opentext™

Quantifying data risk: Visualizing financial exposure

Contents

Introduction

To truly improve data security postures, organizations must be mindful and purpose-driven in understanding where sensitive data resides, who has access to it, and how it is used. The challenge presented to many is that business environments are so vast and sprawling. Prioritizing and understanding where to start takes time. Different strategies to deploy in developing this understanding include data discovery, data sampling, and risk modeling to help assess the value of data across the data ecosystem. Imagine you could inject and visualize the financial risk exposure associated with data during these assessments and tasks.

Data discovery

Not all data is created equal—and not all sensitive data represents the same exposure risk to your organization. Data discovery evaluates data to discover what is sensitive, redundant, outdated, or trivial. Highlighting sensitive data helps target risky applications and locations where data is stored for remediation and protection. A data hygiene program can help mitigate the risks of retaining data with low or no value.

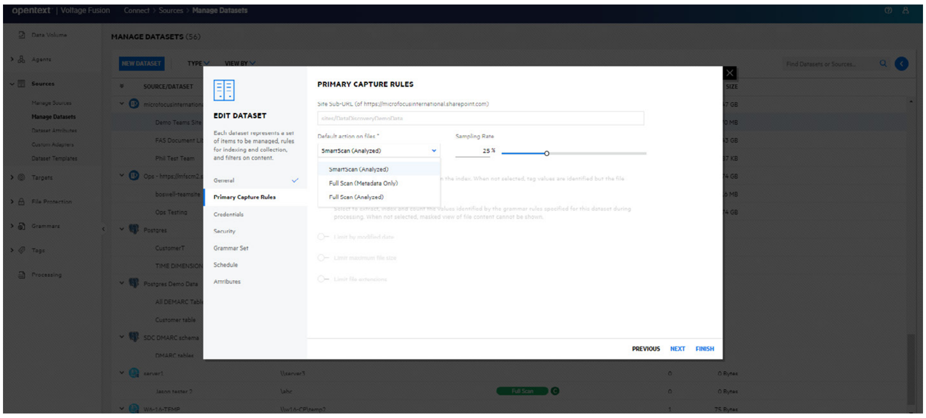
Automating data discovery can also assist in quantifying your exposure risk and help with prioritization. According to the IBM Cost of Data Breach Report 2023, key categories of compromised records include Customer PII with an average cost of US$183 per record, Employee PII costing US$181, Corporate Data costing $168, Intellectual Property costing US$156, and non-PII costing US$138.[1] With this model, a SharePoint site containing 22,000 customer records represents ~US$4 million in financial exposure. A set of local file servers containing 9,000 duplicate corporate records would represent US$1.51 million in financial risk reduction if that data was deleted.

Automating data discovery can also assist in quantifying your exposure risk and help with prioritization. According to the IBM Cost of Data Breach Report 2023, key categories of compromised records include Customer PII with an average cost of US$183 per record, Employee PII costing US$181, Corporate Data costing US$168, Intellectual Property costing US$156, and non-PII costing US$138.[1]

Data sampling

Organizations do not have the time or budget to analyze every kilobyte of data. Intelligent sampling techniques can identify high saturation areas containing sensitive data, giving your security and compliance teams visibility into where to begin.

Metadata scans are not enough to highlight risk, and regular expression-based analysis results in too many false positives and insufficient contextual understanding of data. Intelligent data sampling techniques help prioritize risks based on their potential impact and sensitivity level as "hotspots" by only analyzing a small percentage of data. These techniques allow organizations to allocate resources, spend effectively, and address the most critical issues.



EDIT DATASET

Each dataset represents a set of items to be managed, rules for indexing and collection, and filters on content.

PRIMARY CAPTURE RULES

Site Sub-URL (of https://microfocusinternational.sharepoint.com): sites/DataDiscoveryDemoData

Default action on files *: SmartScan (Analyzed)

Sampling Rate: 25 %

Figure 1: Data sampling screenshot – OpenText Cybersecurity

[1] IBM, Cost of a Data Breach Report 2023

> If you look at the financial element of risk models, as referenced in the IBM report, the average data breach in 2023 was US$ 4.45 million.[2] While that model is notable overall in representing the breadth of the challenge data breaches impose, financial risk modeling needs to go beyond simple linear models and quantify risk more pragmatically.

