevant risks that often fare the best when it comes to defending against and recovering from a data breach.

The **State of Data Breach Intelligence: 2022 Midyear Edition** covers publicly disclosed compromise events first reported between January 1, 2022 and June 30, 2022. Note, the comparisons found throughout this report correspond to the same time period from prior years unless otherwise noted.

Key Findings

- 1,980 breaches were reported in the first six months of the year, approximately 15 percent below 2021's final H1 total.
- The number of records exposed dropped dramatically in the first six months of 2022 compared to the first six months of 2021, falling from 27.3 billion records to 1.4 billion records.
- The decline in records exposed can be attributed to a decline in the number of breaches impacting 100 million or more records. In 2021 H1, there were 13 such incidents. In 2022 H1, only three such incidents have been reported.
- The most prolific breach type remains consistent with prior years, with unauthorized access to systems (aka "Hacking") accounting for approximately 60 percent of breaches reported in 2022 H1.
- The combined Healthcare and Social Service economic sector reported the most breaches in 2022 H1. However on a business group level, Financial Services and Software/Data Services both reported more breaches than Hospitals, the leading reporter of breaches within the Healthcare economic sector.

Number of Breaches and Records Lost
2022 Midyear data breach trends

Frequent readers of this report are familiar with the disruptions to breach reporting that occurred at the onset of the pandemic and continued into 2021. At long last, reporting has settled into a more predictable cadence and while some delays persist, breach disclosures and the data they contain show a return to more conventional patterns. Throughout the report we will be looking more closely at those patterns in terms of the what, how, and who of breach activity.

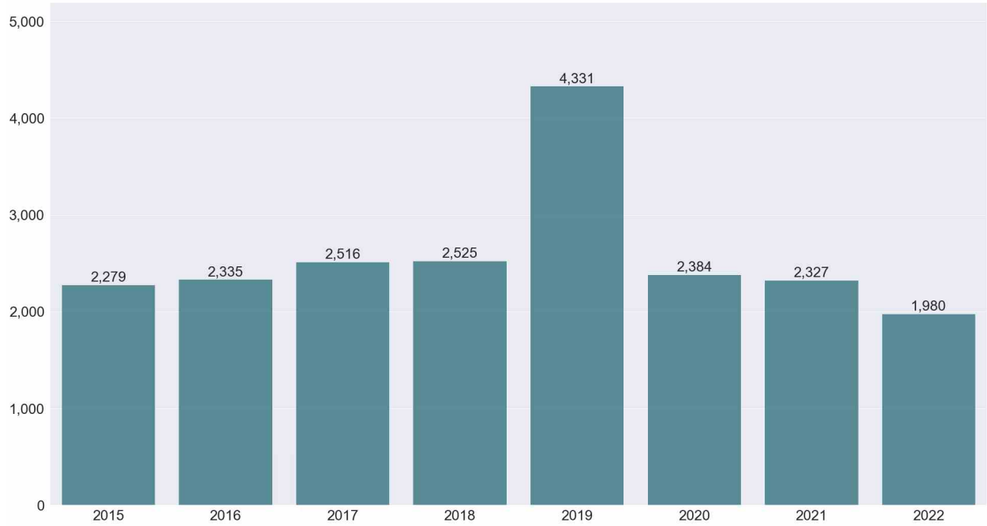

Figure 1: Number of breaches disclosed by H1, in the last eight years

As of the writing of this report, there were 1,980 publicly disclosed data breaches, down approximately 15 percent compared to the same period last year, after accounting for backfill. This should not be construed as a decline in breach activity. Rather, this is a typical pattern when factoring in ordinary data development.

Assuming 2022 H1 sees an additional 20 percent to 25 percent increase of reported breaches after the publication of this report, the final total will likely be on par with the prior two years.

Year	Breaches disclosed at time of midyear report	Breaches disclosed as of 6/30/22	Percentage increase
2021	1,767	2,327	31.7%
2020	2,037	2,384	17%

Table 1: Comparison showing the number of breaches disclosed by H1 to backfilled total, in 2021 and 2020

Shown in figure two, the number of records exposed in the first half of the year dropped to its lowest level since 2015. This is due in large part to an overall decline in the number of very large, open misconfigured services and databases resulting in a breach. These misconfigurations have been the driving force behind the astronomical number of records exposed in recent years.

Such misconfigurations are classified as breach type "Web" and while this allows for a variety of online oversharing activities, reviewing the combination of this breach type and breaches exposing over one million records shows the impact of misconfigured services on total exposed records.

