Deliverable 14

Ein optimiertes Produktivsystem auf Basis eines kontinuierlichen Performance-Monitorings und Auswertens der Benutzerakzeptanz ist vorhanden.

Ein neuartiges DRM-Konzept zum Handling personenbezogener Sensordaten im Rahmen der KI-Entwicklung für autonome Fahrfunktionen ist erarbeitet und prototypisch umgesetzt.

Version	1.0
Editor	Dr.-Ing. Amin Hosseini und Marius Bachhofer
Project coordination	Mercedes-Benz AG und ZF Friedrichshafen AG

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Contents

List of Figures

List of Tables

1 Introduction

The document at hand "Deliverable 14: Ein optimiertes Produktivsystem auf Basis eines kontinuierlichen Performance-Monitorings und Auswertens der Benutzerakzeptanz ist vorhanden" describes the optimisation of the system to produce the required data on the basis of Performance monitoring and evaluation of the user acceptance. It contains results of one subproject; TP1 - Data Production.

The chapters 1.1 to 1.3 of this document give a general introduction into the project and its deliverables. They show how the 18 deliverables contribute to the project aim and how they complement to deliver the project outcome.

Chapter 2 gives the detailed content, it present the following results

Results TP1

- E3.5.1.1 Optimiertes Produktivsystem auf Basis eines kontinuierlichen Performance-Monitorings und Auswertens der Benutzer-Akzeptanz
- E3.2.1.1 Erarbeitung und prototypische Umsetzung eines neuartigen DRM-Konzepts zum Handling personenbezogener Sensor-Daten im Rahmen der KI-Entwicklung für autonome Fahrfunktionen

1.1 Project Description

The goal of the KI Delta Learning project is the development of methods and tools for the the efficient extension and transformation of existing AI modules of autonomous vehicles to the challenges of new domains or more complex scenarios. AI modules are the core of the cognitive intelligence of automated vehicles and thus a key technology for ever higher levels of automation of assistance systems up fully to autonomous driving. Therefore, AI modules are of central importance to the future value creation of the German automotive industry. The market launch strategy of the German automotive industry for these assistance systems is proceeding step by step toward ever higher levels of automation and larger areas of application for automation. The focus of the project is the gradual expansion of the domains of application of assistance systems and the associated AI modules, which will be developed in parallel in six directions according to the most relevant use cases.

Thus, within the project, various deltas - the gaps between known domains of applications and new domain areas - are considered including deltas due to improved sensor technology, due to different traffic areas such as highways or construction sites, due to changes in country and corresponding traffic rules and signage, due to new forms of traffic and road users such as e-scooters, due to changing environmental conditions such as day, night, sun or rain, as well as deltas due to the advancement of neural network designs. A stepwise, structured extension of AI modules towards the six mentioned deltas is called "Delta Learning". This will not necessarily involve all six extensions simultaneously. Rather, the automation of assistance systems will gradually increase through efficient Delta Learning. KI Delta Learning aims to incrementally extend AI modules that have already been trained for limited areas and locations of use without completely re-executing the otherwise usual training and optimization process at a very high cost. So far, no sufficiently efficient and stable methods and tools exist for such Delta Learning.

Hence, the focus of the project is on method development. In two orthogonally operating but interlocked subprojects (TP2 and TP3), these methods are developed on the one hand starting from overcoming the deltas under the focus of transfer learning and on the other hand from the didactic approach. Furthermore, questions of the automotive suitability of the developed methods are investigated (TP4) as well as necessary data generated, recorded and processed (TP1).

The fundamental extension of current generations of AI algorithms expected to be achieved within the project enables a decisive leap towards the large-scale realization of autonomous vehicles. Thus, KI Delta Learning represents an important innovation building block for the competitiveness of the German automotive and supplier industry in an increasingly competitive economy.

1.2 General Deliverables Overview

The project will work on the set goals over a period of three years. It pursues the step-by-step improvement of the models and methods to be developed in four successive project steps, the project increments (PI). At the end of each project increment, which goes hand in hand with the defined project milestones, the (interim) results achieved in the work packages are documented in the form of deliverables. A total of 18 deliverables were defined in the VHB, which serve to communicate the results both internally and externally to the funding agency (BMWi, project sponsor). These 18 deliverables were distributed among the four project increments and can be grouped into following topics.

