Data-driven machine learning and optimization pipelines for real-world applications

Koch, M.

CHAPTER 7

Data-Driven Services in the Car Industry

Numerous recent studies show the prosperous future of data-driven business models. Some key challenges have to be dealt with when moving towards the development of data-driven car services (see Chapter 2). We present a more general approach towards the development of data-driven car services. We point out its main challenges and suggest a method for developing new customer-oriented data-driven services. This approach illustrates key points in developing a practical service, from a technical and business related perspective, which is connected to individualized service examples that would potentially benefit from it. Such data-driven services are developed mostly on a small number of initial test data, which results often in a limited prediction performance. Therefore, based on an optimized Cross Industry Standard Process for Data Mining (CRISP-DM) approach, we propose a methodology for developing initial prediction models with limited test data and stabilizing the models with newly gained data after deployment by online learning. On-board and off-board services are discussed with the result that especially off-board running services offer a large potential for future data-driven business models in a digital ecosystem. The flexibility of such an ecosystem depends on the degree of the integration of the vehicle in the ecosystem - in other words, the car needs to be enabled to deliver data on demand according to General Data Protection Regulation (GDPR) and to any applicable regional law and in cooperation with the customer. The presented method, together with the ecosystem, enables fast developments of various data-driven services.

7.1 Introduction

Robotics and transportation have been underpinned by artificial intelligence since its early beginning. In 1969, Nilsson (1969) discussed the use of artificial intelligence in integrated robot systems. In the late 1970s pioneering discussions were made on the first autonomous vehicles with artificial intelligence (Tsugawa et al., 1979). Across the end of the 1970s to the 1990s, first prototypes were developed by different scientists and organizations (Schmidhuber, 2018). Such technical progresses continued until 2000 and the autonomous driving was feasible for the first time, sparking major developments in both research and industry (Stone et al., 2016; Huber et al., 2008; Aeberhard et al., 2019; Ardelt et al., 2012). In some reports, it is estimated that as of 2020, 10 million cars featuring self-driving will be on the street (Greenough, 2018). In autonomous driving, data from different sensors are combined by computers deployed in the car (Liu et al., 2017). Using methods of artificial intelligence (specifically deep learning techniques), these computers predict the car actions that are required to handle situations. Due to the large data volume, those artificial intelligence models are mainly deployed on on-board-systems (embedded) in the car (Aeberhard et al., 2019). Beyond autonomous driving functionalities, certain types of car data, especially the one related to self-driving, are combined with a car internet interface and a robust internet connection, offering a new era of data-driven services. Most of these services require no additional car hardware and operate only with the vehicle data that is available. As such services are mainly driven by small data volumes, the data set used can be transferred to a back-end system, complying with data protection regulations and customer's consent. This enables running the data-driven service outside of the car (off-board). Hence, new services do not require any changes in the hardware, which significantly simplifies the service development. It allows for the continuous and faster creation of new services, even within the lifetime of cars. Therefore, off-board running services are much more powerful than in-car computations: A car interface sends data on request to a back-end system, which uses data-driven models for a prediction and sends the output, e. g., back to the car. This is combined with full transparency and involvement of the customer regarding certain data. The interconnection of vehicle and back-end system builds a so-called *digital ecosystem*, which integrates all aforementioned methods to provide

car services. It enables faster service development and deployment, even within the lifetime of cars.

The revenue from mobility services and connected car services are projected to reach USD 1,087 billion by 2030 (Seiberth and Gruendinger, 2018). This is not only a large business for OEMs (Original Equipment Manufacturer), but also for suppliers as well as ecosystem developers and other parties involved (Seiberth and Gruendinger, 2018). Data availability, its protection and privacy of an open (e. g. to third parties) digital ecosystem, is of key importance to integrate cars more seamlessly into our lives with more digital services. This study mainly illustrates the technical approach with its challenges for developing new data-driven customer services, going from the idea to a running service.