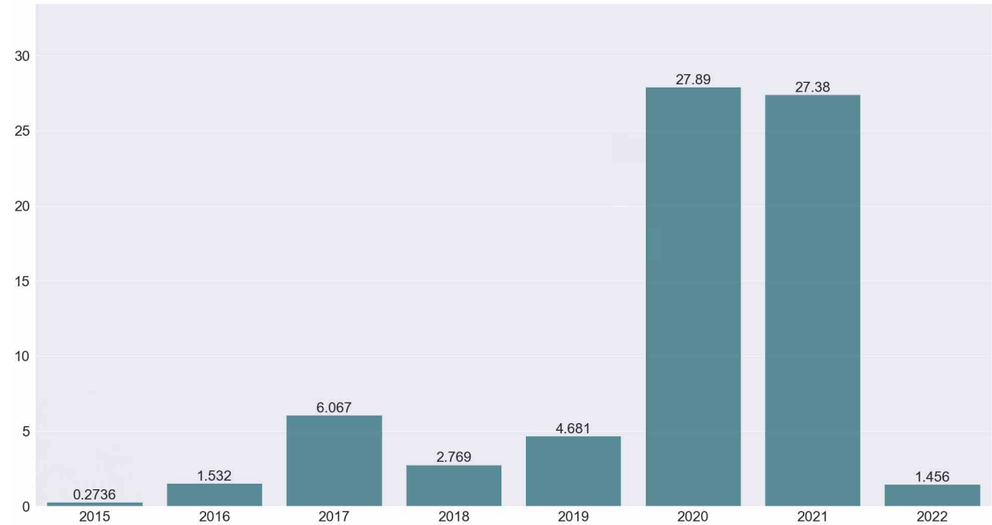

Figure 2: Number of records lost (in billions) reported by H1, for the past eight years

Year	Number of breaches exposing 1M or more records attributed to breach type "Web"	Number of records exposed
2022 H1	5	383,706,267
2021 H1	17	25,368,619,308
2020 H1	34	26,604,704,882
2019 H1	34	3,709,640,000
2018 H1	13	636,583,733

Table 2: Number of breaches exposing 1M or more records attributed to the "Web" breach type by H1, in the last five years

Number of Breaches by Breach Type

2022 Midyear data breach trends

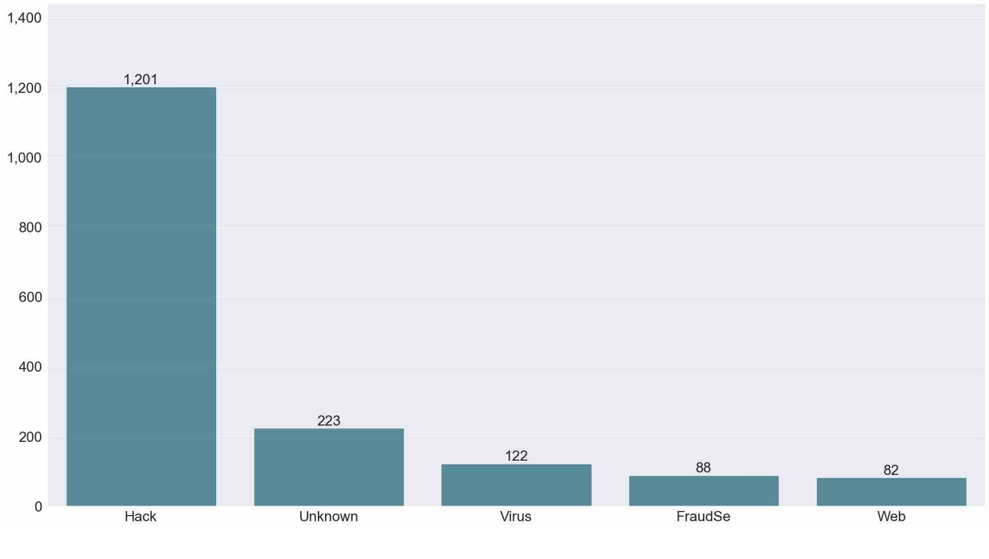

Figure 3: Number of breaches by breach type, reported by 2022 H1

Unauthorized access to systems or services, referred to as breach type "Hack," accounted for 60 percent of breaches reported in 2022 H1. This is not surprising, as hacking has been the top breach type since the first edition of this report.

The more interesting trend is the growing presence of breach type "Unknown." The language in breach notifications and other disclosure reports is increasingly opaque. Phrases such as "cyber attack" and "security incident" are commonplace, with woefully little else provided in the way of explanation. The effect of this shift in language is a key contributor to the steady increase in the number of breaches classified as "Unknown."

Time Period	Number of breaches classified breach type "Unknown"	Rank
2022 H1	233	2nd
2021 H1	156	2nd
2020 H1	119	4th
2019 H1	122	4th
2018 H1	77	9th

Table 3: Number of breaches attributed to the "Unknown" breach type by H1, in the last five years

"Unauthorized access to systems or services, referred to as breach type "Hack," accounted for **60 percent of breaches** reported in 2022 H1."

