De-Identification Engine

Speaker Name – *Speaker Title*

Year Month Day

CUBIC | Transportation Systems

De-Identification Engine

Public & private entities have access to a lot of data that can drive insights and efficiencies

But they also have a responsibility to protect privacy

That's why Cubic developed the De-Identification Engine…

- Transforms sensitive and PII data into an analytics-ready dataset
- Makes data safe for internal and external audiences, including open APIs
- Applies de-personalization techniques based on the needs of use cases and data sets
- Includes removal, hashing, generalization, and aggregation
- Simplifies data automation process with embedded privacy controls

CUBIC | Transportation Systems

De-Identification Engine Evolution

The De-identification engine was originally developed to address privacy concerns for internal and external data usage at the **Massachusetts Bay Transportation Authority (MBTA)**…

But de-identification treatment is applicable to all **transportation agencies and authorities** for data analytics, while protecting the privacy of users…
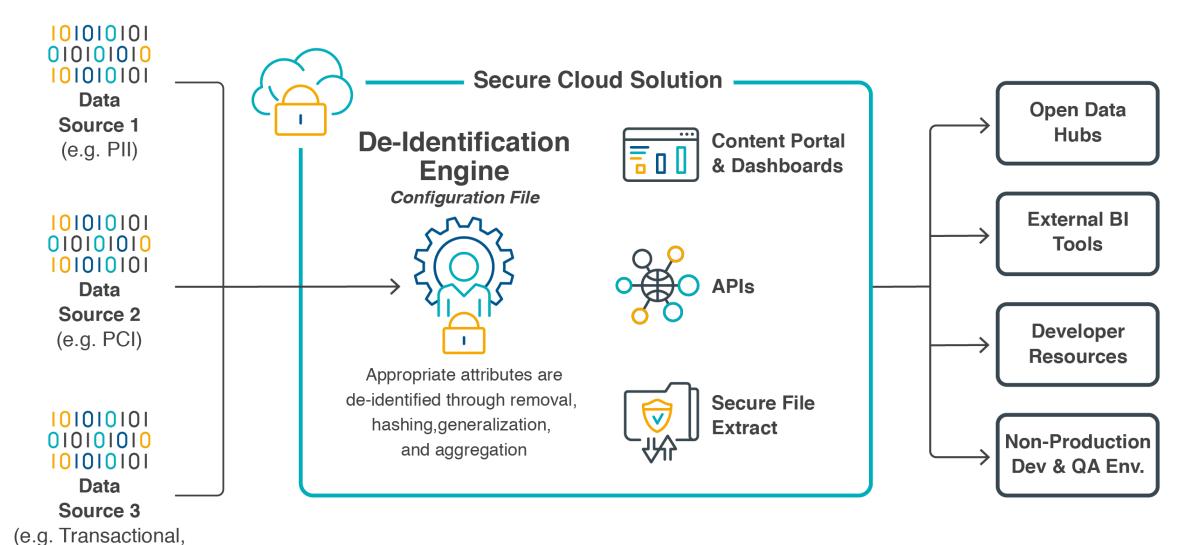
With growing data assets, privacy concerns aren't going away, and de-identification can support a multitude of use cases for **cities** and **private companies** from open government to economic development…

And these de-personalization techniques can be applied **across industries** to make sensitive data safe for internal and external consumption

De-Identification Engine Conceptual Diagram

Data Source 1 (e.g. PII)

Data Source 2 (e.g. PCI)

Data Source 3 (e.g. Transactional, Geo-Locations)

Secure Cloud Solution

De-Identification Engine
Configuration File

Appropriate attributes are de-identified through removal, hashing, generalization, and aggregation

Content Portal & Dashboards

APIs

Secure File Extract

Open Data Hubs

External BI Tools

Developer Resources

Non-Production Dev & QA Env.

CUBIC | Transportation Systems