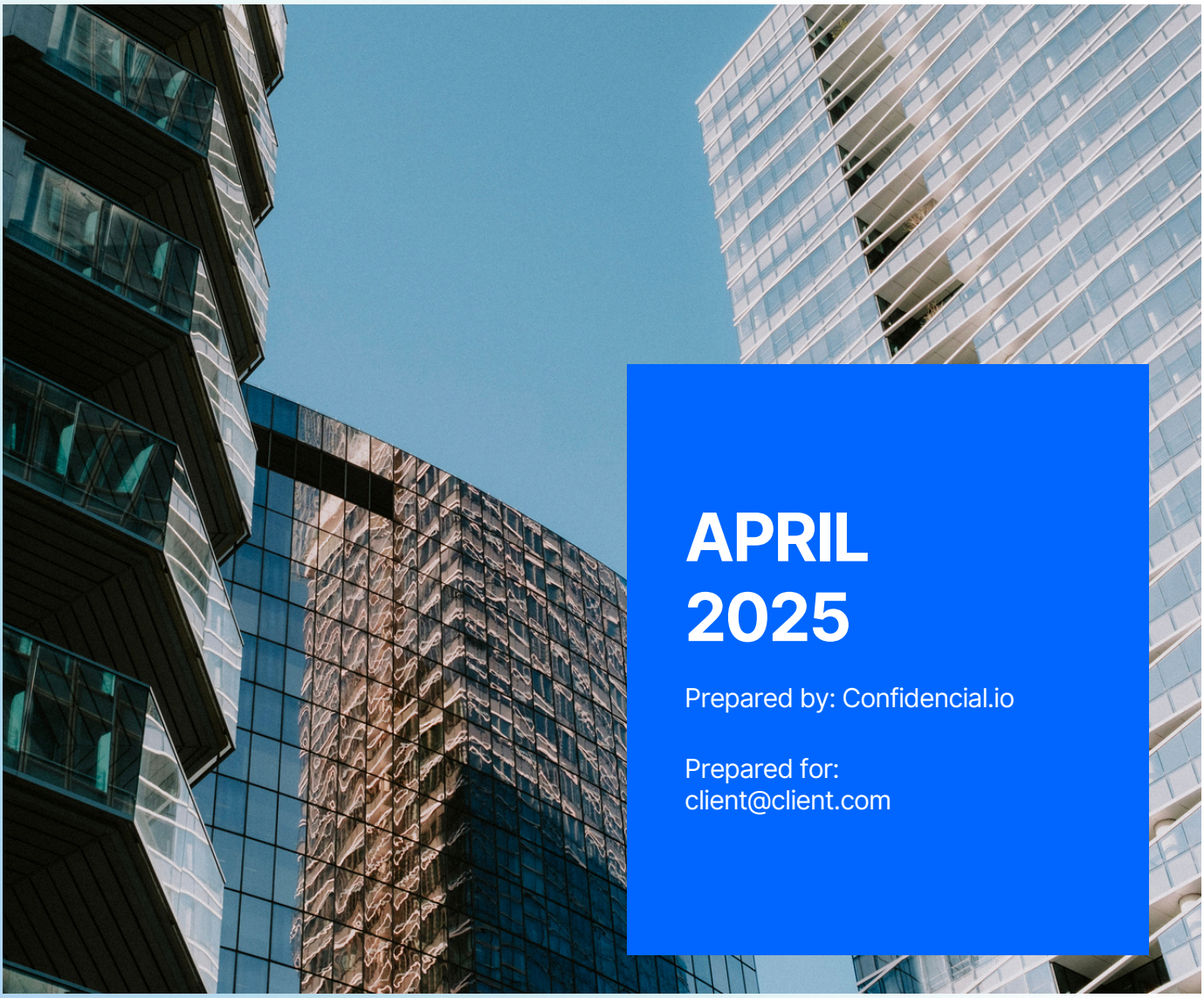
Confidencial

Data Assessment



APRIL 2025

Prepared by: Confidencial.io

Prepared for:
client@client.com

Who we are

Confidencial was born from DARPA-funded research at SRI, developed by a team of cybersecurity engineers and PhDs dedicated to unlocking the massive potential of unstructured data—safely and securely. Built to meet the military's need for protecting sensitive and proprietary data, our technology was forged under the highest standards of resilience and control. With millions invested across two DARPA projects, SRI recognized the broader market opportunity and spun out Confidencial in 2021, exclusively licensing all related IP to us. Today, we're bringing that same battle-tested security to organizations seeking to harness AI's power—without compromise, only control.

Our Mission

Our mission is to be woven into the fabric of how organizations secure sensitive information within unstructured data. Confidencial bridges the critical gap between identifying sensitive data at a granular level and truly protecting it—supporting and accelerating the entire journey from discovery to control. By making data secure and usable, we empower organizations to confidently unlock the full potential of their most valuable information, enabling AI workflows and pipelines to act responsibly and intelligently.