Number of Records Lost by Breach Type

2022 Midyear data breach trends

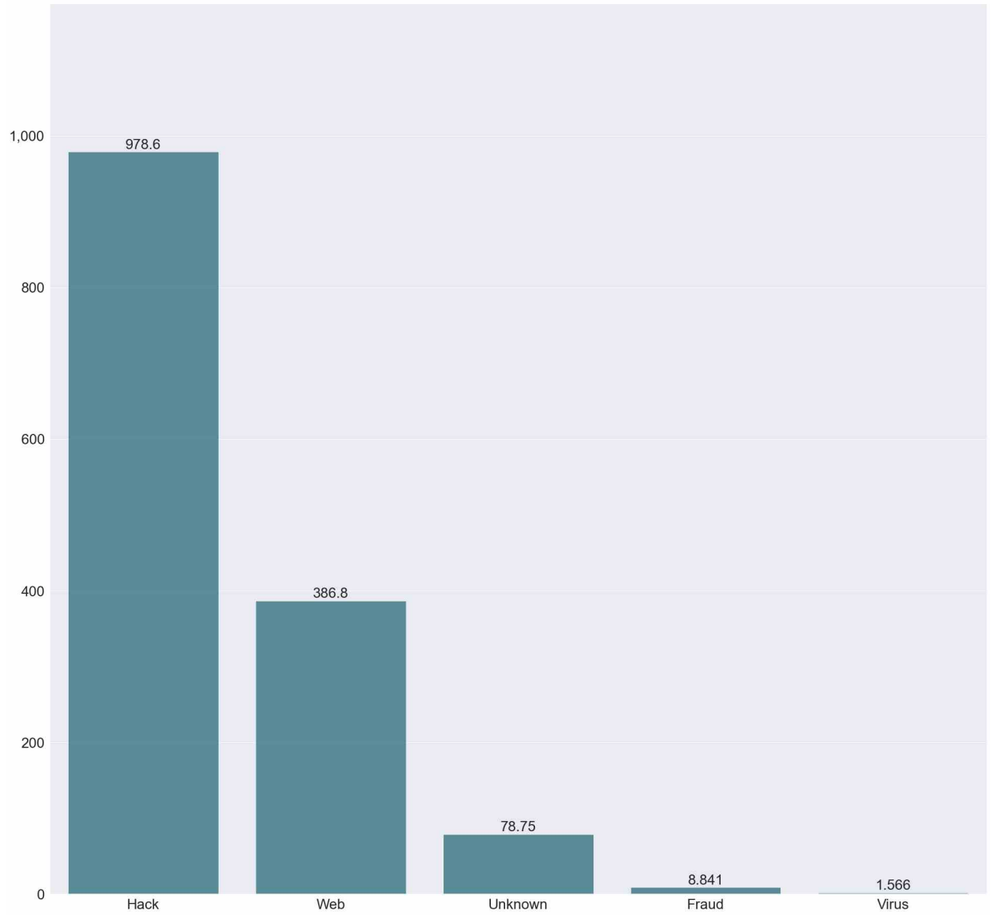

Figure 4: Number of records lost by breach type (in millions), reported by 2022 H1

The breach type accounting for the most records exposed has changed for the first time since 2018 H1, when fraud accounted for 46 percent of records exposed. "Web" has historically been the top breach type and time will tell whether this is a temporary blip or a more permanent change. Unlike stopping skilled malicious actors, web exposure is an issue largely under the organization's control.

Comprehensive and repeatable processes for hardening systems, coupled with periodic audits of controls can substantially reduce the likelihood of a large scale data leak. After several years publicizing the risk, organizations may finally have turned the corner on this issue.

> “Comprehensive and repeatable processes for hardening systems, coupled with periodic audits of controls can **substantially reduce** the likelihood of a large scale data leak.”

