PKWARE

CASE STUDY

Financial services company simplifies regulatory compliance with automated PII discovery and masking

Overview

We've all heard the standard customer-support call message that says, "This call is being recorded."
In the case of a financial services company, those recorded audio files are converted to text and stored. The raw files contain information that can ultimately be used to improve customer service; however, the files also contain highly sensitive personal information including customer names and addresses, device serial numbers, and payment card data. By storing such personally identifiable information (PII) in data lakes or databases, the financial services company could make itself liable and exposed.

That raw data comes in variant forms: It can be structured, semi-structured, or unstructured. More than 16,000 employees oversee 12 PB of data in 18 countries, and the data grows daily.

The company stores multiple petabytes of data, and terabytes of new data are entering the lake on a daily basis, making it is physically impossible to address any data privacy law without first knowing where the sensitive or personal data resides, and then being able to protect and monitor it on an ongoing basis.

As the financial services company saw the emerging General Data Protection Regulation (GDPR) European Union law and the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA), it knew these regulations were not to be taken lightly. Fines for non-compliance with GDPR can total up to 4 percent of a firm's annual global revenue. Due to the nature of its business, the company not only had to ensure data compliance for itself, but also needed to mitigate any risk for its large carrier clients. Just one breach—accidental or otherwise—could trigger a trickle-down effect that could damage all parties' reputations.

At the end of the day, the company had a tall but important order: to ensure that it could account for all sensitive data and that no PII was stored in its Amazon S3 data lake.

PKWARE

- Identifies and locates sensitive personal data across repositories, including on-premises and Amazon S3 and Amazon Redshift repositories
- Protects consumer privacy by masking, encrypting, and pseudonymizing data
- Scales along with Amazon Web Services (AWS) infrastructure

About PKWARE

PKWARE offers the only data discovery and protection solution that locates and secures sensitive data to minimize organizational risks and costs, regardless of device or environment. Our ultra-efficient, scalable software is simple to use on a broad range of data types and repositories, enabling precise, automated visibility and control of personal data, even in the fastest-moving, most complex IT environments. With more than 1,200 customers, including many of the world's largest financial institutions, retailers, healthcare organizations, and government agencies, PKWARE continues to innovate as an award-winning global leader in data discovery, security, and compliance. To learn more, visit PKWARE.com.

Challenge

A financial services company needed to discover, protect, and monitor personally identifiable information (PII) stored across all its AWS data repositories to prepare for data protection laws that are country- and state-specific, including GDPR, CCPA, and PIPEDA.

Results

- Established process and policies to meet privacy regulations including GDPR, CCPA, and PIPEDA
- Scanned 12 PB of data—and growing
- Consistently finds "no results" in searches for PII, indicating that the system is clear of PII

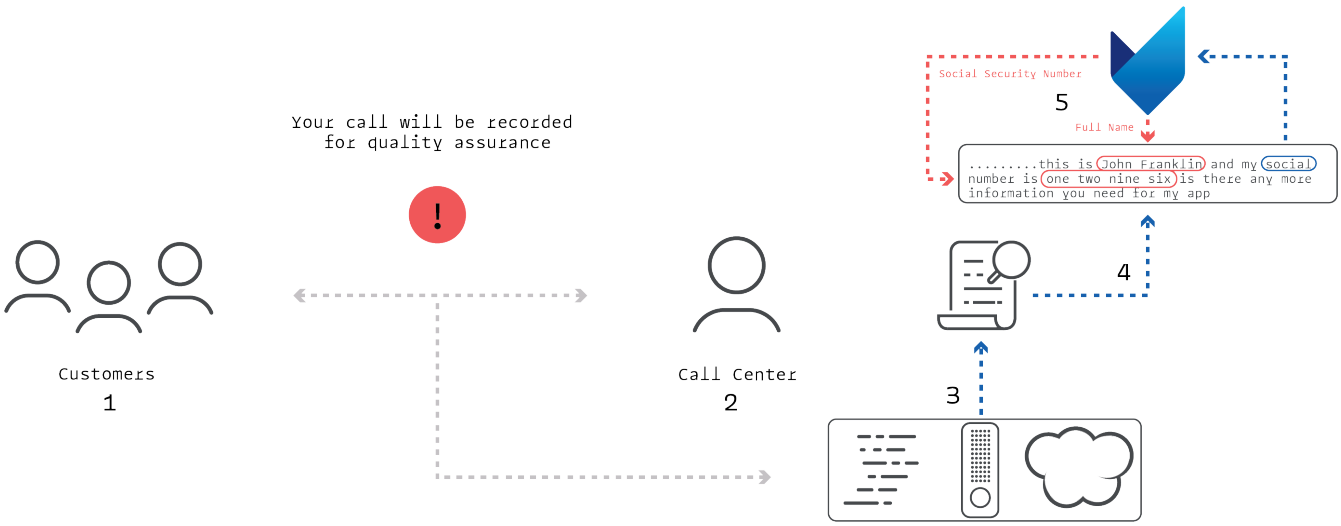
Solution

PK Protect, developed by PKWARE, offers a simple, comprehensive solution for sensitive data governance that detects and protects data across on-premises and cloud repositories.

