

Secure Analytics for Auto Manufacturers Industry Use Case

Auto manufacturers capture massive volumes of data that, with predictive analytics, can produce valuable insights to monetize, improve products, optimize operations, and better serve customers with new features. Automotive IoT or connected cars, is expected to generate significant value to users, companies, and economies.

Secure Analytics and the Auto Industry at a Glance

With massive volumes of data pouring in around the clock from millions of cars on the road, auto manufacturers are implementing data lakes to capture real-time sensor data and other traffic, combined with historic data. But keeping sensitive data secure and private for use in analytics is a major challenge.

There has been a 6X increase in automotive cyberattacks between 2010 and 2018, according to the Upstream Security Automotive Cybersecurity Report 2019.* The rapid growth of cyberattacks in the connected car industry is impacting everyone in the sector, from Tier 1 companies to OEMs, fleets, car rental, insurance companies, and more.

* Source: Upstream Security Global Automotive Cybersecurity Report 2019

Security is the top barrier to IoT adoption across industries, but the connected car market is soaring, promising a continued explosion in data volumes—and risk. The undeniable benefits of IoT are dangerously offset by the heightened risk of data breach and potential non-compliance with data privacy regulations.

Locked-Down Data Protection vs. Open Usability of Data

There's potential conflict between creating new value with open access to data for analytics vs. securing data in a locked-down unusable mode. IT architects and decision-makers must be able to provide access to analytics platforms and data lakes while also safeguarding against cyberattack as well as non-compliance with data privacy regulations such as the GDPR and CCPA. Organizations that invest in IoT, but then lock down access due to security and privacy concerns, can't realize the needed returns on their technology investments.

Auto manufacturers are streaming, feeding, and storing sensitive data including geo-location codes, Vehicle Identification Numbers, and customer personal data that must be protected. For example, multiple geo-location codes for one driver, when combined in an analytics platform, may be used to identify an individual. If data is compromised, this can lead to regulatory penalties and loss of customer trust. There is much at stake if no clear data protection strategy is in place. There are traditional IT security controls that should be put in place, such as perimeter protection, and monitoring user and network

activity. But system-centric controls are unreliable, given the need for data access and use in analytics.

Storage-level data protection may check a box in data privacy compliance, but offers no capability in the analytics environment. Data masking is a one-way transformation and cannot enable replication of results.

The drive for insights and a return on investment cannot be achieved by locking down access to just a few data scientists, in silos, with static controls in place.

Maintaining Privacy with Usable Data

Data privacy regulations such as the GDPR and CCPA, and others, recommend encryption and pseudonymization as mechanisms that can be used to protect sensitive personal data and help enable compliance. Pseudonymization is a term for various data de-identification techniques where the pseudonym or surrogate data can be used in business processes. Field-level encryption and tokenization are such methods.

How Do We Scale Security at a Speed to Support the Growth of the Connected Car Business?

Using data at scale while lowering risk requires protection that scales with the data being ingested and analyzed. This calls for encrypting data as close as possible to its source before ingestion into analytics platforms and data lakes to eliminate gaps in protection, encrypting the sensitive data elements with usable, yet de-identified

surrogate values that maintain format, behavior, and meaning.

The Solution: Embed Protection into the Data Using Format-Preserving Encryption (FPE)

Augmenting infrastructure system and perimeter controls with protection embedded into the data is essential to scale protection, while mitigating risk exposure, in order to securely enable analytics.

Voltage SecureData Enterprise by OpenText™ with ArcSight Secure Data Format Preserving Encryption Add-on for ArcSight Data Platform (FPE) makes secure analytics possible in data lakes and analytics platforms ingesting massive amounts of data. Voltage FPE encrypts sensitive structured data at the field and sub-field level, preserving characteristics of the original data including numbers, symbols, letters, numeric relationships, and referential integrity across distributed data sets.

The protected form of the data can be used in applications, analytic engines, data transfers, and data stores while being readily and securely re-identified for those specific applications and users that truly require access. Yet, in the event of a data breach, the protected data yields nothing of value, avoiding the penalties and costs that would otherwise have been triggered.