The remainder of this chapter is organized as follows: First, existing work that is related to our approach is discussed in section 7.2. Second, we present an example of a data-driven service in section 7.3. Our proposed methodology with its six main steps is then introduced in section 7.4. In section 7.5, we provide conclusions and an outlook.

7.2 Related Work

Recent studies illustrate new data-driven business models in the car industry by means of a digital vehicle ecosystem. In this context, Seiberth et al. present a definition of data-driven business models: "data [...] as primary business resource to deliver value to customers and to convert this value into revenue and/or profit" (Seiberth and Gruendinger, 2018, p. 8). They declare that in 2050 car manufacturers will achieve 50 % of the revenue from data-driven services. The growing digitalization with its disruption process destroys many traditional business models (Weill and Woerner, 2014). The authors also picture more general business models and the possibilities of digital ecosystems for different industries. Car manufacturers have different approaches to deal with digital services and many tech start-ups are already developing sustainable business models with digital services. Furthermore, OEMs enter already strategic partnerships and invest into such connected vehicle start-ups (Kaiser et al., 2017).

Seiberth and Gruendinger (2018) discuss the available car data, e. g. from sensors for autonomous driving, and highlight the possible revenues when creating services

based on it. In addition, there is a growing need for building trust towards the customers regarding the use of their data for services and therefore for the transparency of the data, its use, and privacy and security (Kilian et al., 2017). Beyond that, they present a figure of *the connectivity ecosystem*, which describes roughly a connectivity platform: it communicates with the data source (car) and receives external data like weather, traffic etc. The connectivity platform is connected to the OEM's back-end system, as well as to third party services and apps.

Many studies present new business models enabled by data (Seiberth and Gruendinger, 2018; Weill and Woerner, 2014; Kilian et al., 2017). In most cases, new service ideas are superficially mentioned and it is only briefly discussed how to really benefit from each individual service. Some of the studies discuss the design of (vehicle) ecosystems, e. g. Immonen et al. (2016, 2018), but a methodology for creating data-driven (customer) vehicle services has not been a major topic of scientific research yet.

7.3 A Data-driven Service for Crash Damage Prediction

The variety of possible data-driven services is large. A data-driven car service often assists the customer (like a car pooling service) or the car (like predictive maintenance services). Based on historical data and with methods of artificial intelligence, models are trained to predict behavior, e. g. in car pooling to predict the best possible route to carry the most passengers or if a certain part of the car needs to be maintained in the nearby future.

Another example of a data-driven service is a crash damage prediction system (see Chapter 7). Based on a machine learning model, such a system predicts the damaged parts of a vehicle in a low speed crash. A low speed crash is an accident with a velocity difference below approximately 16 km/h (RCAR, 2018). The baseline of this service is to use only on-board data. Therefore, data from serial car sensors are used for the prediction (e. g. acceleration). To generate an initial data set, low speed crash tests are performed and certain on-board data are recorded. These recordings are used together with the occurred damage on the vehicle for training first initial models. The benefit of such a data-driven service is

e. g. immediate transparency of the damage, which allows initiating a faster and more convenient repair for the customer (Seiberth and Gruendinger, 2018; Koch et al., 2018; Koch and Bäck, 2018).

However, the machine learning model itself is only one unit of the car service. When striving to create seamless customer experiences with a data-driven service, it is essential to consider the whole customer journey. Such a journey describes the way how a customer experiences the whole service. The overall objective should be creating something which is so-called convenient to the customer at all levels. This can be achieved by designing the end-to-end service with its technical challenges like data transfer or intuitiveness of its handling as a whole picture. Based on this, in the following we propose a general path towards data-driven services to tackle and consider the challenges with the one and only goal to create customer value.

7.4 Methodology towards Data-Driven Services

In this section we propose a methodology for developing data-driven car services. This interdisciplinary method is illustrated in Figure 7.1. The horizontal axis shows the time of the development while the vertical one describes the level of development, i. e. the maturity of the service. The origin presents the time of the initial idea about the service and the beginning of its development. The methodological approach consists of six overlapping phases:

1. Idea.
2. Potentials.
3. Modeling.
4. Deployment.
5. Process.
6. Finalization.

All phases are linked to each other. In order to allow short development times the phases are partially executed in parallel. The phases are described in the following sections.

