## Volume

250M Verbundene Autos

150B Umsatz

60-200 Sensoren aus verschieden Bereichen wie

ECUs

Fahrzeugs Plugins

Kameras

## Variety

Verschiedene Sensor Daten:

SQL(Impala)

SDK(Partners)

Streaming(Spark)

Batch (Spark, Hive)

Multiple Formats, wie Data von der ECU

GPS Data

Saftey Data

V2V und V2I

Video

## Value

Aus den Sensordaten Information gewinnen mit hilfe von

Spark und Spark Streaming KDU, cloudera

## Velocity

Realtime data mit Kafka

25 Gb pro stunde pro Auto

130 TB pro jahr pro Auto

## Veracity

Sensordaten können Messfehler enthalten

## 4 Levels of Data Handling

### Data Source Layer

Die Daten warden von einem Conntected Car generiert über verschiedene sensoren/IoT Data.

Dann auch noch weitere Interne und externe Daten Sources

### Data Storage Layer

Die Daten warden auf einem Filesystem HDFS, Relatione Datenbank Kudu und NoSQL DB Hbase gespeichert.

### Processing Layer

Verarbeitet warden die Daten mit Spark, Impala, Solr und Hive

### Data Output Layer

Die daten warden Visualisiert mit Hilfe on Arcadia