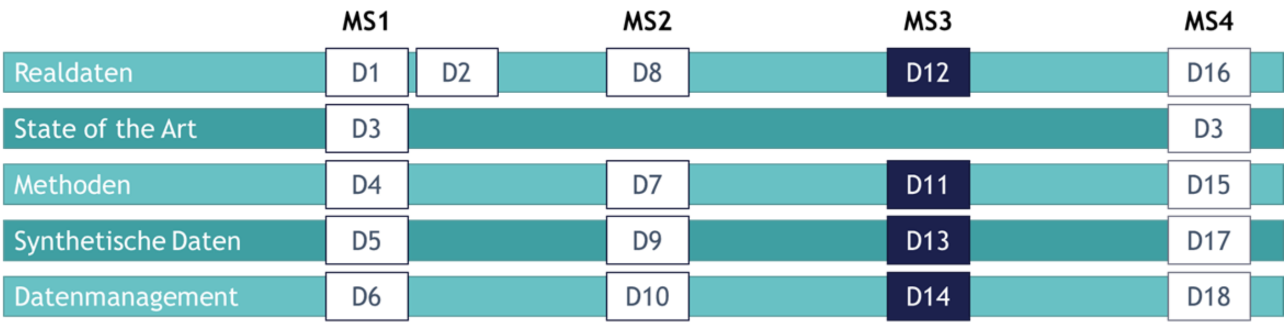

Figure 1.1: Timeline of the Deliverables and grouping to Milestones

The third project increment started and ended with a delay of three months compared the the initial planning as given in the VHB: The end of the third project incerment was shifted from month 27 (03/2022) to month 30 (06/2022). The reason are still the effects of the CoVid-19 pandemic. The results of third project increment are provided in the follwing deliverables, they represent Milestone 3:

1. D11 Eine Defizitanalyse sowie ein Optimierungsplan für die ausgewählten Delta Learning-Methoden mit vorhandenem Datensatz sind vorhanden
2. D12 Datensatz für die Realdaten aus PI3 mit definierten Annotierungen für unterschiedliche Domäne ist vorhanden & für Partner zugreifbar
3. D13 Synthetische Daten zur Erweiterung des Datensatzes für weiterführende Untersuchungen (Corner Cases, etc.) sind vorhanden
4. D14 Ein optimiertes Produktivsystem auf Basis eines kontinuierlichen Performance-Monitorings und Auswertens der Benutzerakzeptanz ist vorhanden

1.3 Deliverable Context

The functional prototype of the system for integrated data management and processing as well as the system for evaluating the user acceptance that was developed and reported was rolled out in PI1 and PI2. In PI3 it was optimised, this is presented in this Deliverable 14, a further optimisation and evaluation will be executed in PI4, the results will be the basis for D18.