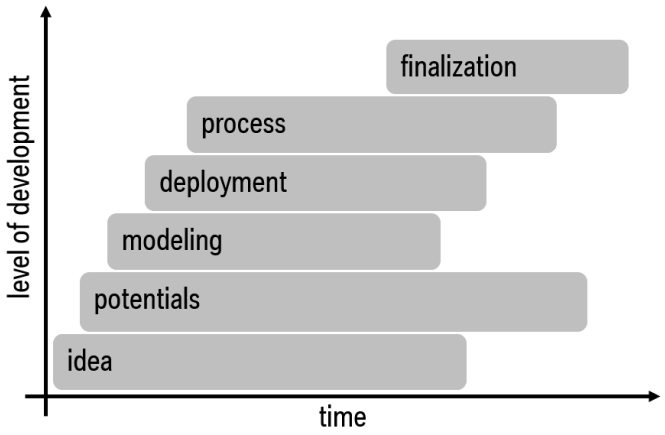

Figure 7.1: The methodological approach: From the idea to a deployed data-driven service.

7.4.1 Idea

The first phase of Figure 7.1 is referred as idea. This pictures the timeline from the first idea about the service to very concrete solution concepts. Principally, there are many motivations or ideas for thinkable services, but for successful and seamless services the business potential and customer benefits have to be evaluated continuously in the next phase, the evaluation of the potential.

7.4.2 Potentials

Seiberth et al. state that new car services follow mainly two objectives: improvement of the brand image or increase of profit (Seiberth and Gruendinger, 2018). This shows that the motivation to create those is based on image or profit reasons or a combination of both. Therefore, data-driven services can have strong impacts on the brand and can be used for strengthening images with creating so-called customer experience by building positive experiences followed by an emotional bond between user and product (Glattes, 2016).

Next to retail customers other stakeholders like, e. g., fleet operators, insurance companies or other parties can strongly support such services with their own advantages (Seiberth and Gruendinger, 2018). Creating a service with many

benefiting parties exploits its potentials and is key for a successful and seamless service. Therefore, in case of promising ideas, in phase 2 of Figure 7.1, it is important to analyze and continuously evaluate all aspects of the data-driven service regarding the own objectives and the targets of partners. However, it is mostly very ambitious to evaluate the real potential of a new service in an ad-hoc manner. Therefore, it is important to quickly develop prototypes for experiments, get customer feedback and constantly monitor the need for the service and decide continuously to proceed or cancel the development.

In this context, after revealing an initial potential of the idea, data scientists begin the phase of modeling with collecting data and designing first models.

7.4.3 Modeling

In the beginning of the modeling phase, data scientists have to prove the feasibility of describing the desired relations by the available data with methods from the field of artificial intelligence. A feasibility study helps to quickly assess the practicability of the idea.

To start the modeling phase, an initial data set is crucial. In some cases, the data has been already collected and is available or can be gathered quickly. However, in most cases the data has to be generated manually. When considering the damage prediction system, data from low speed crash tests are required. Performing large numbers of such tests is very tedious and expensive. Therefore, in such cases only small initial data sets are generated in order to evaluate the feasibility. Prediction models based on small data sets are often of poor prediction quality. In order to increase its quality and especially to have informative results for the feasibility study, the use of optimization techniques is key.

Shearer proposed an approach to run data mining projects in industry, the CRISP-DM (Shearer, 2000). This approach has become a very known standard process to perform industrial data science projects. Roughly, it describes the process from the business understanding to data understanding, data preparation, modeling, evaluation until its deployment. We have modified parts of the CRISP-DM and added optimization between modeling and evaluation in order to enhance the model performance. Furthermore, we have separated the data into initial data and field data, as well as the process streams into offline learning (black) and online learning (yellow) (see Figure 7.2). Offline learning describes the process

of learning models with an offline generated (initial-) data set (Service Data) to create an (initial-) prediction model. In addition, car information data like the car type, the equipment of the vehicle and geometry information, as well as external data like, e. g. weather or traffic are used as additional data resource, because such data often contain valuable information for the service. After deploying the initial model in an offline learning process, we are updating it by online learning (yellow stream). This means that the initial model is stabilized step by step after deployment with newly generated field data.

