



Data Handling and Sharing

L3Pilot Final Event

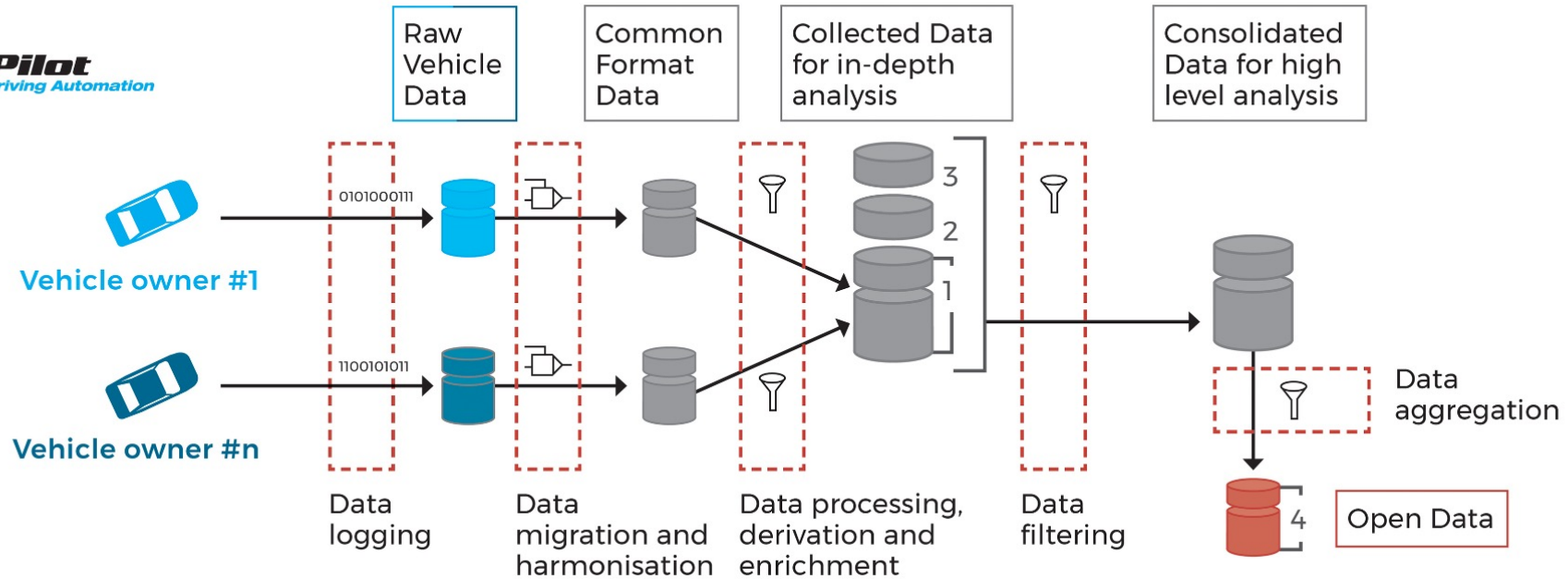
Johannes Hiller, ika RWTH

On behalf of data team (Natalia, Erik, Frederic, Sami, Markus and many more)



L3Pilot Common Data Format

Data Flow



← Vehicle owner specific - activities in SP6 → → SP6 data management - SP7 analysis →

Categories of data:

- 1 Derived Vehicle Data (CAN, GPS, Pls, video, and/or video annotations)
- 2 Subjective Data (interviews, questionnaires, simulator studies)
- 3 External Data (weather, map, ...)
- 4 Open Data (aggregated data)

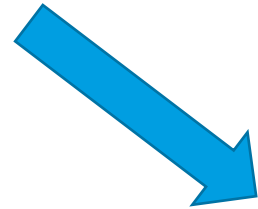
Tools provided by SP5

© L3Pilot

L3Pilot Common Data Format Signal Requirements

RQ Level 3	Logging requirements / sensors available	Performance Indicators Required
The ADF will work as expected in its planned driving and traffic scenario	Vehicle data + additional data sources (if information are not provide by the car)	System status, distribution of longitudinal velocity per driving scenario and situation variables
How often do unexpected take-over requests occur?	Vehicle data + additional data sources (if information are not provide by the car)	System status, distribution of longitudinal velocity per driving scenario and situation variables
Does the function initiate a take-over request if required by the boundaries of the ADF?	Vehicle data + additional data sources (if information are not provide by the car)	System status, distribution of longitudinal velocity per driving scenario and situation variables
What is the impact of ADF on the frequency of traffic violations?		Distribution of difference between speed and speed limit, distribution of distances to other objects, frequency of overtaking manoeuvres in overtaking prohibitions
How do take-over situation affect the driving dynamics of the vehicle?		Distribution of lateral and longitudinal acceleration and velocity
What is the impact of ADF on longitudinal acceleration in defined driving situations?		Distribution of logitudinal acceleration
What is the impact of ADF on lateral acceleration in defined driving situations?		Distribution of lateral acceleration