Risk modeling

There are different ways to model risk around enterprise data. Risk scoring can help an organization gauge the riskiness of an individual record or file that contains one or many sensitive data entities. It can prioritize data discovery tasks for clean-up, risk assessments, or data protection activities. Risk scores can, however, be skewed by false positives and poor implementation practices.

If you look at the financial element of risk models, as referenced in the IBM report, the average data breach in 2023 was US$4.45 million.[2] While that model is notable overall in representing the breadth of the challenge data breaches impose, financial risk modeling needs to go beyond simple linear models and quantify risk more pragmatically.

How can you quantify financial risk exposure?

Many factors are at play when assessing the cost of a breach, including the costs associated with data loss, detection and escalation, post-breach response, notification, and the lost business cost.

Risk on a linear scale

IBM's Cost of Data Breach Study 2023 looks at the average breach loss on a linear scale. In this case, the average cost of a breach is US$4.45 million. The cost breakdown for 2023[3] was:

- ■ 36 percent of the expenses are sunken into breach detection and escalation—up 8 percent from 2018.
- ■ 27 percent of costs are based on post-breach response—up slightly from 2018.
- ■ 8 percent of costs are in breach notification—up 4 percent percent from 2018.
- ■ 29 percent of the cost can be associated with lost business—down 8.5 percent from 2018.

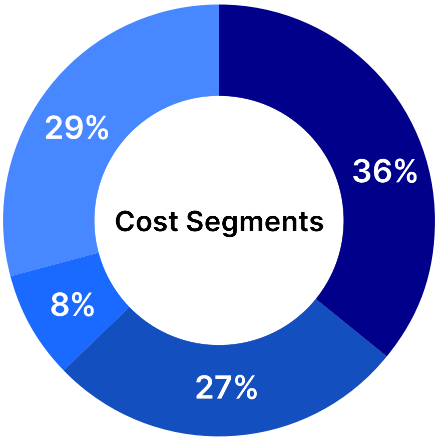


Figure 2: Data breach cost segments

[2] IBM, Cost of a Data Breach Report 2023

[3] Ibid.

Representing potential for financial loss based solely on the number of records lost is likely too simple to be meaningful. Evidence today shows that representing loss on a linear line underestimates the cost of smaller breaches and consequently overestimates the cost of large breaches.

Not all breaches are created equal: the risk is on a logarithmic scale

Representing the potential for financial loss based solely on the number of records lost is likely too simple to be meaningful. Evidence today shows that representing loss on a linear line underestimates the cost of smaller breaches and consequently overestimates the cost of large breaches. The Cytenia Institute's 2022 Information Risk Insights Study uses a model that makes it easier to visualize breach data, analyze trends, and accurately represent financial losses. The Cytenia study also provides cross-industry estimates on breach events' size, frequency, and probability.