Confidencial

Assessment Overview

Confidencial

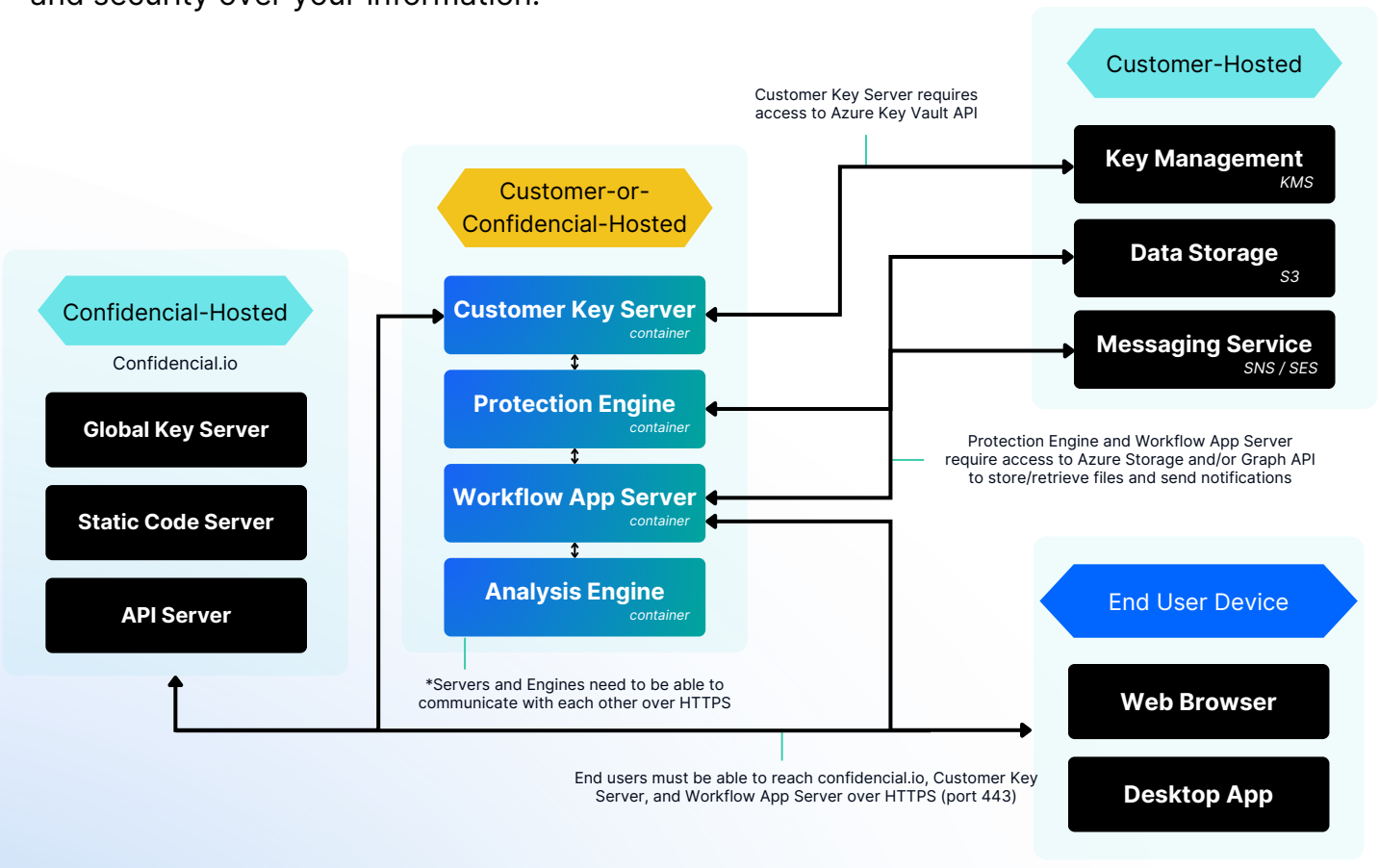
Your privacy comes first

Data Assessment Retention and Deletion Policy

During your scan, Confidencial collects aggregate data to support reporting and analysis. This includes file and folder names, along with summary counts of detected sensitive data types. However, no actual sensitive data is stored—only volume metrics are retained.

Once the scan is complete, your independent database is permanently deleted, ensuring that all recorded counts are fully erased from our system.

With the purchased version of Confidencial, no data is stored on Confidencial's servers. Instead, all data remains within your own hosted environment, maintaining full control and security over your information.



Confidencial

Executive Summary

This report provides an in-depth analysis of the data discovered across your connected storage environment, evaluates associated risks, and delivers actionable insights to prioritize and address critical vulnerabilities, enhancing data protection, ensuring compliance, and reducing financial and operational exposure.