Research questions specified with logging requirements by Methodology Partners



External	Precipitati	Precipitation	-	The intensity of precipitation at the current time step	1	-	zoh	-1	-1	-1	-1	-1	1	-1	-1	?					
Ego vehic	Steering ra	SteeringRackForce	N	Steering force applied by the EPAS system on the steering rack	10	0.1N	linear	1	-1	-1	-1	-1	1	-1	-1	-1					
Ego vehic	Vehicle ba	VehiclebasedNDRA	-	Type of non-driving related activity performed by the driver using a (informi	1	-	-	-1	1	-1	-1	?	-1	-1	-1	?					
External	Level of se	LOS	-	The level of service according to LOS A (0) to LOS F (5), not applicable (6)	1	-	zoh	-1	-1	-1	-1	?	1	-1	?	-1					
External	Traffic den	TrafficDensity		Number of vehicles per km in a bounding box around the vehicle	1	1 vehicles / km	zoh	-1	-1	-1	-1	-1	1	-1	?	?					
External	Type of ne	TypeIntersection			1	-	-	-1	-1	-1	-1	-1	1	-1	-1	-1					
External	Distance tr	DistIntersection	m		10	-	linear	-1	-1	-1	-1	-1	1	-1	-1	-1					
External	Priority rule	Video	Back outsi	Back outside view	-	Driving scene in behind the vehicle; angle should be wide enough to see	10 fps	640x480					1	1	1	?	?	1	1	1	1
Ego vehic	Energy cor	Speed limit	SpeedLimit	The speed limit at the current location of the ego vehicle	1		zoh	1	1	1	-1	1	1	1	1	1	1	1	1	1	
Objects	Heading	Ego vehic	ESC interu	ESCIntervention	-	Whether the Electronic Slip Control / Traction Control system is intervening	10	-	zoh	1	1	1	1	1	1	?	1	1	1	1	
External	Road surf	Ego vehic	Steering w	SteeringAngleRate	rad/s	The turning rate of the steering wheel	10	0.1rad/s	linear	1	1	1	1	1	1	1	?	1	1	1	
Ego vehic	Rear wiper	Objects	ID	ID	-	Unique ID of the object	10	-		1	1	1	1	1	1	?	1	1	1	1	
External	Precipitati	Objects	Longitudin	LongPosition	m	Position of the object in longitudinal direction in the ego coordinate system	10	0.1m	linear	1	1	1	1	1	1	?	1	1	1	1	
External	Constructi	Objects	Lateral pos	LatPosition	m	Position of the object in lateral direction in the ego coordinate system	10	0.1m	linear	1	1	1	1	1	1	?	1	1	1	1	
Meta dat:	Driver ID	DriverID	-	The driver ID. Coding is up to the OEM - should be agreed with selected ps	-		-			1	1	1	1	1	1	?	1	1	1	1	
Meta dat:	Test site ty	TestSiteType	-	Type of test site for the trip	-		-			1	1	1	1	1	1	?	1	1	1	1	
Meta dat:	Eligible for	AnalysisEligible	-	Whether the logfile should be used for analysis. For example, logfiles may	-		-			1	1	1	1	1	1	?	1	1	1	1	
Ego vehic	Time	Time	s	UTC timestamp since epoch	10	0.1seconds	-			1	1	1	1	1	1	1	1	1	1	1	
Ego vehic	Odometer	Odometer	km	The vehicle odometer reading	1	0.001km	linear	1	1	1	1	1	1	1	1	1	1	1	1	1	
Ego vehic	Speed	VehicleSpeed	m/s	Speed of the ego vehicle as reported by the ABS / wheel sensing module	10	0.1m/s	linear	1	1	1	1	1	1	1	1	1	1	1	1	1	

Signals derived and sent out to discussion with Pilot Leaders



Available signals assessed and fixed for Common Data Format

ADFunctionActive	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ADFunctionAvailable	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TOR	-	-1	-1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FrontWiperStatus	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FrontFogLightStatus	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RearFogLightStatus	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

L3Pilot Common Data Format Requirements

- Portability
 - Transfer of data from vehicle owner to analysis partner
- Compatibility
 - Many different platforms used by vehicle owners
 - Many different languages used in development and analysis

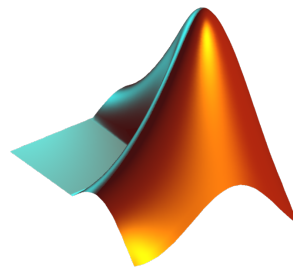


[1]

[2]



[3]



[4]



[5]



Discover the Experience

[6]



