BIG DATA IN DER AUTOMOBILINDUSTRIE
Daten aus dem Fahrzeug nutzen

Norbert Widmann I 11.07.2018

BMW GROUP
THE NEXT 100 YEARS

Rolls-Royce
Motor Cars Limited
AGENDA.

1. BMW, BMW Group IT, Big Data @ BMW.
2. Big Data Platform & Data Lake.
3. Use Case: Driving Patterns.
4. Use Case: On Demand Mobility.
THE BMW GROUP IN NUMBERS (2017).

- 129,932 employees worldwide
- 2,463,526 sold automobiles worldwide
- 98.7 Billions EUROS revenue worldwide
- 30 production locations worldwide
The BMW Group IT consists of more than 4500 employees from over 60 nationalities working in 29 countries all around the globe. We manage 450 software projects for all BMW business areas and speak 47 languages.
“BIG DATA, MACHINE LEARNING, ARTIFICIAL INTELLIGENCE” AT A GLANCE.

~100 Employees 20+ Divisions, for which use cases are currently being built

150+ conducted Use cases

Prospects, Use Cases, Projects

150+ conducted Use cases

Data Lake

Plattform, Architecture, Tools

Data Scientist ~35%
Big Data Architect ~25%

Roles

Data Engineer ~20%
Big Data Project Manager ~20%

Data Scientist ~35%
BUSINESS PROCESSES USING ADVANCED ANALYTICS AND MACHINE LEARNING

<table>
<thead>
<tr>
<th>Use Case</th>
<th>Digital Services</th>
<th>Aftersales</th>
<th>Production</th>
<th>Vehicle Development</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Location-based Services</td>
<td>Predictive Maintenance</td>
<td>Realtime Failure Detection in Assembly</td>
<td>TraceNet</td>
</tr>
<tr>
<td></td>
<td>Geo/Location-Data, Fleet-Data</td>
<td>ECU /Vehicle Data</td>
<td></td>
<td>Vehicle Trace Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Images, Order Data</td>
</tr>
</tbody>
</table>

Images, Order Data
ECU /Vehicle Data
Geo/Location-Data, Fleet-Data
Vehicle Trace Data
Realtime Failure Detection in Assembly
Predictive Maintenance
AGENDA.

1. BMW, BMW Group IT, Big Data @ BMW.
2. Big Data Platform & Data Lake.
3. Use Case: Driving Patterns.
4. Use Case: On Demand Mobility.
BIG DATA PLATTFORM. HAUPTBESTANDTEILE.

- Analytics Workplace (FG-24)
- BMW Clients / BI Frontends
- Real-Time Log Analysis
- Large-Scale Distributed Storage and Processing
- Ingest Platform
- Scalable Event Hub (Kafka)
- External Systems (databases, machines, cars, etc.)
## PLATFORM SIZING.

<table>
<thead>
<tr>
<th>Location</th>
<th>Servers</th>
<th>CPU Cores</th>
<th>RAM</th>
<th>Disk Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMEA</td>
<td>256</td>
<td>10.240</td>
<td>49.152 GB</td>
<td>5.1 PB</td>
</tr>
<tr>
<td>USA</td>
<td>24</td>
<td>1.016</td>
<td>4.864 GB</td>
<td>1.1 PB</td>
</tr>
<tr>
<td>China</td>
<td>16</td>
<td>680</td>
<td>3.328 GB</td>
<td>0.7 PB</td>
</tr>
</tbody>
</table>
Conclusion:
- Automated provisioning and configuration of data science environments help data scientists to focus on their job.
- Isolated environments reduce risk of unwanted impacts among projects.
- Portability through containerization
- Integration with our Big Data Platform (e.g. Spark, Hive, Kafka) + GPU Support (DGX-1, Cluster GPUs)
DATA LAKE: BREAKING UP DATA SILOS ACROSS THE COMPANY.
FACTS AND FIGURES.

3 PB of Data

300 TB of Data Assets

Growth per Day

1 TB Data Assets

40 Areas as Users

3 Platforms on 3 Continents

1 Big Data Catalog

55 Data Assets live

26+ Data Assets on the Roadmap

32 Data Stewards

161 Data Scientists and Business Analysts as Users

From 45 Departments

Data from all Business Areas

Vehicle Development 40%

Connected Vehicle 27%

Aftersales 18%

Production & Logistics 13%

Other
AGENDA.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>BMW, BMW Group IT, Big Data @ BMW.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>Big Data Platform &amp; Data Lake.</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Use Case: Driving Patterns.</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Use Case: On Demand Mobility.</td>
</tr>
</tbody>
</table>
FROM IDEA TO IMPLEMENTATION. „AI DELIVERY MODEL“.

Data discovery

- Top-down
- Bottom-up

? + !

Data exploration

- "Fail fast"
- Further analysis, proof of large-scale implementation feasibility

Decision gate 1

Decision gate 2

Agile

~20% throughput

Implementation

Business Value

Potential and feasibility analysis within 10 days

Potential and feasibility analysis within 10 days

Top-down

Bottom-up
EXAMPLE: THE VALUE OF OUR BMW FLEET DATA.

COLLECT DATA

ANALYZE

ACT

- Derive requirements for HAD/FAD
- Evaluate design specifications
- Marketing hotspots in cities
- Launch area for parking app
- Robot Taxi Strategies
- Reduce Product Complexity (based on usage)
- ...
MAP MATCHING – MOTIVATION.

Map

- Complex data model
- 100 Mio. street segments (Europa)

Map-Matching

- Statistical Problem
- Depends on map and geo data

Scaling

- Distributed computing
- Scales horizontally

Figure 1: Map matching consists of matching measured locations (black dots) to the road network in order to infer the vehicle’s actual path (light gray curve). Merely matching to the nearest road is prone to mistakes.
LAUNCHING PARKING APP IN CITIES / AIRPORTS.

– Which cities / airports should be prioritized when launching a new parking app?
– How many BMW’s enter certain cities / airports per day?

1. DEFINING GEOFENCES
2. COUNTING ACTIVITY
3. RANKING AREAS
AGENDA.

1. BMW, BMW Group IT, Big Data @ BMW.

2. Big Data Platform & Data Lake.

3. Use Case: Driving Patterns.

4. Use Case: On Demand Mobility.
On-demand mobility

Hey BMW, what does the yellow light in my cockpit mean?

Hey Alexa, open my garage door.
Demand Prediction

Where, when and how many ride requests do we have at any given moment in time?

Fleet Intelligence

Which driver picks up the passenger and which route does the driver take?

Success factors
Price, QoS, ETA, …
PROBLEMS

Undersupply → Waiting time

Oversupply → Utilization

Mid- and short-term demand prediction

1. Match supply with demand distributions.
2. Incentivize drivers moving from oversupply regions to overdemand regions.
Thank you.