This assessment **reveals a high risk associated with sensitive data** stored in files across your connected storage environment. Health Insurance Portability and Accountability Act (HIPAA), Personally Identifiable Information (PII) and Payment Card Information (PCI) were found in PDF and word documents.

Based on our analysis of your data and industry benchmarks, your organization faces an estimated financial exposure of $3.5 million due to potential data breaches, including regulatory fines, legal expenses, and operational costs.

**The total dollar figure represented above is broken down on page 13.*

Confidencial

Scope of the Analysis

WHAT WE SCANNED:

1 Data Source

Name:

Type:

Path:

X

Detected Suites

PII, HIPAA, GDPR

xGB

Unstructured Data

X

Folders

X

Documents

Confidencial

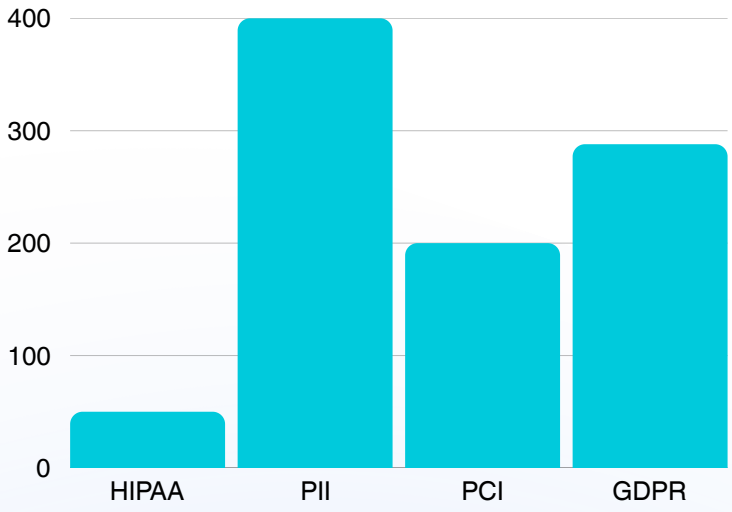
High-Level Results

308

documents containing sensitive information

4,063

sensitive data objects discovered



The assessment revealed that your connected data source primarily contained Personally Identifiable Information (PII), such as names and Social Security numbers, followed by GDPR-regulated data, including email addresses and financial details. PCI and PHI were less common, with HIPAA-regulated data found in the least amount.