Primary Use Case: Predictive Analytics, Vehicle Maintenance

Ecosystem Scope: Big data analytics, mission-critical IT, cloud, IoT data

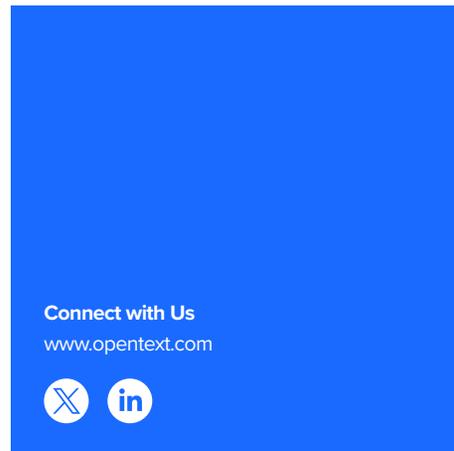
Predictive analytics enable manufacturers across many use cases, from demand sensing to anticipating supply chain issues, identifying emerging safety and quality issues. Business results include improving customer satisfaction, vehicle safety, and brand perceptions. They can also realize reductions in total cost of spending on quality.

With the rise in threats from cyber-attacks, and data privacy regulations, a key goal is to ensure that all sensitive data is protected in

analytics. Automotive manufacturers need a proven, unified architecture for secure use of sensitive information by data scientists, DBAs, and third parties across relational and non-relational databases and storage.

Today, that vision is achieved by auto manufacturers with petabytes of data secured across unlimited data types, hundreds of applications, and the data lake with Voltage SecureData Enterprise. The OpenText Simple API easily integrates into a wide range of Big Data and ETL products. Off-the-shelf code provides templates for Hive, Spark, cKafka/NiFi, Kafka/Storm, Map Reduce, Impala and Sqoop. Edge nodes may be used as a landing zone for sensitive data then stored in HDFS.

One example use case is to protect real-time sensor data for performance analytics. Data is



protected at the IoT edge, streaming in real-time from vehicles on the road, encrypted prior to ingestion in the analytics platforms, followed by orchestration of protected data into HDFS and other environments.

Use Case: auto manufacturer data flow

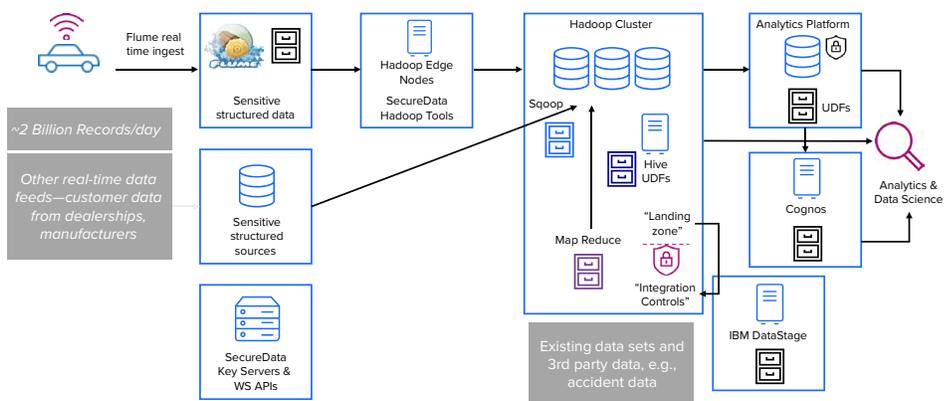


Figure 1. Auto manufacturer data flow

Safely Unleash the Power of Big Data for Customer Analytics

Voltage SecureData with FPE and SST enables automotive manufacturers to:

- Extract value safely with analytics from protected data in Hadoop data lakes, relational and non-relational database management systems
- Comply with GDPR, CCPA, and other data protection regulations, such as PCI DSS
- Deploy data protection from ingestion at the source and throughout its end-to-end lifecycle

Find out how to unleash big data for analytics insights to more users and applications for increased business value creation.

Learn more at www.microfocus.com/sdhadoop