Inside and Outside Threats

2022 Midyear data breach trends

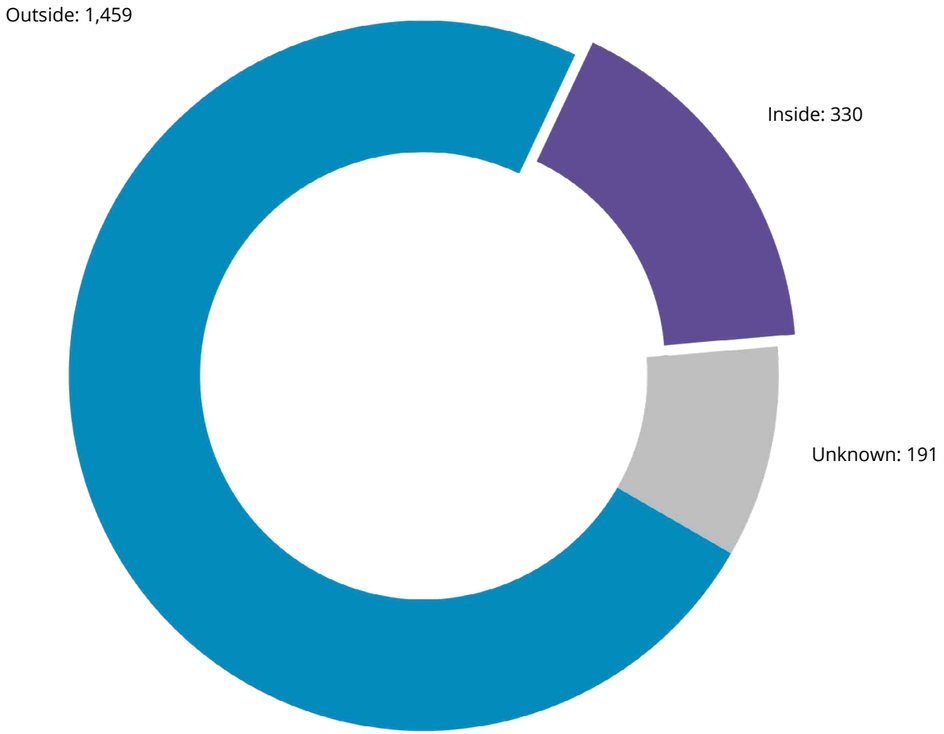

Figure 5a: Number of breaches by attack vector, reported by 2022 H1

Insider risk is a perennial topic of discussion. Is the insider threat fact or fiction? The answer is not a simple yes or no. Of the breaches with a confirmed origin, only 23 percent of incidents originated from within the victim organization and of that 23 percent, the majority, 61 percent, were attributable to data handling mistakes.

Of the 54 breaches confirmed to have originated with a malicious insider, the incidents range from the banal, such as small scale theft of credit card data from customers at the point of sale, to the potentially catastrophic such as theft of investment intensive technological innovations and proprietary source code.

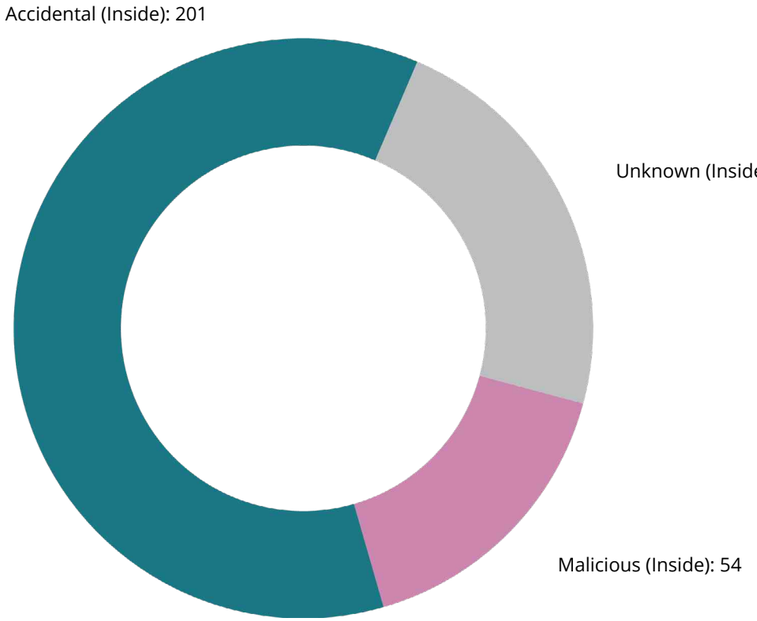

Figure 5b: Breakdown of "Inside" breaches, reported by 2022 H1

Confidentiality Impact

2022 Midyear data breach trends

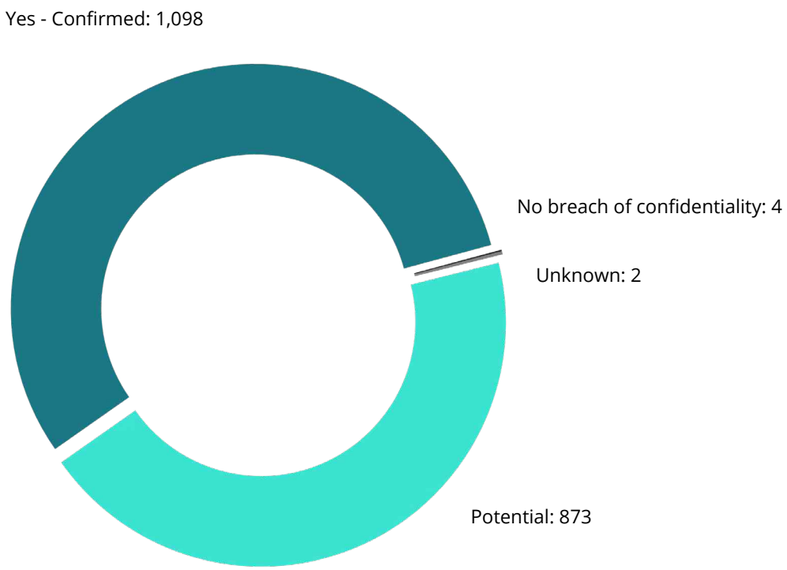

Figure 6: Number of breaches by known confidentiality impact, reported by 2022 H1

The C-I-A triad lies at the heart of information security and of the three, none is more central to data compromise events as Confidentiality. Even a casual reader of breach notifications will quickly notice distinctions are being made between confirmed data theft and the inability to confirm whether data was accessed or exfiltrated.