Discovered Data Objects

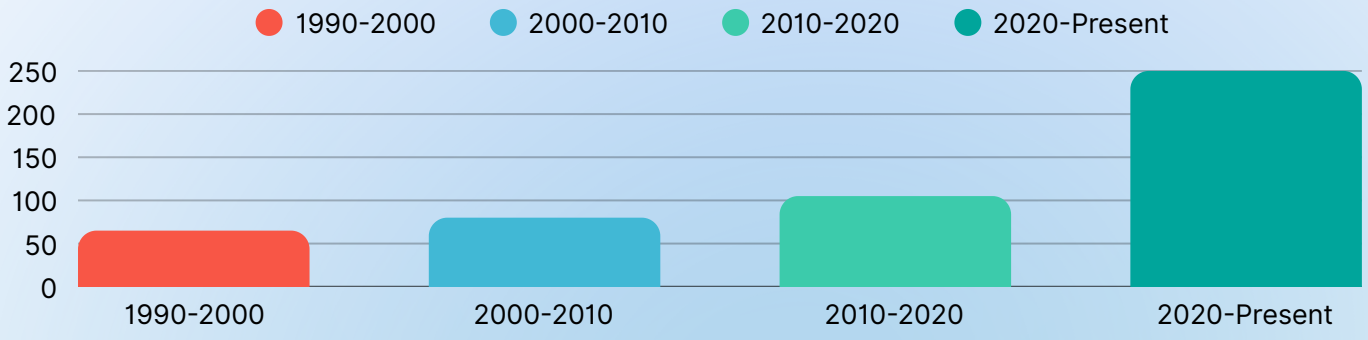
55 Personally Identifiable Information	30 Payment Card Industry
15 Protected Health Information	5 Credit Cards
55 General Data Protection Regulation	7 US Social Security Numbers

Personally Identifiable Information (PII)

This section exposes the full scope of your PII risk—breaking down the types of sensitive data discovered, their volume, where they're most concentrated, and the age of those documents. It's a clear roadmap to your most vulnerable points.

Label	Count	Hotspots
Name		
Date of Birth		
Email		
Social Security Number		
Drivers License		
Email		
Gender		
Religion		

Sensitive Data in Aging Documents

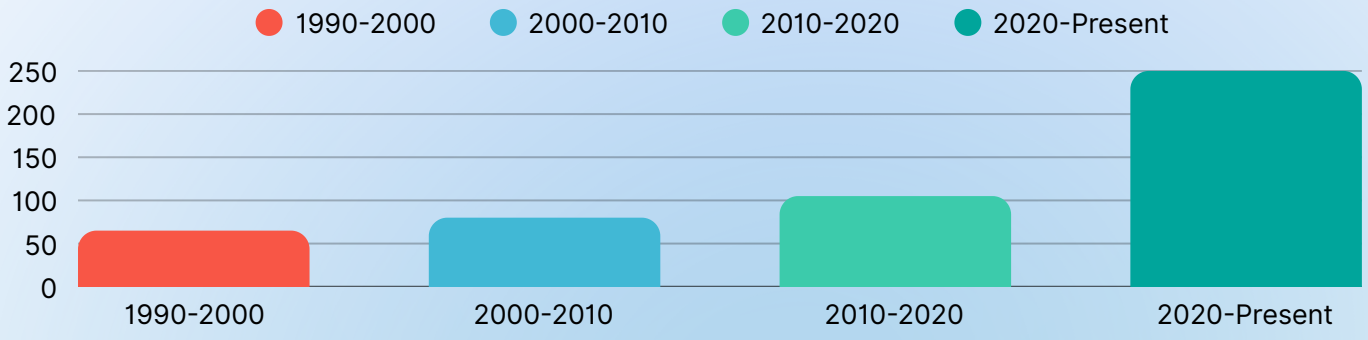


General Data Protection Regulation (GDPR)

This section exposes the full scope of your GDPR risk—breaking down the types of sensitive data discovered, their volume, where they're most concentrated, and the age of those documents.