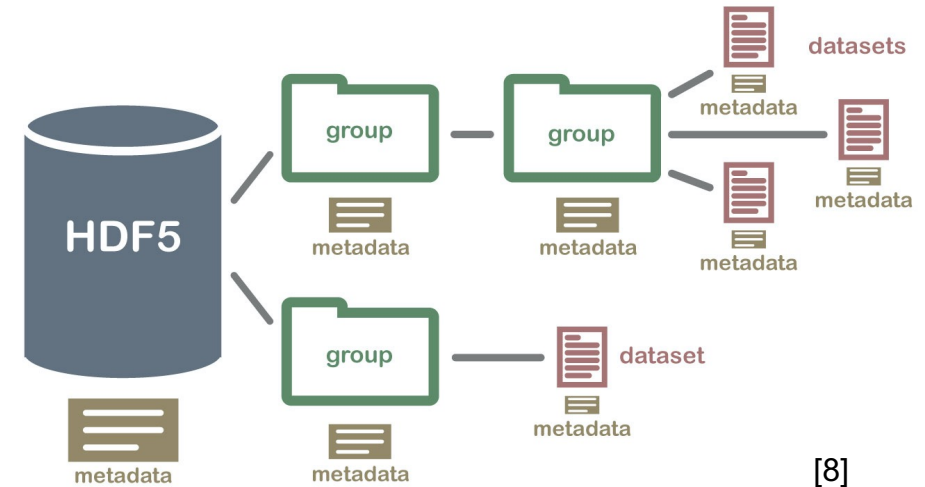
[7]



L3Pilot Common Data Format

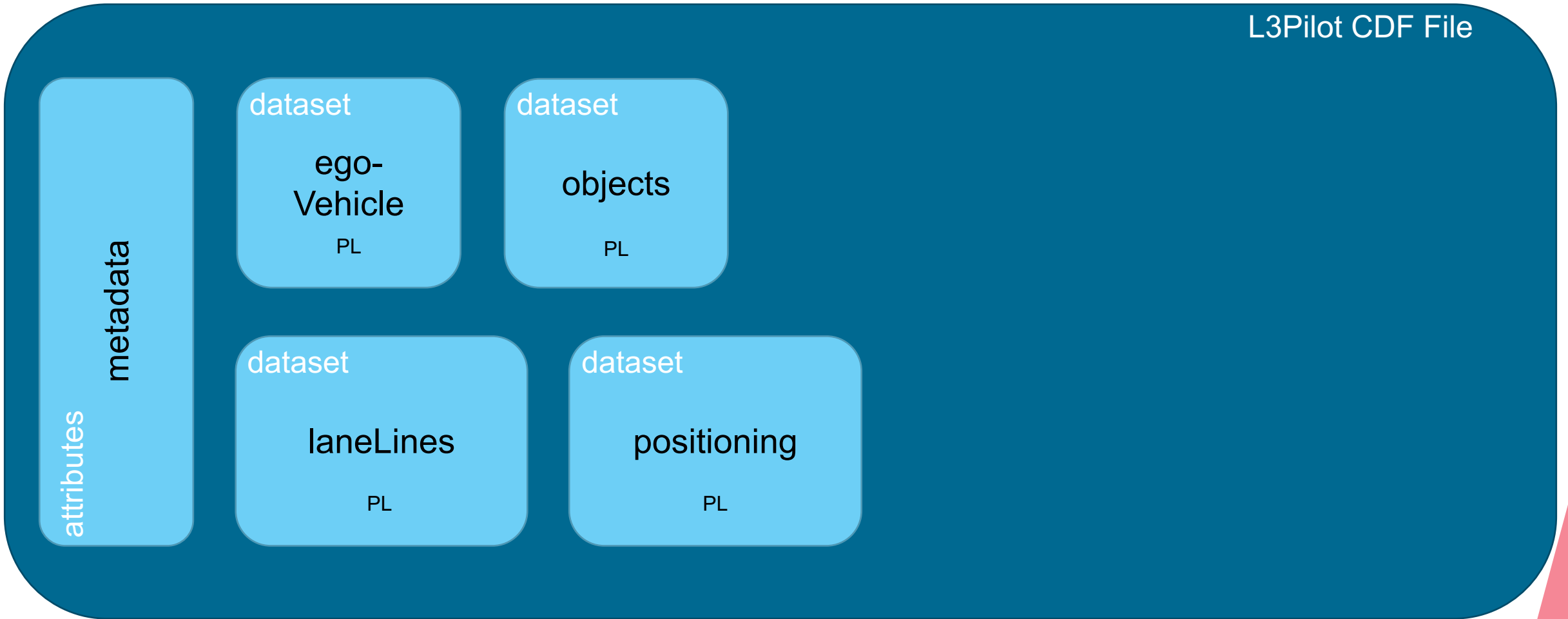
HDF5

- Many file formats were checked
- Hierarchical Data Format (HDF) was selected
- Portable, binary format
 - Compression optional
- Open source and free to use
- Available for many platforms & languages
 - Windows, Linux, ...
 - Matlab, C/C++, Python, Java, ...