Even after investigation, it could not be determined whether data was compromised in approximately 44 percent of incidents. Should this be interpreted as a positive statistic? Likely the answer is no, as the high percentage of potential exposure indicates a lack of visibility into activity taking place on systems or within services.

Breaches That Interrupted Operations

2022 Midyear data breach trends

It would be an understatement to say ransomware operators have altered the breach landscape. The threat of impaired operations coupled with data theft has made ransomware into one of the most discussed issues in security today. Beginning in early 2020, the research team began tracking breaches that also resulted in an interruption to business operations. While breaches coupled with downtime is a relatively small percentage of reported breaches, these are potentially one of the most damaging types of incidents an organization can experience.

Time period	Number of breaches	Percent of total
2022 H1	232	11.7%
2021 H1	354	15.2%

Table 4: Number of breaches reported to interrupt business operations disclosed by H1, in the last two years

Severity of Breaches

2022 Midyear data breach trends

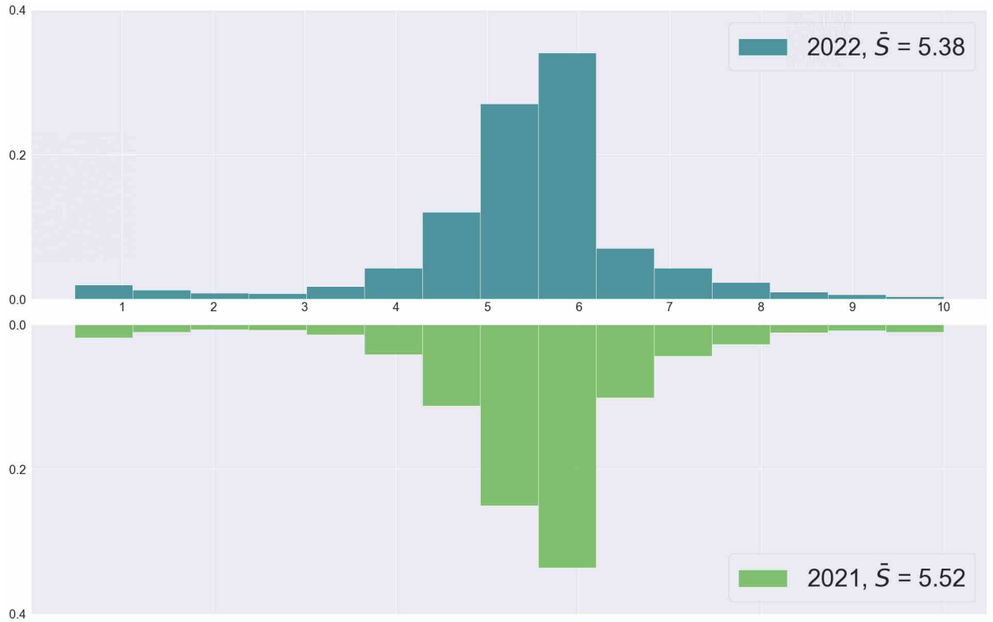

Figure 7: Severity distribution of breaches in 2022 H1

Severity score is a function of the total number of records compromised in the breach combined with other factors such as type of data lost, how the breach occurred, and follow-on events such as lawsuits filed due to the incident. Another bright spot of 2022 H1 is the distribution of scores, with the majority of breaches falling at or below an average severity of 5.38.

Records compromised	Number of incidents
Unknown	1,057
1 to 9	84
10 to 99	37
100 to 999	141
1,000 to 9,999	277
10,000 to 99,999	202
100,000 to 999,999	128
1,000,000 to 9,999,999	41
10,000,000 or above	13

Table 5: Number of incidents with records lost in these ranges reported by 2022 H1

Despite a decline in the overall number of records exposed, over 1.4. billion records were exposed in 2022 H1, including three breaches that each exposed over 100 million records. Once again, it is the "Unknown" that stands out, with approximately 53 percent of reported breaches exposing an unknown—or confirmed—number of records. It is possible as data develops this percentage will drop.