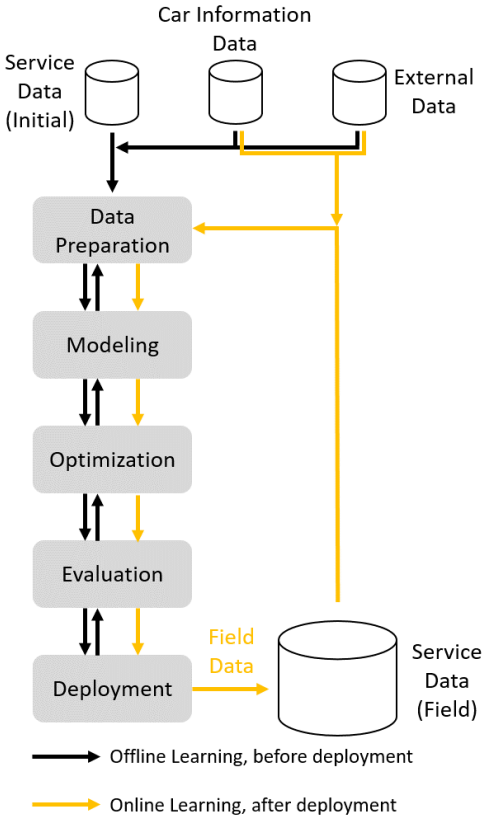

Figure 7.2: The modified CRISP-DM approach with optimization and online learning components. Note that the offline learning part follows the methodology proposed in (Shearer, 2000).

We have developed this modified CRISP-DM approach, when we were dealing with the modeling of the crash damage prediction system. Its required crash data

is extremely difficult to generate at large volume, because either crash tests or simulations have to be performed. Therefore, we only created a small test data set containing just enough observations to verify whether it is feasible to use on-board data to predict the damaged parts. We obtained a data set with 100 observations. The number of damages of some parts is less than 5 among the 100 tests. This indicates a very small and class-imbalanced data set. Due to the character of our data set, first results with, e. g., multi-label classification methods were not leading to promising results. More and more we have tailored our approach: we developed a part-wise classification, i. e., we generated individual prediction models for each part of the vehicle. This was very promising, because each model has its own set of characterizing features and its own set of hyperparameters (see Chapter 5.2). However, creating hand-crafted predicting models for each vehicle part is a very time consuming process. As a result, we developed our own automatic approach for time series classification, a so-called machine learning pipeline. The input of our pipeline are time series with the corresponding label. The outputs are predictive model performance measures such as accuracy or $F1$-score, which describe the quality of the prediction.

In Chapter 4.3 our pipeline approach is shown in detail. This pipeline describes the modeling and optimization part of our modified CRISP-DM approach more in detail. Our pipeline is computational relatively cheap and shows promising result (see Chapter 5 and 6). We developed this pipeline to efficiently generate individual models predicting the damage for each part and, more importantly, the pipeline can be used for automatically enhancing and stabilizing the initial model performance after deployment by online learning following the methodology of our modified CRISP-DM approach.

Our initial pipeline models have achieved $F1$-scores between 0% and 94%. This indicates, that based on the small number of data points the predictability depends strongly on the part, i. e., the damage of some parts can be predicted more precisely than of others. Additional methods like frequent pattern mining could help analyzing which parts are likely to be damaged together within one crash. This is especially important for parts with a low prediction quality. Furthermore, by considering, e. g., the learning curves from our results we were able to foresee an improvement of the performance with increasing data (see Chapter 5). Among other things, this helped us to evaluate the feasibility for practical use.