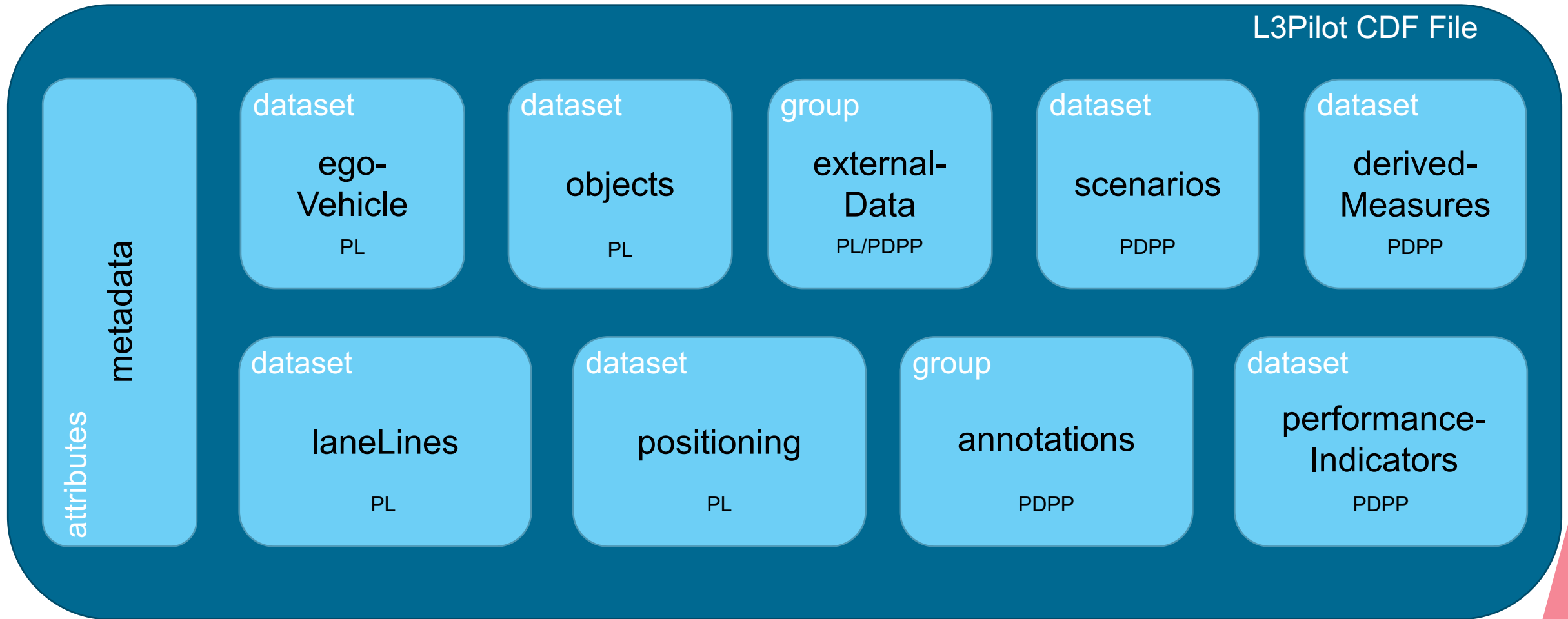
L3Pilot Common Data Format Structure with Vehicle Signals

L3Pilot CDF File



PL – Pilot Leader

L3Pilot Common Data Format Structure with Vehicle Signals

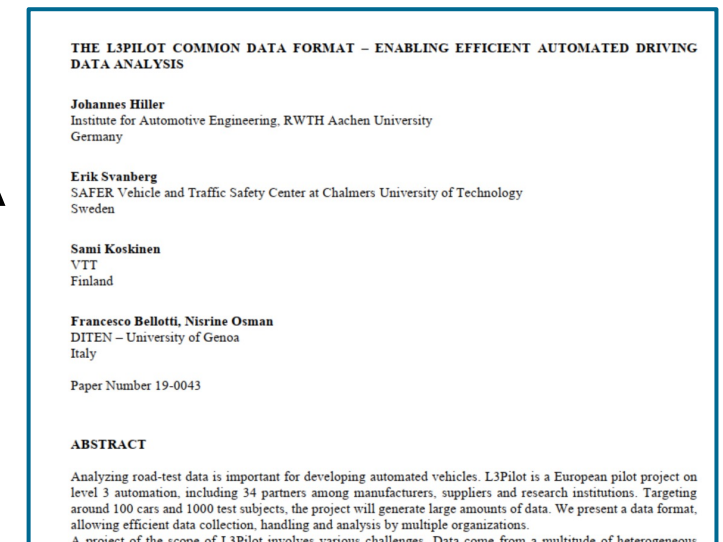
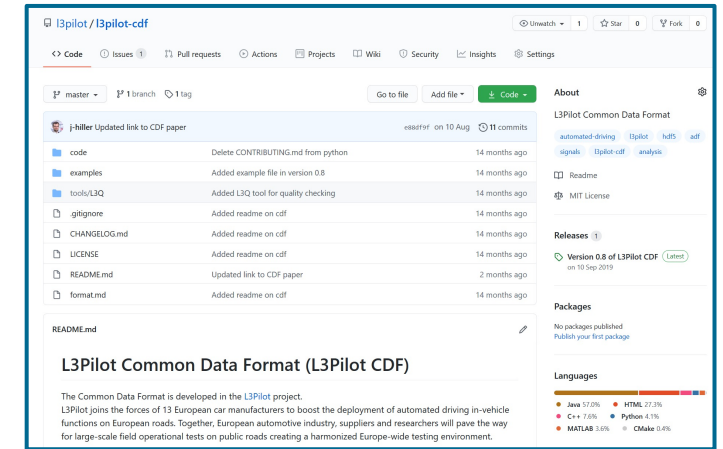