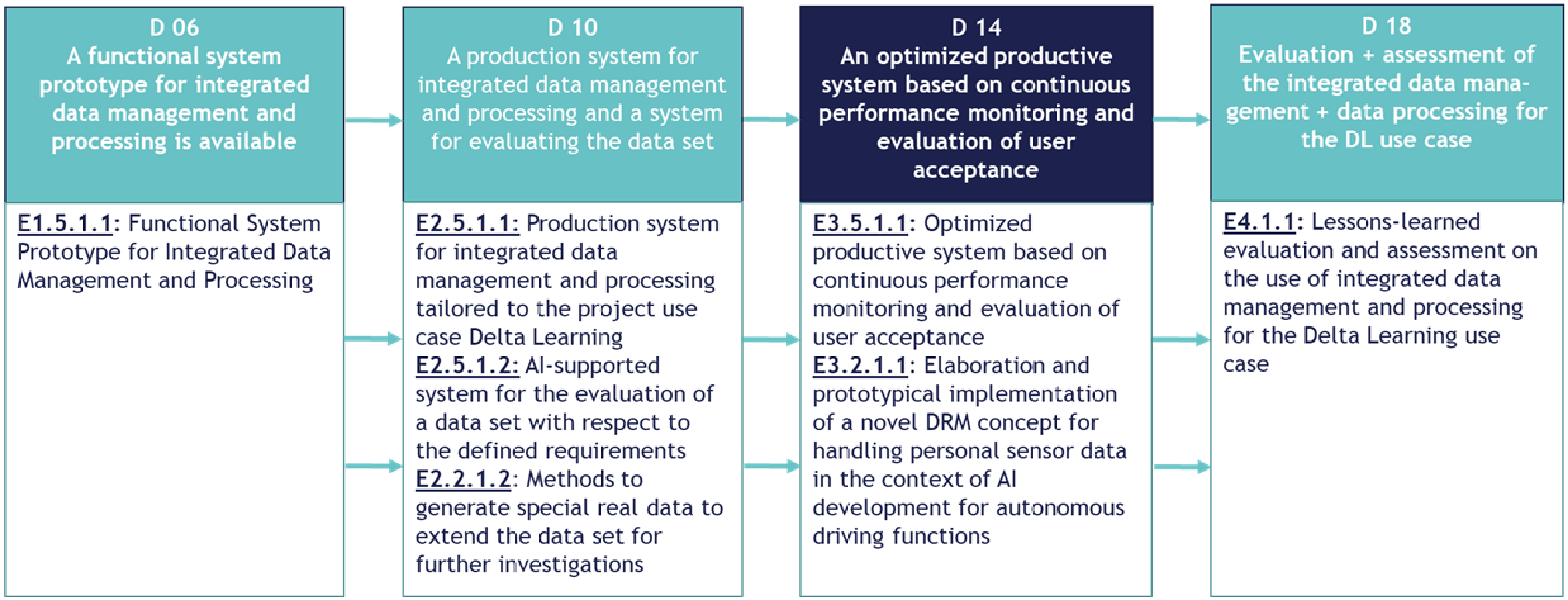

Figure 1.2: Developement of results in the D06/10/14/18 chain over the 4 project increments (PIs)

2 Results TP1

2.1 E3.5.1.1: Optimized productive system based on continuous performance monitoring and evaluation of user acceptance

2.1.1 Introduction

The platform for data management and processing, which was described in the previous Deliverable D10, was optimized and extended, based on user feedback. Changes were made to the browser based frontend, to improve usability, and to the backend where the data is stored and processed. Ultimately, the goal is to improve the performance, flexibility and usability of the platform, to allow for providing high-quality datasets which are the basis for all delta learning activities, in the KI DL project and beyond.

2.1.2 Method

In regular exchange with users of the platform, potential fixes or new features were identified and subsequently implemented to a large extent. Furthermore, statistical data was integrated into the web frontend, so that the users can get an overview of the performance and monitor the process of data processing.

2.1.3 Experiments

The overall layout of the platform for data management and processing as it is currently implemented is shown in the following figure

This schematic overview shows the main process steps for automatic data processing, as well as the interaction by expert users that manually label the data.

The improvements of the platform compared to the initial description in Deliverable D10, as they are described below, refer to this schematic structure.

Data pre-ingest

Newly ingested data is validated and subsequently moved into a dedicated folder structure on the server. In addition, a "markerfile" is created to initialize the following processes.

Data ingestion

For every new recording on the file server, a data set is created in the Meta Data Store. All tagging information of the driver that exists in an accompanying json file is added as meta information. After this process step, all recordings, files, streams and their meta data is visible in the web front-end of the platform, and can be selected and processed.

