

### Use-Case in Delta Learning

Collected and generated data needs to be managed in order to facilitate the data management and processing for the consortium.

End users need insight into the content of available data, allowing them to determine and search for relevant time sequences within the recorded data. Once a set of recorded data is selected, the data sets can be used for further processing.

### Technical Problem

As several members of the consortium need to be integrated at this stage, a common notion of data is needed. Moreover, for the AD use case, with related sensor data collection, arises the need to cope with large amounts of data. This requires being able to scale workloads on the compute platform without making the integration for the participants too complex. Additional constraints deriving from the use case are privacy related restrictions, e.g. raw data containing sensitive data like number plates or faces can't be used directly.

### Technical Solution

The next generation C.DATA solution has the control center as central coordination station for the end user. The Javascript based modular application is running in a browser and allows the user to manage data as well as related workflows.

The underlying data has been modeled with a non-strict hierarchical structure, containing AD domain specific elements like projects, recordings, files, streams and attribute definitions.

The developed backend system can also attach metadata in form of attributes to recordings and streams, so that the data can be enriched, and the data lake becomes searchable.

### Recording

#### Attribute Value

Driver = Thomas Mustermann  
State = annotated

#### Scene Attribute Value

Scene = Cut-In, from 1000 – 1500  
Key-Frame = true, 300-300  
Key-Frame = true, 400-400

Figure 2: Metadata Structure

The system also enables the user to define re-usable static and dynamic data-sets which are used in data driven workflows. To evaluate the coverage of a training or validation data set, a statistic API was developed.



Figure 3 Statistic API

### Evaluation

So far, performance has been evaluated with large synthetic datasets. The meta data store (MDS) has elastic search as backend and has proven to be able to scale and deal with large amounts of data. The MDS domain model is also used to manage the fully automated annotation workflow and the anonymization workflow for the recorded video streams.

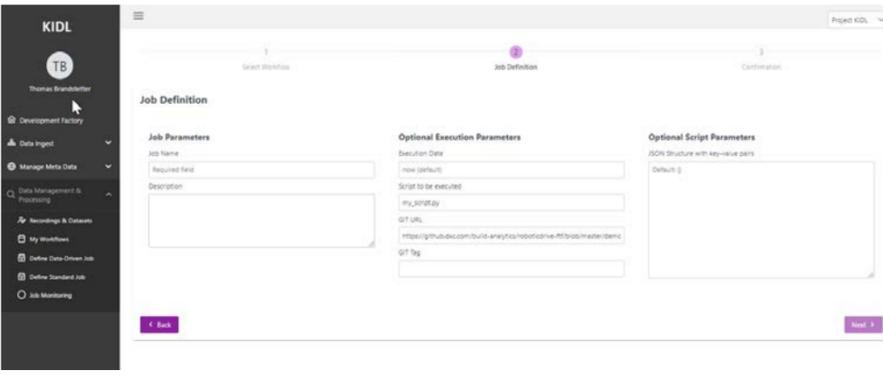


Figure 1: Control Center, UI for data and workflow management

**Luxoft**  
A DXC Technology Company

For more information contact:  
thomas.brandstetter@dxc.com,  
emanuele.dettori@dxc.com,  
roshan.muthaiya@dxc.com

### Partners



### External partners



KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Energy.



Supported by:  
  
Federal Ministry  
for Economic Affairs  
and Energy

on the basis of a decision  
by the German Bundestag