PL – Pilot Leader

PDPP – Pilot Data Processing Partner


L3Pilot Common Data Format (L3Pilot CDF)

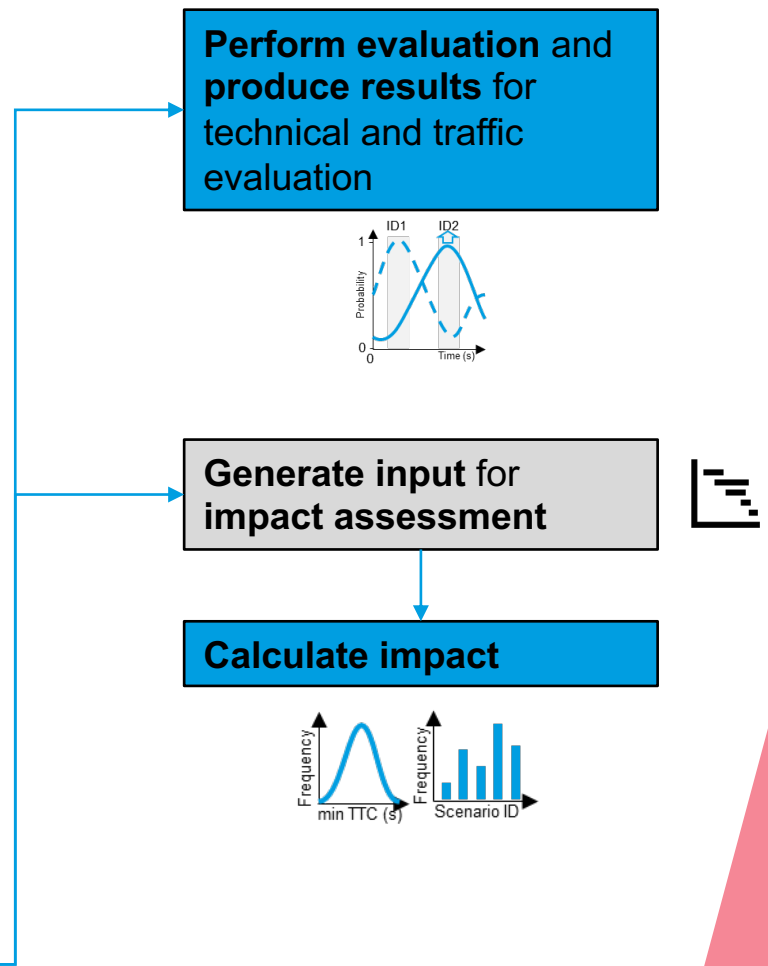
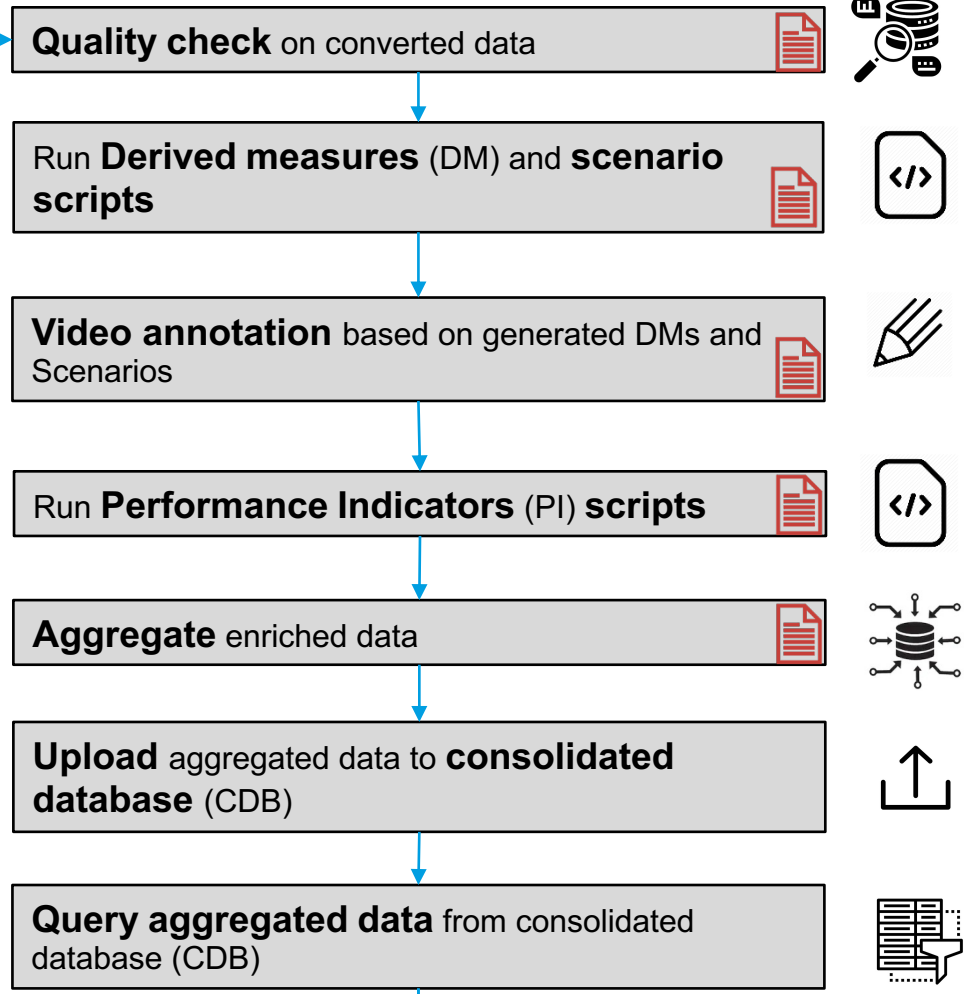
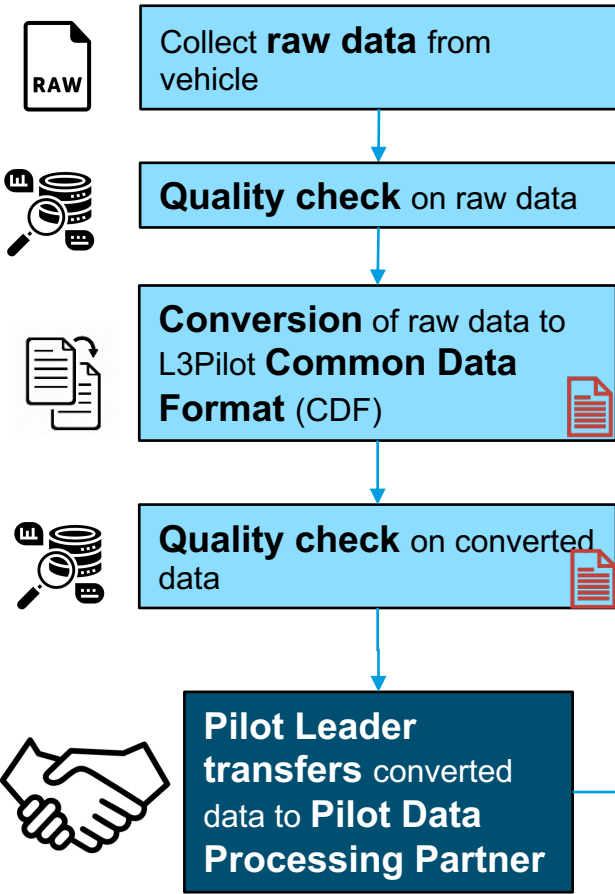
- The CDF is **made available to the public** via Github: *l3pilot/l3pilot-cdf*
 - Everyone is invited to use the format and contribute to it
 - Use open source tools and formats to facilitate use in other projects
- Detailed information on the format can be found on NHTSA website (public access)
 - “The L3Pilot Common Data Format – Enabling Efficient Automated Driving Data Analysis”