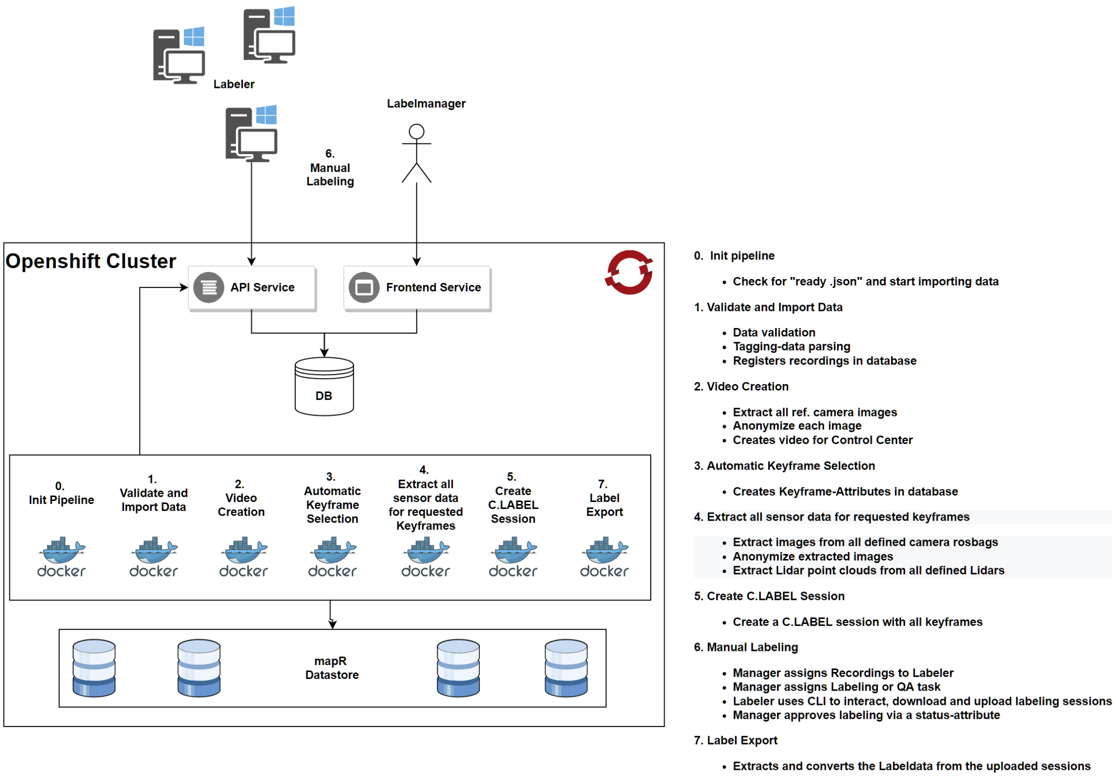

Figure 2.1: **CMore-E3.5.1.1-1:** Schematic overview of docker based post processing

Improve the anonymization process

The anonymization process of video data originally consisted of two individual scripts (models), one for anonymizing faces, and one for the anonymization of number plates. This was leading to memory and performance problems. A new model, which integrates both tasks, was trained, implemented and deployed. This new model led to improved *performance*, as it is approximately 2x faster.

- The model was trained on an extended data set, which was collected and labelled. As a result, the model' s detection *accuracy* was increased.
- More information about the anonymization is described in the result E3.2.1.1

Integrate Key-Frame selection script (Valeo/DLR) to pipeline

Key-frames are those video frames which contain a driving situation which is particularly relevant, and hence need to be marked for further analysis. Key frames can be selected manually or automatically.

- A script to select key-frames was provided by Valeo and DLR. This script was integrated into the pipeline in order to select important frames automatically. The keyframe selection script uses meta-data such as road surface, weather, road/traffic information, and GPS locations to select the most interesting areas of the environment. The aim is to generate as much different data as possible from individual recordings and then use it for labeling

Control Center test and maintenance

The web based user interface ("Control Center") for the data management and processing platform was further tested and improved. In particular, the following issues were identified and resolved:

- Data tags missing
- Video preview missing
- Big recordings (full day drives) are not processed
- Video lacking frames / jumping
- Clean up unnecessary features in Control Center

Control center authentication

In the initial release of the Control Center, there was no dedicated user rights management. In order to define user groups and assign rights to different groups, an authentication was introduced.

- On the web frontend, each user now has to sign in with a username and password, see the screenshot below

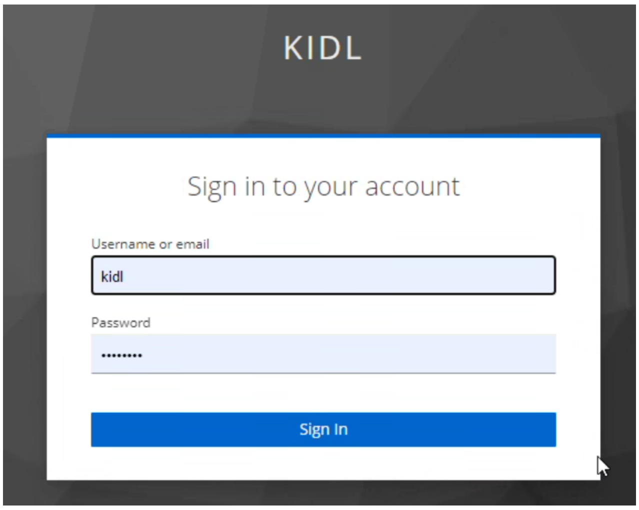

Figure 2.2: **CMore-E3.5.1.1-2:** Screenshot of Log In screen of the control center

- Each user is assigned to one user group. Depending on the user group, the user has different rights to view or edit data. At the moment, two user groups are configured, while in principle, more user groups are possible to manage the access rights in a more fine-grained manner, for example when the workflow is scaled up and more user groups are involved.