Label	Count	Hotspots

Sensitive Data in Aging Documents

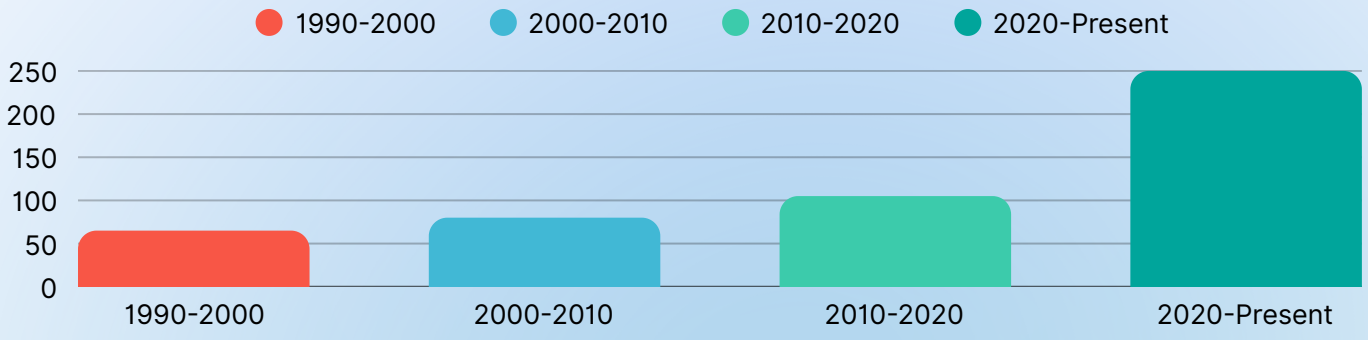


Protected Health Information (PHI) & HIPAA

This section exposes the full scope of your PHI & HIPAA risk—breaking down the types of sensitive data discovered, their volume, where they're most concentrated, and the age of those documents.

Label	Count	Hotspots
Patient Name		
Date of Admission		
Medical Records		
Health Plan Beneficiary		
Finger Print		
Name of Employer		

Sensitive Data in Aging Documents

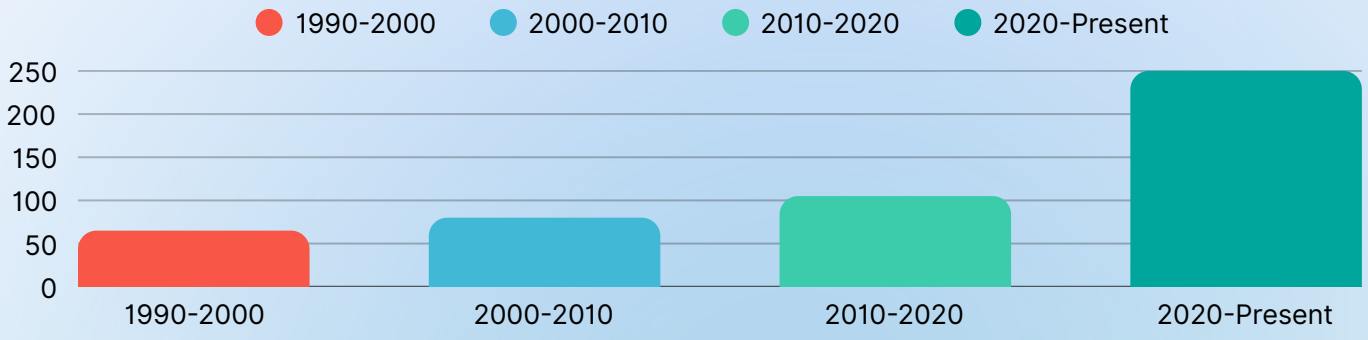


Payment Card Industry (PCI)

This section exposes the full scope of your PCI risk—breaking down the types of sensitive data discovered, their volume, where they're most concentrated, and the age of those documents.