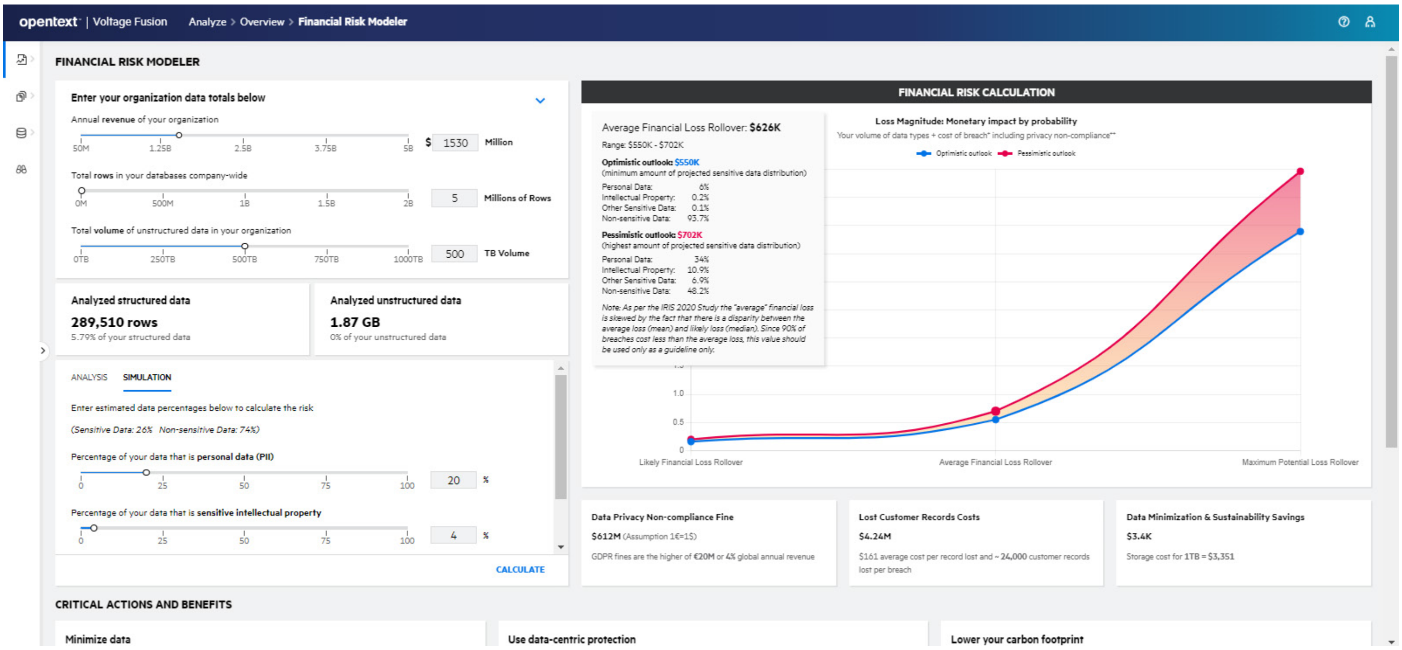
The Cytenia study estimates, based on historical losses, the likely breach cost is US$200,000, but 10 percent of breaches are extreme and exceed US$20 million.[4] On the end of the spectrum, they found Fortune 250 organizations' financial exposure risk is US$100 million or greater.[5] That is a large disparity when comparing data across this wide range. As a result, Cyentia uses a log-normal scale to represent a breach's probability and financial implications. This model provides better visibility than a linear model that loses the magnitude of the upper end and poorly represents the actual probable losses.

Interpreting financial risk exposure

The IBM study does a good job of assessing the value of lost records and even the breakdown of cost centers associated with a data breach. They also made strides in the 2023 study looking at breaches across industries and geographic regions, yet the average cost of a breach is still laid out on a linear line. However, Chief Security Information Officers and C-level executives still need more depth to assess the financial exposure around the data they process, use, and manage.

[4] Cyentia Institute, IRIS 2022 Information Risk Insights Study

[5] Ibid.


opentext | Voltage Fusion
Analyze > Overview > Financial Risk Modeler
50M
1.25B
2.5B
3.75B
5B
0M
500M
1B
1.5B
2B
0TB
250TB
500TB
750TB
1000TB
ANALYSIS
SIMULATION
0
25
50
75
100
CALCULATE
FINANCIAL RISK CALCULATION
Optimistic outlook
Pessimistic outlook
1.0
0.5
0
Likely Financial Loss Rollover
Average Financial Loss Rollover
Maximum Potential Loss Rollover

FINANCIAL RISK MODELER

Enter your organization data totals below

Annual revenue of your organization: $ 1530 Million

Total rows in your databases company-wide: 5 Millions of Rows

Total volume of unstructured data in your organization: 500 TB Volume

Analyzed structured data
289,510 rows
5.79% of your structured data

Analyzed unstructured data
1.87 GB
0% of your unstructured data

Enter estimated data percentages below to calculate the risk

(Sensitive Data: 26% Non-sensitive Data: 74%)

Percentage of your data that is **personal data (PII)**: 20 %

Percentage of your data that is **sensitive intellectual property**: 4 %

Average Financial Loss Rollover: **$626K**

Range: $550K - $702K

Optimistic outlook: $550K
(minimum amount of projected sensitive data distribution)

Personal Data:	6%
Intellectual Property:	0.2%
Other Sensitive Data:	0.1%
Non-sensitive Data:	93.7%

Pessimistic outlook: $702K
(highest amount of projected sensitive data distribution)

Personal Data:	34%
Intellectual Property:	10.9%
Other Sensitive Data:	6.9%
Non-sensitive Data:	48.2%

Note: As per the IRIS 2020 Study the "average" financial loss is skewed by the fact that there is a disparity between the average loss (mean) and likely loss (median). Since 90% of breaches cost less than the average loss, this value should be used only as a guideline only.