Table 2.1: **User groups, access rights**

user group	keyframe setting	keyframe status	attribute management
admin	view, edit	view, edit	view, edit, delete
KIDL-user	view, edit	view, edit	view

Enhance and speed up Labeling Process
The process for image labelling was improved, to make it more efficient. In particular, this was done by the following changes:

- Extend session creation script for automated session split (split into max. 6 frames)
- Extend session export script to session merge
- Enhance the labeling process by integrating Jira into the pipeline.

2.1.4 Discussion & Conclusion

Based on the previous work on the platform for data management & processing, significant steps were taken to improve its usability and performance. This enables CMore to provide a reliable and powerful platform to other partners in the KI DL project.

Furthermore, the platform integrates multiple aspects of the KI DL project which are necessary to generate a high-quality data set, such as data recording & ingestion, editing of the data, and labelling.

2.2 E3.2.1.1: Elaboration and prototypical implementation of a novel DRM concept for handling personal sensor data in the context of AI development for autonomous driving functions

2.2.1 Introduction

Collecting data from the roadways, in particular images/video data, raises data privacy concerns. Government regulations do not allow the use/access to acquired data without protecting personal data. Faces and number plates of traffic participants are of particular interest since the data relates to a person whose identitiy is discernible.

This raises the need for image and video data to be anonymized in an automatic fashion, to allow for efficient processing of the acquired data. Furthermore, it needs to be ensured that key features, which are relevant for analysis of the video data, are maintained, while privacy is ensured.

Finally, the raw video data before anonymization needs to be stored in such a way that it is not accessible to unauthorized people.

2.2.2 Method

Anonymization can be done in various methods. The most well-established ones are:

1. Deep learning approach: a subset of machine learning where neural networks — algorithms inspired by the human brain — learn from large amounts of data. Deep learning algorithms perform a task repeatedly and gradually improve the outcome through deep layers that enable progressive learning. It's part of a broader family of machine learning methods based on neural networks.

2. Computer vision approach: a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs.

For face and license plate detection, we used the deep learning approach because, unlike deep learning, computer vision methods struggle to generalize to new orientations and positions. Deep Learning methods have proven to have better accuracy, speed of training, and normalization. The only drawback of deep learning, however, is the need for labeled data. Nevertheless, we had already 5000 labeled images of license plates, and we labeled around 3000 labeled images of faces.

The anonymization algorithm identifies the region of interest using deep learning-based object detection. In this context, the region of interest is a vehicle' s number plate, or faces of humans. The model (i.e., neural network) takes the frame as an input and outputs rectangular bounding boxes around the detected objects with its class. The bounding boxes are blurred using a Gaussian function.

Model

Transfer learning was used to train the model, i.e., a model already trained for detecting objects was used as a starting point to generate a new model. The baseline model was Faster-R-CNN with FPN backbone (https://arxiv.org/abs/1506.01497v3), which uses standard convolutional layers for mask prediction and fully connected layers for box prediction. The models are trained on COCO 2017 dataset which had 123,287 images and 886,284 instances (https://paperswithcode.com/dataset/coco).

Dataset

We have collected around 6000 images and labeled the faces and license plates in those images. The dataset was split into training (80%), and validation (20%). The model was fine-tuned to detect the two classes (faces, and license plates) and validated on the validation dataset.

2.2.3 Experiments

To demonstrate the approach, the following images can be compared before and after the anonymization process. After the anonymization, a rectangular area around the number plate is blurred.

2.2.4 Discussion & Conclusion

From the concept and the prototypical implementation of an anonymization method for images, which was described above, the key results can be described as:

Table 2.2: **Average Precision in detecting faces and number plates:**

category	AP
License	52.856
Face	23.847

Performance

The model was tested on a GPU RTX 2080 TI and the inference time was recorded to be 0.13 sec/image.

Figure 2.3: **CMore-E3.2.1.1-1:** Example of anonymized licence plate