The solution space for such modeling problems is usually large and often the combination of different methods is leading to usable results in practice. In our opinion automatic machine learning methods (AutoML) like our pipeline approach are very promising, because of its practical use and due to its performance and its efficiency (short computation times).

Modern vehicles employ many different sensors and can produce large amounts of data. Sometimes it is not obvious what data would be promising for modeling. Therefore, from a possibly large number of sensors, the most promising ones for the task at hand have to be discovered.

As mentioned, after the data generation a feasibility study shows the practicability of the data-driven service. When receiving results matching the expectations, the models for the serial application can be developed.

In conclusion, the key of the modeling phase is to generate efficient predictive models and to check the validity of the service model approach by assessing the quality of the models that can be learned from the data. After the feasibility is identified and confirmed, the deployment of the service should be prepared.

7.4.4 Ecosystem/Deployment

The deployment of a data-driven service in an automotive environment depends mainly on whether it requires on-board or off-board running services. On-board services are deployed on embedded systems in the car. This needs data storage and computing power on control units of the vehicle. Off-board services are running in back-end systems. In this case, data is transferred via the internet interface from the car to the back-end, provided a sufficient bandwidth is available. In both off-board and on-board running services the car needs to provide the required data. In this regard, in the deployment phase the electronic components of the car have to be enabled to deliver the needed data. Figure 7.3 shows a typical vehicle network of a passenger car. Such architecture consists of different communication systems like Ethernet and more traditional *bus systems* like *FlexRay, CAN* or *LIN*. Ethernet is a local area network (LAN). It is designed to transmit data between computers (Spurgeon, 2000). BES are bridged end stations (switches), which can send and receive transmissions. Bridges communicate to other bridges, to the gateway (router) and to end stations (ES), which is in an automotive environment, e. g., the head unit (Spurgeon, 2000; Spurgeon and Zimmerman,

2013). Bus systems like *FlexRay, CAN* or *LIN* differ in various bandwidths and each system transfers data between components, called Electronic Control Units (ECU). ECUs are embedded systems, which control electrical systems in the vehicle. A car contains many ECUs like the engine control unit, the airbag control unit, the battery management system or the telematic control unit, which sends and receives data via the mobile network. All bus systems are connected via gateways (Robert Bosch GmbH, 2014; Matheus and Königseder, 2015). Sensors are connected to the ECUs, which process the sensor raw data and route it partly to the bus system. This bus data can be used from other ECUs within the connected bus or by the gateway. In some cases, data from one bus is required on another bus. Then, the gateway routes this data from one bus to another. However, mostly data is only available in the ECUs or on the initial bus system.

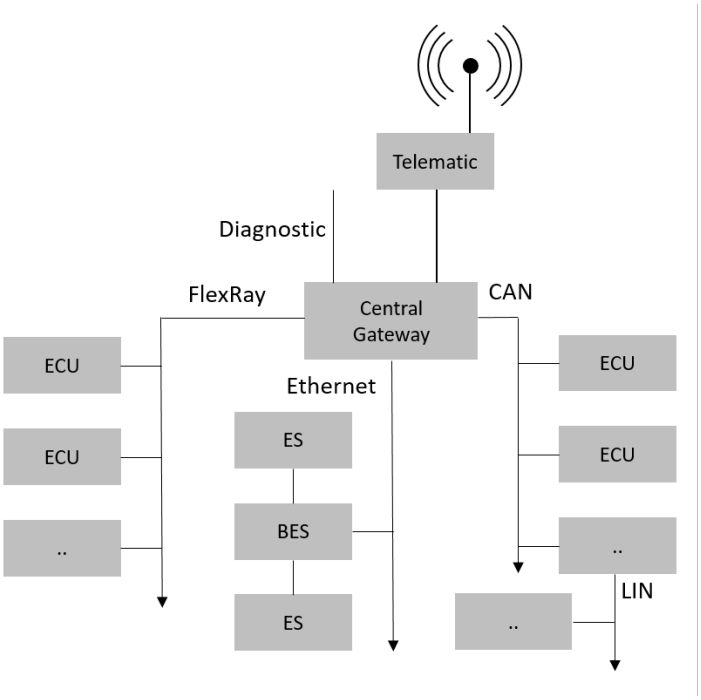

Figure 7.3: A typical schematic in-vehicle network.