For comparison purposes, the number of breaches with unknown number of records exposed for the same time period in prior years is as follows:

Time period	Number of breaches	Percent of total
2021 H1	1,214	52%
2020 H1	934	39%
2019 H1	876	20%
2018 H1	911	36%

Table 6: Number of breaches with unknown number of records exposed, reported by H1

Types of Compromised Data

2022 Midyear data breach trends

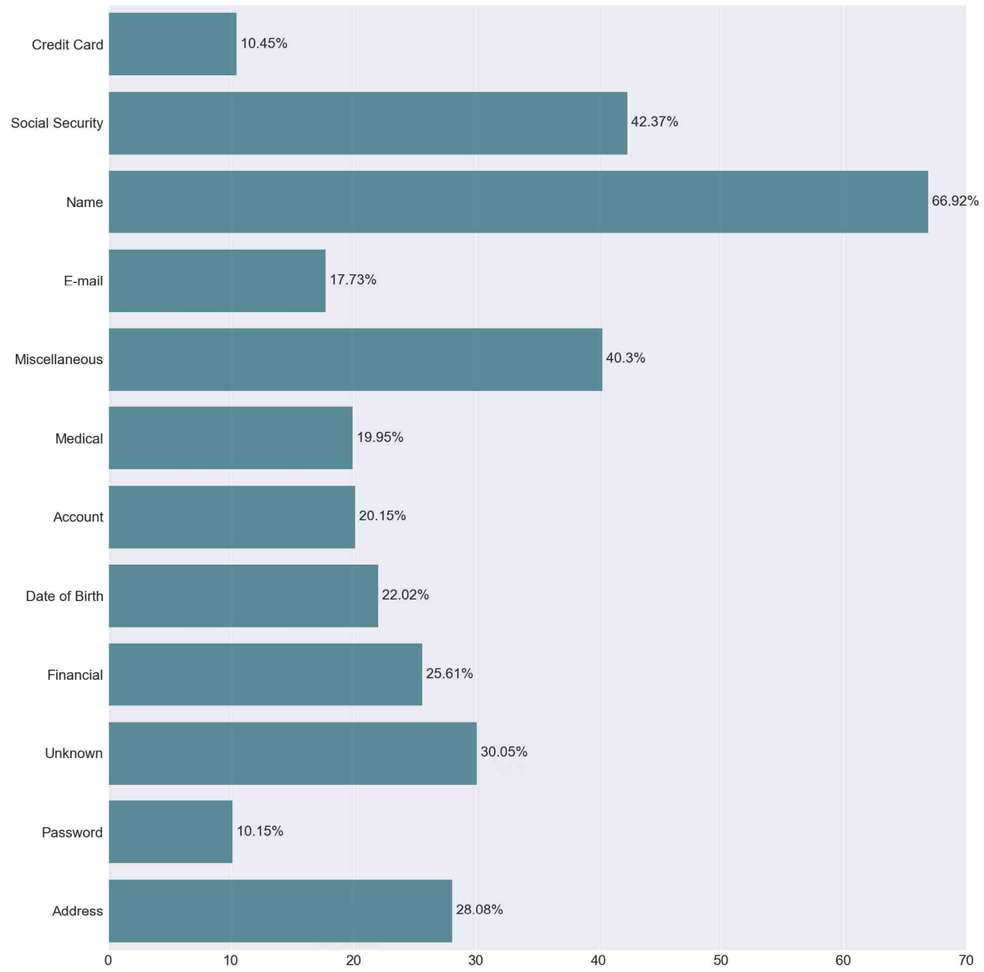

Figure 8: Data types exposed in breaches reported by 2022 H1

The types of data compromised has remained relatively unchanged since 2021. That said, two data classes stand out due to the steadily increasing percentage of breaches in which they are exposed. Those classes are Miscellaneous and Account data.

In this classification schema, Miscellaneous is used to capture data elements that may be useful for identification such as ethnicity, gender, marital status, or place of birth as well as governmental issued identifiers such as drivers' license numbers, passport numbers or taxpayer identification numbers.

Account data, not to be confused with financial data or payment card information, pertains to information such as membership ID's, billing account numbers, insurance identification numbers, and loyalty account data.

Data type	2022	2021	2020
Name	67%	65%	47%
Social Security Number (SSN)	42%	41%	28%
Misc.	40%	37%	28%
Unknown	30%	31%	24%
Address	28%	30%	23%
Financial	26%	26%	17%
Date of birth	22%	25%	16%
Accounts	20%	17%	10%
Medical	20%	15%	12%
Email	18%	19%	40%
Credit Card Number	10%	11%	12%
Password	10%	11%	33%

