OneDPO Case Study

Find privacy risks and stale data

Industry: Automotive

A large automotive company in the US

1 Million+
DATA FILES & TABLES

Billions
PERSONAL DATA ELEMENTS

The Problem - Identify privacy risks and stale data

An automotive company's European division had recently migrated a large volume of data from on-premises into Microsoft Azure blob. The team was concerned that the blob storage might have PII information that could lead to any potential privacy violations. Any privacy violations could result in penalties and loss of consumer trust. Their analytics team believed their data storage contained mainly backup data, but their privacy managers wanted to understand what personal data were in the cloud storage. Additionally, they wanted to make sure if employees are actively using the data. If the data is not actively used, they wanted to archive it.

The Challenge

The storage contained close to a million files in a variety of formats. Any manual data audit effort might take several weeks, consuming time and resources. Moreover, with a large volume (14 TBs) of data, locating personal data was complex and required advanced AI/ML techniques to identify personal data. Even if they had created an inventory of data, they might have to spend many more days analyzing user activities and finding the stale data assets.

OneDPO Found PII and Stale Data

The Solution

OneDPO offered a turnkey solution that came with pre-built integration to Microsoft Azure blob storage. Within a few hours of the deployment, OneDPO started analyzing TBs of data to produce the results.

The data objects were large tables, CSV, XML files that contain thousands of lines of data. Other standard methods or tools would have taken months for the team to analyze. OneDPO did in a fraction of that time. OneDPO's distributed processing and smart sampling helped process the 1M data objects in a few days.

Apart from standard personal data, the customer's privacy team has classified a few personal data types as personally identifiable data. OneDPO team customized the ML models to scan for these additional data types that the customer considered personal data and personally identifiable data.

Our solution showed that 99% of their data in this instance of the storage was never used. OneDPO's workflows helped the company to get approvals on data archival and collect privacy metadata. OneDPO maintained an audit trail for any future reviews.

The Results

OneDPO saved several months of work that the privacy team would have spent in analyzing and discovering issues. Our platform found that 99% of the data was never used. The customer reduced their data protection surface area and risks tenfold by archiving these stale files.

100%
PERSONAL DATA

Discovered and mapped all the personal data across 1 million data objects.

10X
MINIMIZED DATA & RISKS

99% of the data was stale that no one was using. The company could delete this data and reduce its privacy risks by tenfold.

"We implemented OneDPO to get visibility into privacy risks in our cloud data. Finding unused data was a big bonus."
- Privacy Manager

www.onedpo.com