Deploying a service on-board in an ECU demands a high effort regarding receiving / delivering the needed data, matching the data quality requirements and the general deployment of the software within the automotive system. This additional software needs to harmonize with all car systems and thus implies very costly

technical security. In addition, the capacity of storage within the ECU and its computing power is limited for additional services due to the fact that ECUs serve most likely other essential vehicle functions. Furthermore, it is very challenging to deploy on-board services in the lifetime of cars due to compatibility issues and the additional technical security required.

A more flexible way is established by transmitting the data from the car to a back-end system and running the data-driven system off-board. As soon as the required data is available on a bus system, this data is routed by the gateway to the bus where the sending unit (telematic) is located and it transfers the data from the car to a back-end system. The back-end computes the results and transmits it to the involved systems of the stakeholders. In this regard, the internet interface (telematic) of the car needs to be enabled to transmit different data package sizes in order to provide efficiently the required data. One crucial baseline is, that the bandwidth of the mobile network allows such data transfers.

As mentioned before, the deployment of a service on-board (embedded) is complex, time consuming and not as flexible as a data-driven service is meant to be. Some services need to run on embedded systems like autonomous driving (Liu et al., 2017). The functionality of most of the other services allows running outside of the car like in case of the damage prediction system. Off-board car services provide, next to their simpler deployment, more flexibility regarding faster model updates and adaptations. The key for off-board services is the availability of data: The car has to be able to send the required data on demand, according to GDPR and any applicable regional law and after the confirmation of the customer, i. e., the electronic architecture of the vehicle must be enabled to provide the requested data. This is the foundation of a digital ecosystem, which can collect demanded data and allows deploying new services and interactions with the car and other involved parties quickly.

Figure 7.4 shows the basic principle of a vehicle ecosystem for data-driven services. When developing a new service, the requested data is sent via mobile network from the car to a data layer, called service data, of a protected service cloud (Security Layer). This service cloud is a part of the whole vehicle back-end system. Next to the service data, the back-end system receives external data from third parties like traffic, weather and other service important information. Furthermore, the back-end contains car information like, e. g., the car type, the equipment of the vehicle, the drive technology, geometry information, service information. Such

information are often very valuable for a data-driven service. Hence, using these additional data can increase the prediction performances. The data-driven model (AI-Unit) is deployed to the back-end system, as well as other units (Flexible Units) like data processing. This back-end system communicates on demand with the car. Beyond that, the architecture of the ecosystem allows training the models with newly collected data from time to time or automatic (see Figure 7.2). Such a digital ecosystem gives even the possibility to provide partly accesses to third parties to create new valuable services, e. g. BMW Group (2018). Generated customer information from the service cloud are provided to customer devices in the car (control panel) or outside the car, e. g. mobile applications.