Table 7: Top data types lost in breaches reported by H1 for the past three yeras

Breaches by Economic Sector

2022 Midyear data breach trends

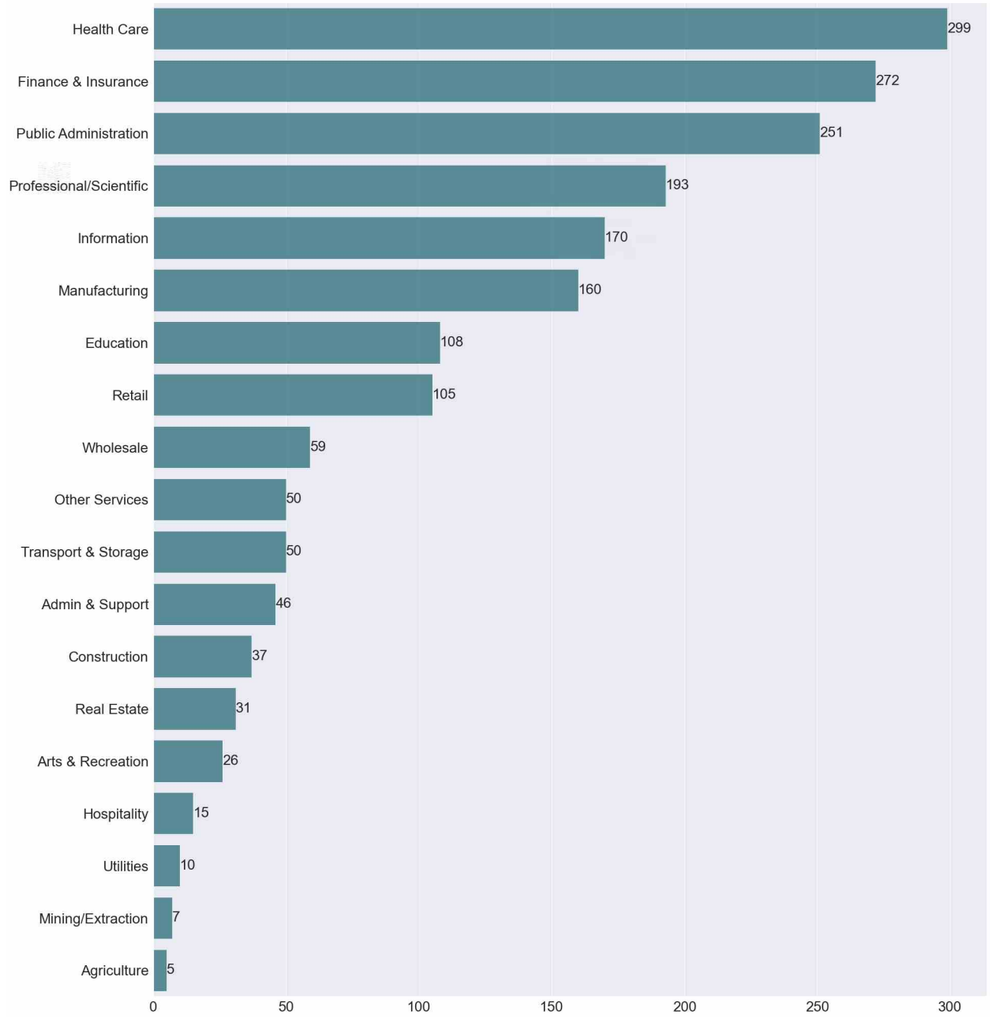

Figure 9: Number of breaches by economic sector, reported by 2022 H1

Continuing the theme of 2021, Healthcare remains at the top of the list of economic sectors experiencing the most breaches. For reference, economic sector classifications follow the 2017 North American Industry Classification System, also known as NAICS, and this graph represents the classifications of compromised organizations.

For example, the incident at Professional Finance Company (PFC) is captured here under the Administrative and Support and Waste Management and Remediation Services. So while the breach exposed medical information provided to PFC by healthcare clients, for this graph, the breach is linked to the classification of the organization that was breached.

Regular users of the NAICS system will likely recognize that while useful for classification purposes, the sectors do not align neatly with risk profile. After all, collection agencies and landfill operators, both included in the Admin & Support sector, are two very different operations. Because of these differences, each economic sector is subdivided into business groups that share similar data risk profiles. The top five most compromised business groups are:

Business group	Percentage of breaches in 2022 H1	Associated economic sector
Financial	10%	Finance and Insurance
Software Development/SaaS	6.6%	Information
Hospitals	5.3%	Healthcare
Healthcare Facilities (non-hospital)	4.8%	Healthcare
Manufacturing	4.3%	Manufacturing

Table 8: Percentage of breaches reported within various business groups and associated economic sector