PKWARE

PK Protect is a simple and comprehensive solution for sensitive data governance. PKWARE benefits include:

- Use a single solution to discover sensitive personal data across all data repositories, including Amazon S3
- Improve data privacy by masking, encrypting, and pseudonymizing data across the organization
- Audit and monitor sensitive data with automation—all from a single dashboard view
- Address data privacy and compliance regulations, including GDPR, PIPEDA, CCPA, and PCI-DSS
- Scale the process of discovery, management, and protection even as data grows



The first step in protecting data is identifying and locating sensitive data across the organization

PKWARE

Why PKWARE

The financial services company had existing systems and processes built on Amazon Web Services (AWS), including apps that pulled personal data into data streams, pushed the data into an Amazon S3 data lake, and performed queries on Amazon Redshift.

The company had simple but critical requirements for a solution. It needed to be compatible with existing AWS infrastructure, scalable, and cost-effective. In the end, the company needed to harness all of that data and ensure that the data was stripped clean of personal information before it was even extracted from Amazon S3 or queried in Amazon Redshift.

That's where PKWARE came in. "What helped make this decision clear was reviewing the initial test results from PKWARE and confirming there were sensitive elements within our Amazon S3 environment that were outside their originally perceived locations," said a senior big data architect at the financial services company. "PKWARE's broad platform support covered all of our disparate hybrid environments at an overall cost lower than their competition."

When examining its data privacy needs, the financial services company needed to instill proper controls to govern its data to meet in-house information security (InfoSec) requirements and regional regulations. "In an attempt to protect sensitive information on Amazon S3, we evaluated over a half-dozen technology vendors. In the end, it was clear that PKWARE has the only solution on the market that can find sensitive data in the specific formats we store and at the speed in which we need to achieve our business goals."

PKWARE and AWS

Deployed across the organization, PKWARE enabled the financial services company to audit and clean all of its at-rest data.

The solution inventories the location of sensitive personal information contained across all Amazon S3, Amazon Redshift, Amazon EMR, and Amazon Aurora databases, as well as third-party repositories running on AWS on a virtual machine, including Oracle, SQL Server, and many more. Once the sensitive data is identified, PKWARE protects the data elements via encryption, masking, and pseudonymization, then monitors who accesses this information over time.

The PKWARE solution also helped the company locate PII data embedded in its audio files. "If people give sensitive information over the phone that is converted from voice to text, we should be able to make sure that those things are encrypted and removed off the data lake," says the big data architect at the company. PKWARE found that PII and masked it—a critical piece of the client company's compliance.

Not only did the PKWARE solution meet the company's parameters for compatibility, scalability, and cost, but it was easy and fast to launch. PKWARE was up and running within a few hours after the initial proof of concept.

Results and Benefits

Today, the financial services company uses PKWARE on AWS to discover, protect, and monitor sensitive data across the organization and ensure all processes for compliance are in place for its growing data. If something breaks, or if the company should miss anything, the PKWARE tool should quickly identify the issue—and help the company close the gap. Yet, as the company's big data architect says, that rarely happens. When running the processes and tools, most of the time the tool yields "no results"—meaning no PII data has been found in the company's systems.

"PKWARE was able to locate specific sensitive elements in locations no other solution was able to find," says the big data architect. "The alternative of manually sifting through the data warehouse or implementing a solution that required extensive, time-consuming professional services was simply not feasible."

So far, the company has met GDPR and PIPEDA requirements. With PKWARE in place, the company now has the ability to shift its focus away from the daunting task of locating sensitive data and now focus on other key aspects of its business—and look ahead to future regulations. PKWARE and the company both know that GDPR is just the start, with the regulation likely to trigger privacy laws across the globe. It's already beginning, with the new California Consumer Privacy Act (CCPA) and the Australian Government Agencies Privacy Code both establishing stringent regional requirements for data usage and consumer privacy.

"PKWARE was able to locate specific sensitive elements in locations no other solution was able to find. The alternative of manually sifting through the data warehouse or implementing a solution that required extensive, time-consuming professional services was simply not feasible."