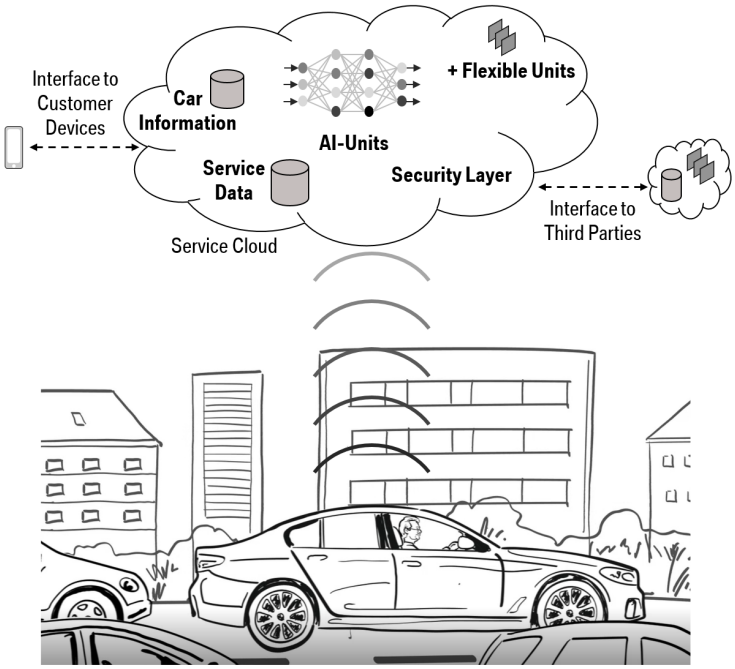

Figure 7.4: A digital ecosystem to run data-driven service. Note that this illustration follows partly the methodology of Kilian et al. (2017).

When considering an example like the damage prediction system, this would in concrete terms imply the following sequence of events: after a low speed crash event and a confirmation of the customer, a small data package is transmitted to the data-driven prediction model in the back-end. With the data package as input the model predicts, based on historical data, the damaged parts and the repair

costs. This information can be provided to participants like, e. g., the customer, the insurance companies for seamless claim settlements or the workshops for faster repair (see Chapter 5.2).

In all cases, before any data is transmitted, the customer has to confirm the certain service with an overview of the transmitted data. Furthermore, a general transparency of the service and its intuitiveness in understanding and handling must be provided within seconds to the customer. A confirmation can be canceled anytime. In this context, recent studies show that 94 % of connected car owners are interested in apps and services. Out of those 94 %, 84 % are willing to share personal automotive data for new services (Otonomo, 2018).

An ecosystem with seamlessly operating data-driven services requires data exchange from the ecosystem not only to the car but also to other stakeholders/participants. These processes need to be designed in regards to the business processes. This is described below.

7.4.5 Process

Running a service requires an interconnection of all stakeholders/participants. Without data transfer to all involved parties the potential of the service cannot be exploited. Therefore, shortly after having a rough idea about the deployment, the business process needs to be designed with taking all necessary stakeholders into account.

When looking at the example of the damage prediction system, the data of the damaged parts and the cost for repair are computed in the back-end system. It can be beneficial for the customer to send certain information to other participants like the workshop to order the parts immediately and to prepare the workshop visit. Beyond that, with detailed damage information the insurance company could approve the repair immediately, which simplifies the whole insurance claim settlement and would avoid an interaction of customer and insurance company. Such connections are identified and designed in the process phase. Furthermore, business architects establish customer oriented processes for running the data-driven system with all necessary parties connected. Often, the whole potential can be only reached when all parties are connected in a beneficial way.

7.4.6 Finalization

Figure 7.1 indicates that the finalization phase starts approximately half-way of the process phase. More precisely, when having first working systems, the finalization phase starts with testing, validating and improving the service. In most cases, it is indispensable to test the service with, e. g. defined customer groups to use this feedback for further improvements. In this phase it is key to consider and connect the five previous phases seamlessly with each other in order to create a customer experience.

7.5 Conclusions and Outlook

Nowadays, the expectations regarding data-driven business models in the car industry are massive. This chapter illustrates a track towards an efficient development and deployment approach for data-driven services in vehicles. It presents the important steps, as well as the main challenges. Through the explanation of the methodology, examples are drawn to show precisely the key points. A flexible and various service generation can be reached with a full integration of vehicles in a digital ecosystem, which means that the car delivers data according to GDPR and any applicable regional law and in cooperation with the customer to a back-end system. The main service runs on this back-end system, processes the data and transfers the results to the participants like the customer or other involved parties like, e. g., fleet operators. The shown method enables generating fast data-driven services in order to integrate cars more seamlessly into our lives.

As an outlook, we mention that data enables much more than creating data-driven services: data is transforming car manufacturers from traditional engineering companies to data-driven companies. This indicates that not only service creation but also car development in general is progressively based on data.