Loss Magnitude: Monetary impact by probability

Your volume of data types + cost of breach* including privacy non-compliance**

Data Privacy Non-compliance Fine
$612M (Assumption 1€=1$)
GDPR fines are the higher of €20M or 4% global annual revenue

Lost Customer Records Costs
$4.24M
$161 average cost per record lost and ~ 24,000 customer records lost per breach

Data Minimization & Sustainability Savings
$3.4K
Storage cost for 1TB = $3,351

CRITICAL ACTIONS AND BENEFITS

Minimize data

Use data-centric protection

Lower your carbon footprint

Figure 3: Financial risk modeler screenshot – OpenText Cybersecurity

Since 2012, incidents of data breaches have increased 44.7%.[6] Some industries also have a disproportionate number of breach events as well. For instance, the Healthcare and Financial Services sectors, according to Cytenia, have 76 times the incidents of corresponding Mining or Agriculture sectors.[7]

Probability and frequency of a data breach

Since 2012, incidents of data breaches have increased 44.7 percent.[6] Some industries also have a disproportionate number of breach events as well. For instance, the Healthcare and Financial Services sectors, according to Cytenia, have 76 times the incidents of corresponding Mining or Agriculture sectors.[7]

Additionally, this report estimates Fortune 250 (organizations with greater than US$100B in revenue) have a greater than 29 percent chance of more than one breach event in a year.[8] In comparison, smaller firms (organizations with between US$10M and US$100M in revenue) have a less than 0.5 percent chance of a breach.[9]

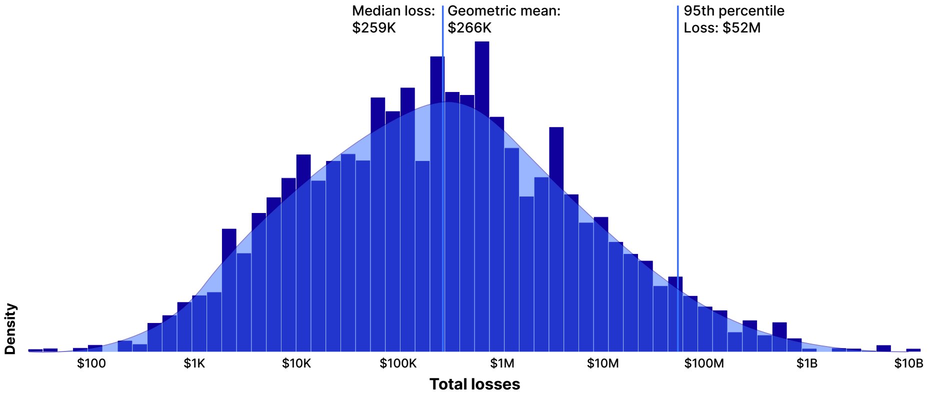


Figure 4: Distribution of reported losses for security incidents from 2012 to 2021[10]

In the same report, Cytenia indicates that the industry your organization belongs to can also suggest loss frequency. Using the public sector as the baseline for comparing breach events across industries, Hospitality, Financial Services, Retail, and Healthcare have a notably greater probability of greater than one breach event annually. Other sectors like Education, Real Estate, Manufacturing, and Construction have a probability of less than one breach event per year.

Mitigating the risk of managing sensitive data

Mitigating the risk and remediating exposed data during data discovery phases is essential to reducing threat landscapes, operating costs, and financial risk exposure while reducing the impact of a breach event.

So, how can you manage data risk?