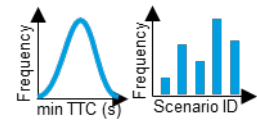
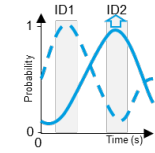
 **Pilot Leader**

 **Pilot Data Processing Partner**

 **Evaluation Partners**

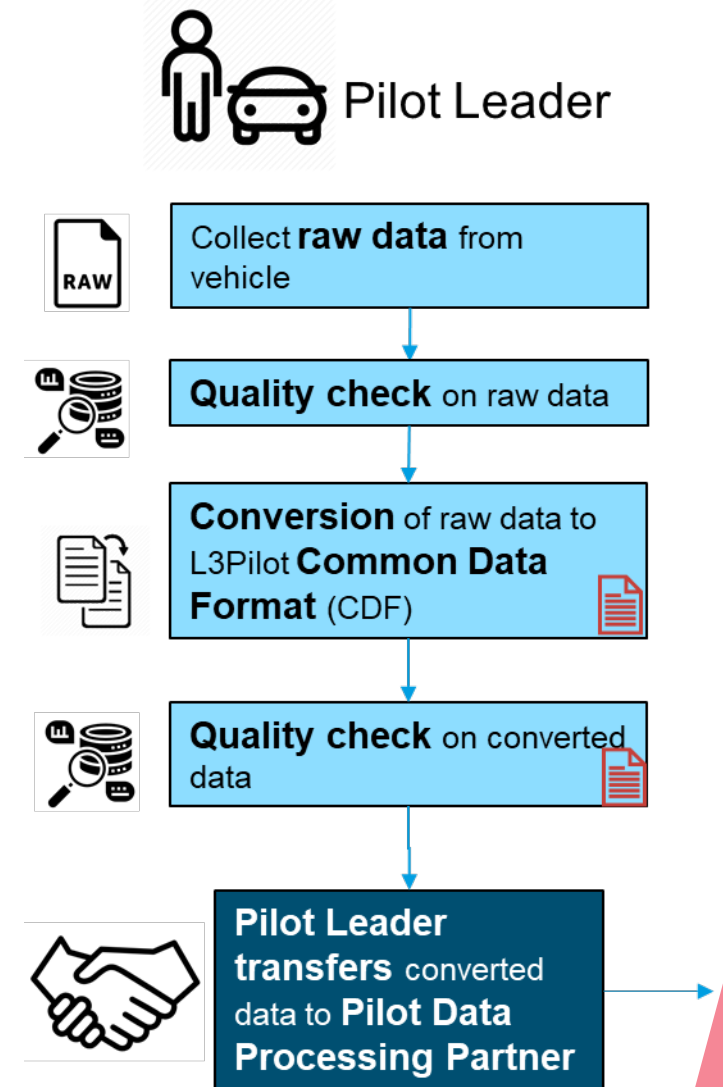


 - L3Pilot CDF



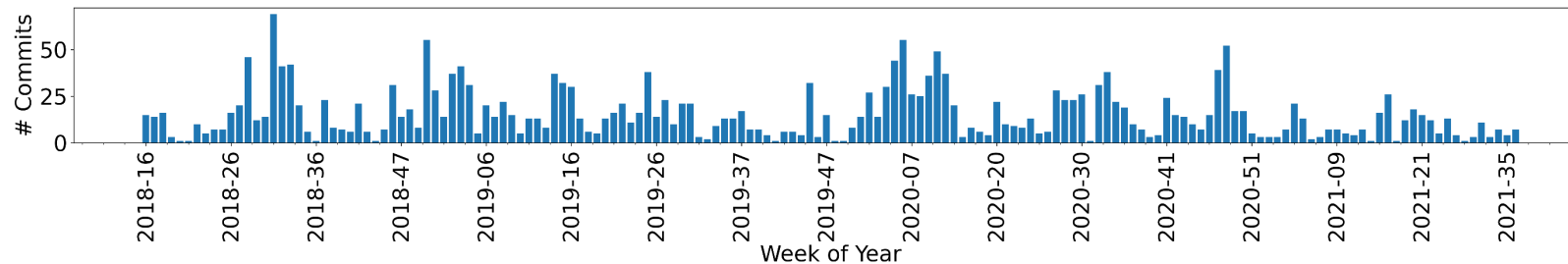
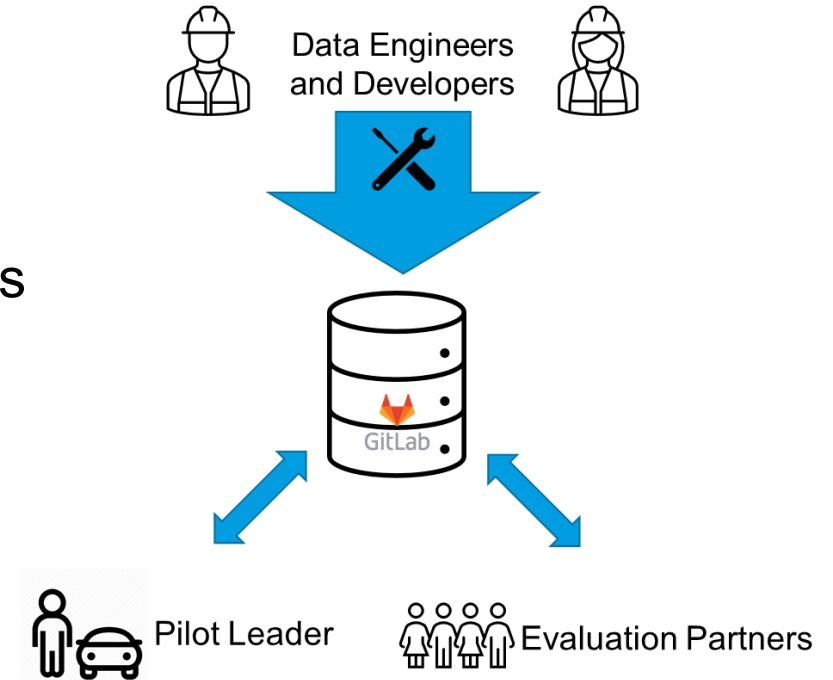
Data Quality Checks