Location of Breaches

2022 Midyear data breach trends

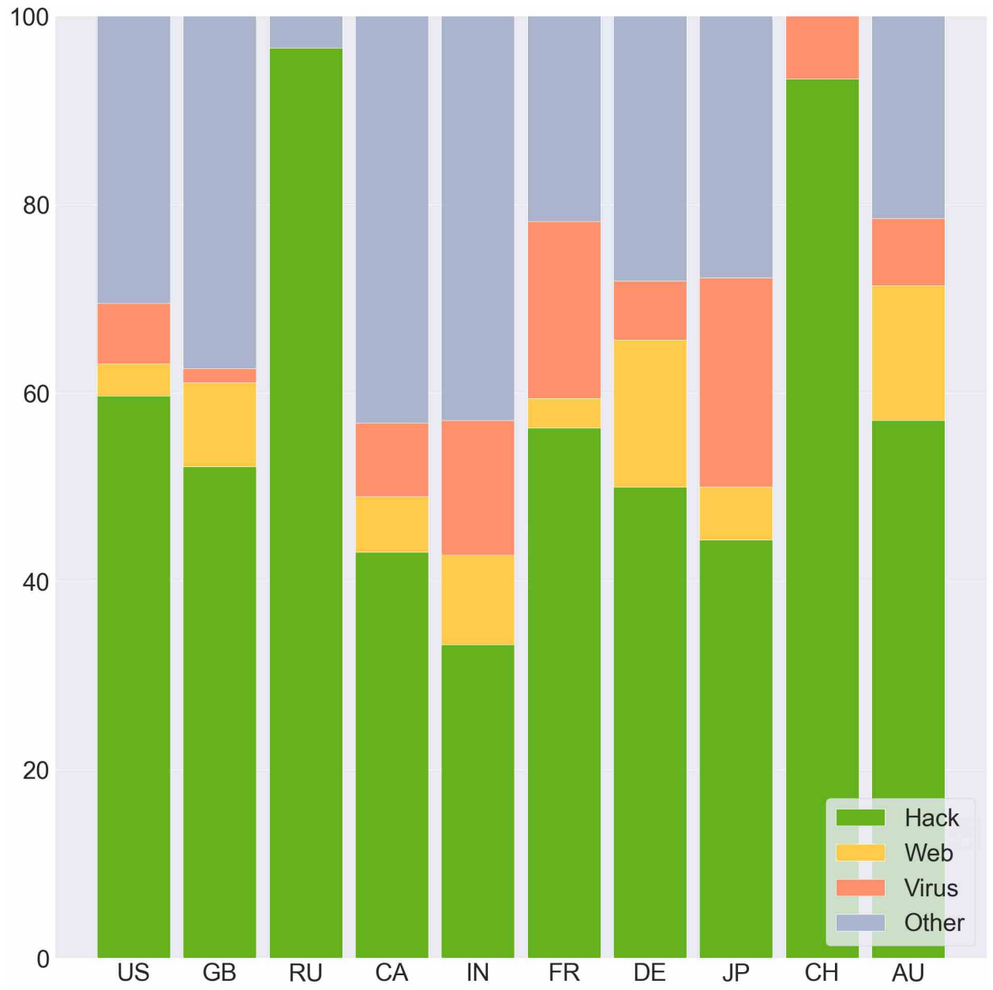

Figure 10: Distribution of breaches reported by 2022 H1

Just as any organization can experience a breach, so too can any location. Of particular interest in 2022 H1, is the increase in the number of breaches at Russian organizations. Russia's invasion of Ukraine has highlighted the new reality that war is no longer fought in physical spaces alone.

Cyber attacks from presumed state-sponsored actors, as well as from supporters on both sides have been a persistent component of the conflict. Russia has made frequent appearances in the top ten locations for breach activity—but typically in the lower quarter of the list. Propelled by war, Russia has moved into third place for breach activity.

> "Russia's invasion of Ukraine has highlighted the new reality that **war is no longer fought in physical spaces alone**."

Conclusion

After nearly two years of volatility in the breach reporting space, more predictable patterns have emerged in 2022 H1. While some resources remain slow to update, a more regular cadence of reporting has returned. This is welcome news, despite a trend toward less candid notifications. Also welcome news is the drop in the number of records exposed. It appears likely that malicious activity will continue unabated and the final tally of publicly reported breaches will likely surpass 2021. However, the corner may have finally been turned on the surfacing of large, open, unprotected datasets containing hundreds of millions of records.

The Cyber Risk Analytics research team will continue to monitor these trends and any new patterns that emerge throughout the year. Tune in for the year end report to catch up with the full 2022 breach landscape.

Methodology and terms

Flashpoint'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 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, Flashpoint requires a high degree of confidence in the accuracy of information reported, as well as the ability to reference a public source. 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.

Turn headlines into intelligence with Flashpoint

Avoid costly risk assessments while acting quickly to proactively protect your most critical information assets. Sign up for a **free Cyber Risk Analytics trial** to learn the security posture of your supply chain. Don't let security gaps of other organizations affect you.

Credits

Thank you to Inga Goddijn, Steven Weinstein, Ben Haynes, and Curtis Kang for their contributions, along with the entire Flashpoint Intelligence Team. Thank for your tireless research and analysis, which make reports like these possible.

About Flashpoint

Trusted by governments, commercial enterprises, and educational institutions worldwide, Flashpoint helps organizations protect their most critical assets, infrastructure, and stakeholders from security risks such as cyber threats, ransomware, fraud, physical threats, and more. Leading securtity practitioners—including physical and corporate security, cyber threat intelligence (CTI), vulnerability managment, and vendor risk management teams—rely on the Flashpoint Intelligence Platform, compromising open source (OSINT) and closed intelligence, to proactively identify and mitigate risk and stay ahead of the evolving threat landscape. Learn more at **www.flashpoint.io**.

FLASHPOINT