Label	Count	Hotspots

Sensitive Data in Aging Documents



Confidencial

Sensitive Data Distribution

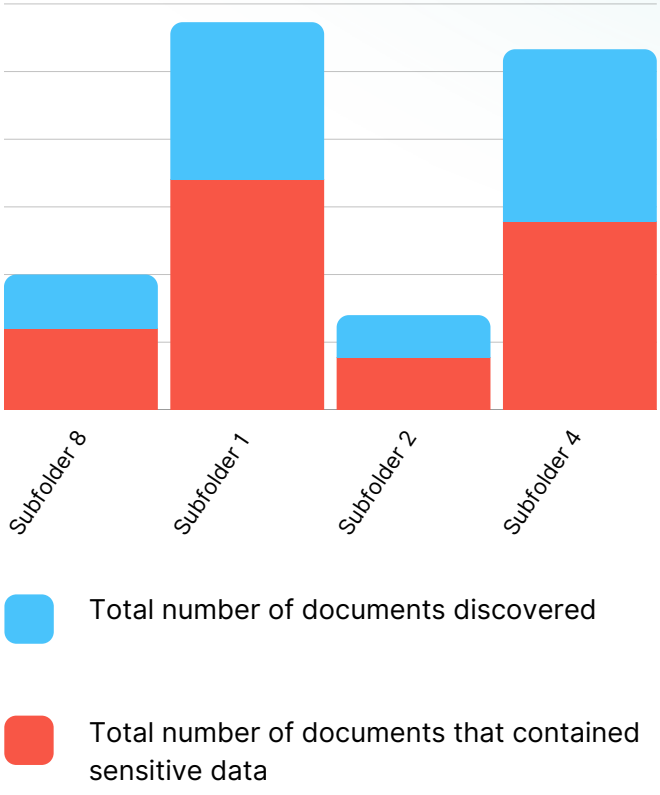
This section pinpoints exactly where your sensitive data resides across connected storage—categorizing file formats like Word, PDF, and Excel.

Where is sensitive data stored?

The analysis showed that the majority of sensitive data in your connected data source was concentrated in subfolder 1, making it the largest point of exposure. Subfolder 4 contained the second-highest volume, followed by subfolder 8, while the remaining data was more evenly distributed across other areas.

Top Locations with Highest Concentration of Sensitive Data:

1. Subfolder 1
2. Subfolder 4
3. Subfolder 8
4. Subfolder 2

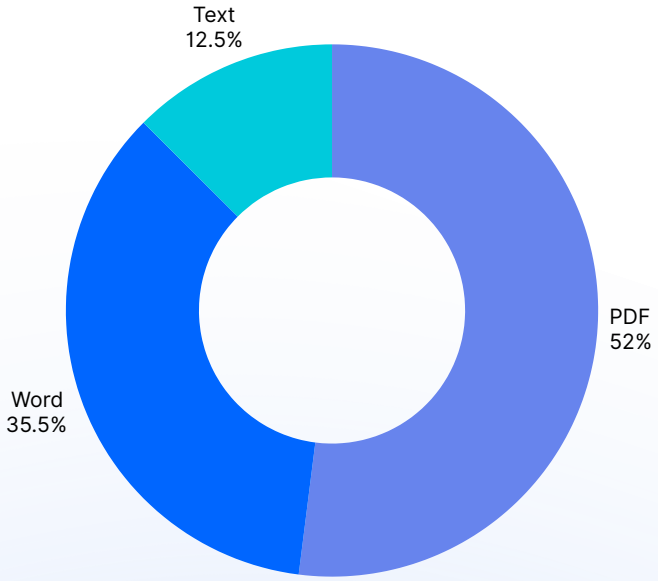


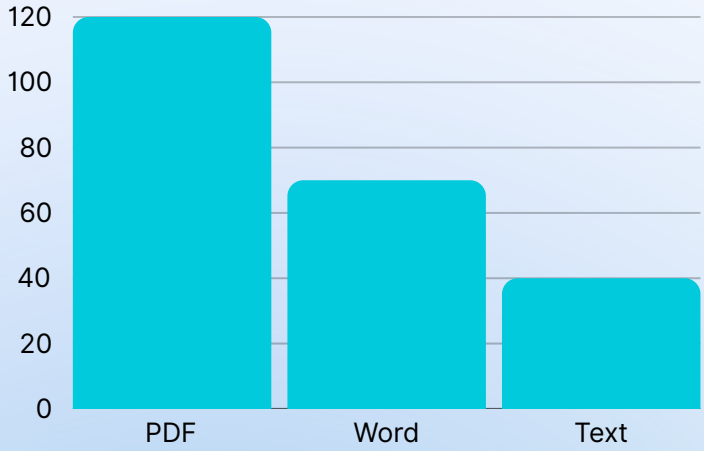
Confidencial

Document Formats

Your connected data source primarily contained sensitive data in PDFs (120, 52% of all sensitive documents), Word documents (70, 35.5%), and text files (40, 12.5%). This suggests a mix of finalized reports, internal documents, and structured text-based information.

As PDFs are the biggest hotspot for sensitive data in this data source, it could be an area of focus to inspect the security of workflows that produce and consume PDFs within your business.





Confidencial

Financial Risk Projections

ESTIMATED PROJECTION:

$3.5M

A breach isn't just a security failure—it's a financial catastrophe. Black market exploitation, regulatory fines, lawsuits, response costs, and reputational damage can cripple growth and drain revenue. Below is an estimate of these costs based on industry benchmarks.

Black Market Ransom

Estimate demand for payment by cybercriminals in exchange for not selling or exposing stolen sensitive data on illegal marketplaces.

Credit Card Risk	$ 1,200,300
SSN Risk	$ 234.678

Identification & Containment

Processes of detecting a data breach and taking action to isolate affected systems, preventing further exposure and limiting impact.

Labor Risk	$ 45,700
Cost Risk	$ 56,798

Fines & Lawsuits

Financial penalties from regulatory non-compliance and legal expenses resulting from data breaches or misuse of sensitive information.

PII Risk	$ 2,340,987
HIPAA Risk	$ 245,987
PCI Risk	$ 45,700
GDPR Risk	$ 234.678

Incident Response

Time, labor, and resources required to investigate, contain, and remediate a data breach, and restore operations.

Labor Risk	$ 234.678
Cost Risk	$ 245,987

Source 1: https://www.ibm.com/reports/data-breach
Source 2: https://www.comparitech.com/blog/vpn-privacy/dark-web-prices
Source 3: https://www.ekransystem.com/en/blog/cost-of-a-data-breach
Source 4: https://www.hipaasecurenow.com/breach-cost-calculator
Source 5: https://blog.rsisecurity.com/pci-fines-and-penalties-for-non-compliance

Confidencial

You've identified the risks, now it's time to act.

We know that identifying your risks is only half the battle. Eliminating them is where true security begins.

Confidencial makes remediation fast, effective, and seamless—and is critical to facilitating both remediation and ongoing monitoring. From protecting sensitive data to strengthening defenses and maintaining continuous visibility, Confidencial drives results across the entire incident response lifecycle.

Let’s breakdown how...

