satori

ACME Enables Analytics for PII Datasets with PII Aware Masking

Accelerating analytics while staying compliant.

Goals

ACME helps growing businesses meaningfully connect and engage with their customers. With their SaaS platform, they help businesses go beyond marketing automation to optimize their customers' experiences. Part of ACME service includes storage and analytics of consumer personal information and it is a top priority goal for ACME to make sure that data is safe and secure.

ACME's goal was to allow analysts to work with data to create new analytics based services and insights while maintaining a high level of security around sensitive data.

Challenges

ACME's service relies on advanced analytics to provide the service and to monitor business performance. That means that multiple teams need access to a central Snowflake data warehouse and an S3 based data lake.

The traditional way of providing analytics on top of PII is to create a masked copy of the data for analytics that is separated from the raw data.

The challenges in this approach was twofold. First, the PII in ACME data stores is collected through forms their customers publish on their websites. Neither ACME nor their customers have control over what data is shared in these form fields and any of them could contain PII. Second, ACME has set up an S3 data lake for the data science team to build data models which require direct access to the data. Thus, access to PII is required to achieve business goals.

These two aspects of ACME's environment made it impossible to create a masked schema as identifying all PII up-front was not feasible and working with PII at some capacity was required.

Solution

Satori provides a secure data access service to monitor, classify and control access to PII across cloud data stores. With Satori, controls are decoupled from the data and do not require changes in data stores and users have the full context of who is accessing, what data and how it is being accessed to enable visibility and control.

To solve the PII challenge, ACME implemented Satori in their data lake and data warehouse. Satori is used for two main purposes: (1) Identify PII in real time as it is being retrieved and (2) Apply data masking based on ACMEs' policy for PII access. Satori also audits all PII access and its usage and provides a platform for meeting regulatory requirements and enables internal reviews of data access and security.

Outcome

With Satori, ACME can comply with existing and future privacy laws and regulations in an agile way, with less engineering overhead. ACME can stay agile in how they operate their data pipelines and analytics while maintaining a high level of security and compliance.