1. **Identify and classify data and exposed data sources**

   Sensitive data sprawl is widespread across email systems, file shares, cloud repositories, business applications, and databases. You should act today to understand where sensitive data is and prioritize your threat remediation tasks based on risk tolerance. By improving data visibility, you increase your ability to respond to breach events, prioritize data remediation, and make better, informed decisions that can drive innovation and growth.

   Capitalize on this insight with discovery solutions that help streamline data protection, preserve privacy, and ensure compliance. Additionally, proactive approaches around data discovery and classification can help adapt to new cybersecurity threats as well as augment and streamline regular security assessments.

[6] Cyentia Institute, IRIS 2022 Information Risk Insights Study

[7] Ibid.

[8] Ibid.

[9] Ibid.

[10] Ibid.

Breaches will happen and you will have to investigate them – but proper discovery and protection techniques help offset that disruption and minimize the associated cost by **≥64%** for the event.[11]

2. Act on legacy data

If you uncover data that no longer has business value, has reached its required retention period, or is duplicated it's time to deal with it. Review the information with business owners, delete what isn't valuable to the business, and archive sensitive data into a long-term repository so it can be protected and managed more effectively over time. These approaches reduce operation and run-time costs and lower the threat landscape for breach events.

3. Protect data that is being used

Data can be vulnerable to attacks while being processed or accessed by applications. As a part of the post-discovery process, there are some valuable techniques to support Zero Trust security practices. Encrypt and tokenize data using format-preserving encryption (FPE), ensuring that protection follows the data persistently (in lead applications and databases), consider using secure and isolated environments for processing data (e.g., secure enclaves), implement tight access controls enforcing least privilege models, and de-identify sensitive data used in your analytics and AI pipelines in building greater risk tolerance and resiliency.

Percentage of Cost[12]	36%	27%	8%	29%
Mitigating strategies	Detection and escalation	Post-breach response	Breach notification	Lost business
Identify and classify data and exposed data sources	■			■
Act on legacy data	■			■
Protect data that is being used		■	■	■
		← 64% →		

Figure 5: Risk mitigating strategies

Key takeaways for risk and security managers:

- More than assessing flat cost per record is required to identify potential financial losses and to evaluate your true financial risk exposure.
- Although most breaches fall into the lower end of the financial loss model, you shouldn't ignore the possibility of severe financial losses in uncommon and extreme cases.
- Based on your industry and the size of your organization, predicting the probability of a breach by sector is difficult. However, financial risk and event probability can be a guardrail for assessing security postures and risk tolerance.
- Look for tools that help assess risk and financial exposure to help prioritize your data security practices and breach defense.
- Data protection techniques like data de-identification, encryption, and tokenization can greatly reduce the cost of a breach by minimizing the time required to respond to the event, breach notification, and downstream impacts of lost business.
- Breaches will happen and you will have to investigate them – but proper discovery and protection techniques help offset that disruption and minimize the associated cost by **≥64 percent** for the event.[11]

[11] Cyentia Institute, IRIS 2022 Information Risk Insights Stud

[12] IBM, Cost of a Data Breach Report 2023

opentext™

Learn More

https://www.opentext.com/products/voltage-fusion-platform

https://www.microfocus.com/en-us/products/voltage-fusion/request-demo

Further guidance

The relationship between financial risk exposure and data security is symbiotic. A data breach can lead to significant financial losses, damage to brand reputation, and legal consequences. On the other hand, financial instability can hinder an organization's ability to invest in and maintain robust cybersecurity measures, leaving it vulnerable to cyber threats. Therefore, a holistic risk management strategy should address both your financial and cybersecurity risks.

As illustrated, navigating the nexus between breach readiness and risk exposure can be challenging. At OpenText Cybersecurity, our Data Security and Identity and Access Management solutions help customers assess financial risk exposure, improve breach defense, and safeguard sensitive data with industry-leading data discovery, protection, and access governance capabilities.

About OpenText

OpenText, The Information Company, enables organizations to gain insight through market leading information management solutions, on premises or in the cloud. For more information about OpenText (NASDAQ: OTEX, TSX: OTEX) visit: opentext.com.

Connect with us:

- OpenText CEO Mark Barrenechea’s blog
- X (formerly Twitter) | LinkedIn

opentext.com/contact