Confidencial

Confidencial Powers Your Journey From Discovery to Remediation

STEP 1

Govern

Confidencial establishes and enforces automated policy frameworks tailored for unstructured data across hybrid environments. Supports compliance requirements like GDPR and HIPAA by enabling granular policy creation, enforcement, and auditing. Ensures only the right individuals and teams are granted access to sensitive data, down to the data type, compliance suite, or customized policy of your choosing. Policies can be customized to address data access, encryption requirements, retention schedules, and remediation workflows.

STEP 2

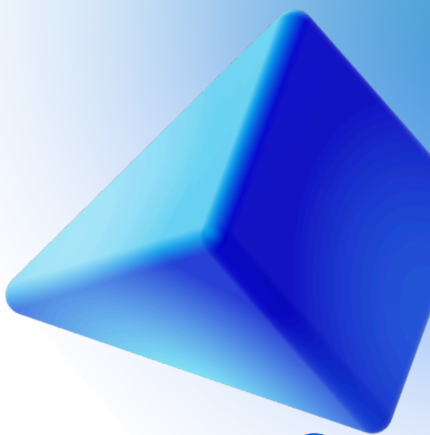
Identify

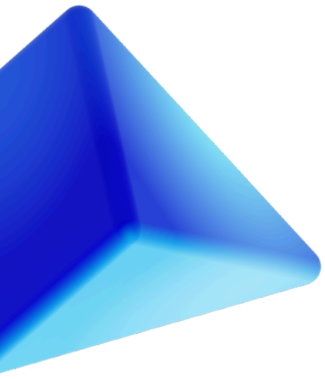
Confidencial continuously scans and indexes unstructured sensitive data across on-premises, multi-cloud environments to discover, classify, and label sensitive data (PII, IP, financial data) to ensure compliance with frameworks. Leverages a multilayered analysis engine incorporating regex, natural language processing (NLP), machine learning, and AI for precise and accurate identification and categorization.

STEP 3

Protect

Confidencial applies selective encryption through policy-based automation, using strong encryption standards (RSA-2048/4096 and AES-GCM-256) and dynamic access controls to secure high-risk data, preventing unauthorized access or exfiltration. Selective encryption secures only sensitive portions of documents—down to specific fields or data types—while keeping the rest accessible. Encryption keys are stored in metadata, enabling different individuals or teams to access different parts of the same document based on permissions. This approach ensures seamless collaboration and sharing without disrupting workflows, while enforcing protection automatically as data is created.





STEP 4

Detect

Confidencial continuously monitors for unauthorized access attempts, unusual data movement, and abnormal sharing behaviors that may indicate insider threats or credential compromise. Provides real-time visibility into unprotected sensitive data across the organization and integrates with cyber threat intelligence feeds for enhanced risk detection. Supports integration with SIEM/SOAR tools to correlate data risks with broader security threats, ensuring comprehensive detection and response.

STEP 5

Respond

Confidencial employs built-in revocation mechanisms to protect sensitive data against insider threats and unauthorized access. Even if data is exfiltrated, it remains inaccessible and unusable due to persistent, policy-enforced encryption. Confidencial generates detailed access logs and audit trails, enabling rapid forensic analysis, compliance verification, and incident response. Its proactive encryption neutralizes breaches before they can cause harm, ensuring data security remains intact.

STEP 6

Recover

Confidencial ensures your business can bounce back faster and with less disruption. Even when attacks occur, your protected unstructured data remains secure, intact, and accessible—minimizing operational downtime, revenue loss, and compliance risks. Confidencial safeguards continuity by preventing data leaks, unauthorized use, and unnecessary setbacks. With detailed activity logs and validation reports, you have proof of ongoing protection, making recovery less painful and your rebound far smoother than it would be without Confidencial.

Confidencial

Where to Start

Recommendations

Stop guessing. Start protecting.

Confidencial gives you real-time risk visibility and automated protection—see it in action.

confidencial.io

linkedin.com/company/confidencial-inc/

hello@confidencial.io