- List of common data quality mishaps and tests
 - Best practices
 - Pilot Leaders encouraged to follow checks
- Consisted of 28 mandatory and 4 optional checks
 - Sensors, Enumerations and Format
 - Aggregated Data and Scenarios
- In the end, some data quality issues were still discovered during evaluation phase
 - Very time-consuming, if vehicle data must be re-processed



L3Pilot Data Tools Development

- Various software tools were developed and improved within L3Pilot
- 2481 single commits to the Gitlab Server in 21 repositories
 - Over 2 commits per day on every weekday in project duration
- Nearly 700 issues opened, discussed, fixed and closed
- International team of over 20 developers



Pseudonymization Process

Personal information

First name	Last name	Driver ID	SHA256(Driver ID + salt)	Age	Gender	Nationality	...
David	Davidson	001	1a064a72...1afe5341	26	Male	Earth	...
Stan	Stanson	002	b2452fbb...02753647	38	Male	Mars	...
Nelly	Nelson	003	339b0d9a...212960be	29	Female	Venus	...

Master table of participants which only Pilot Leader has access to

SHA256 Driver ID is only link between these

H5 file	Start timestamp	End timestamp	Driver ID	Trip ID (SHA256 of start timestamp + salt)	var
2019-04-23.h5	1530364592574	1530364652474	1a064a72	691ea24d...b79b5524	...
2019-04-24.h5	1530364651438	1530364677843	b2452fbb	4a8c43dd...1d228db6	...
2019-04-25.h5	1530364651778	1530364651234	339b0d9a	dc0d005d...4e26967a	...

H5 data for Pilot Data Processing Partner

Consolidated database (Consortium has access)

Trip ID	Max/min/avg var
691ea24d	...
4a8c43dd	...
dc0d005d	...

Needed for database mechanics

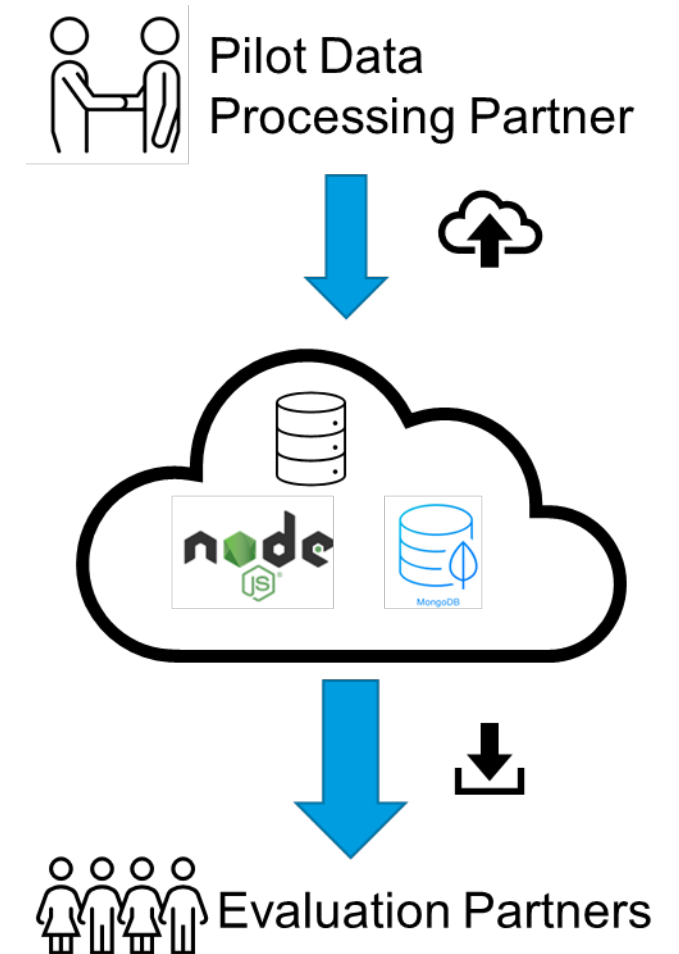
Confidentiality

Data Detail

Consolidated Database

Overview

- Interfaces for storing, managing and retrieving vehicular measurements.
- Big data management based on non-relational database
- Includes run-time checks for data integrity
- User roles, to guarantee the needed confidentiality levels
- Made available through cloud provider
- Tools for accessing and up- and downloads available to evaluation partners
- The L3Pilot CDB is a configured instance of the open source Measurify IoT development framework: <https://measurify.org/>



Conclusions

- Collaborative work of **over 20 international developers**
- Developed a harmonized **Common Data Format** used by pilot leaders and pilot data processing partners
- Made Common Data Format **open source** and available to interested parties
- Developed a set of **harmonized tools** for **converting, transferring and evaluating** piloting data



Thank you for your kind attention.

Johannes Hiller
ika RWTH Aachen
johannes.hiller@ika.rwth-aachen.de



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723051.