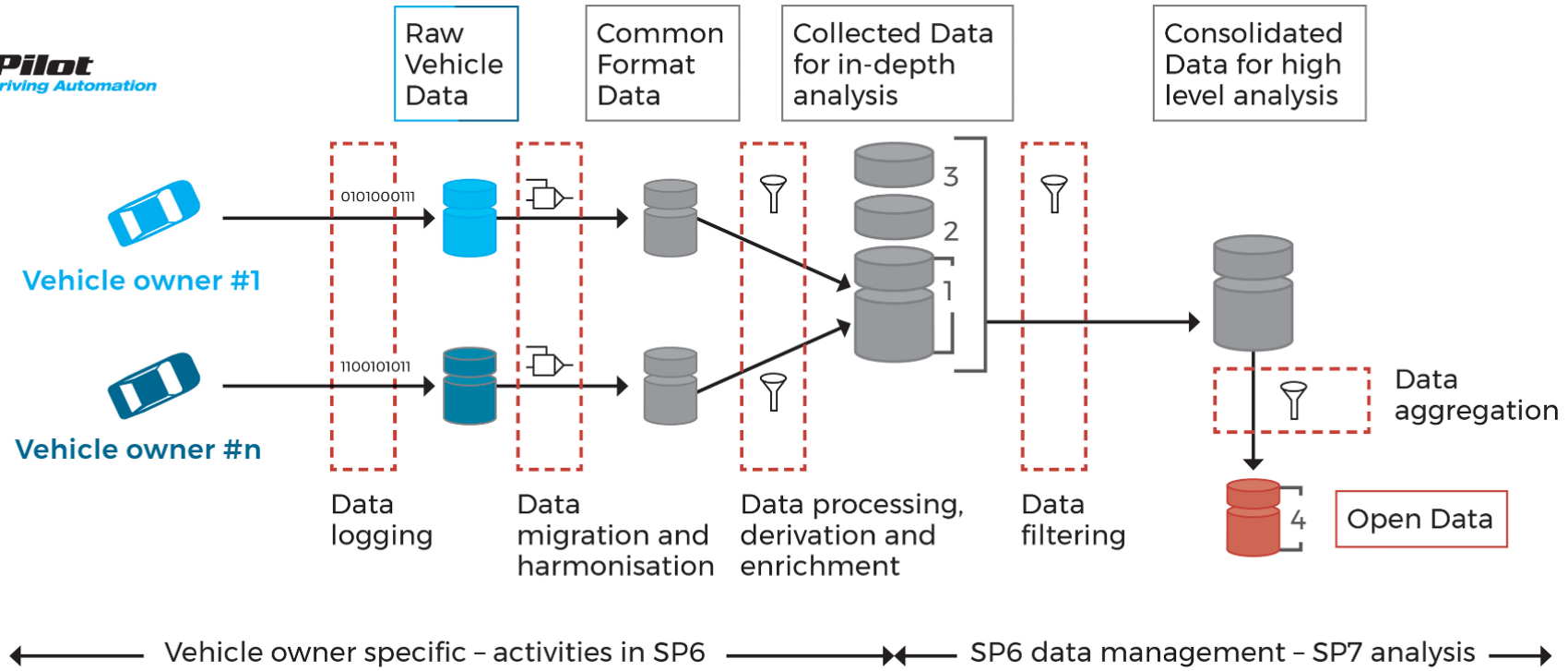
Common Data Format for AD Testing

PS1122 - Comprehensive tools for advancing Automated Driving systems,
Virtual ITS European Congress, 9 November 2020

Johannes Hiller, ika RWTH Aachen

L3Pilot Common Data Format

Data Flow



Categories of data:

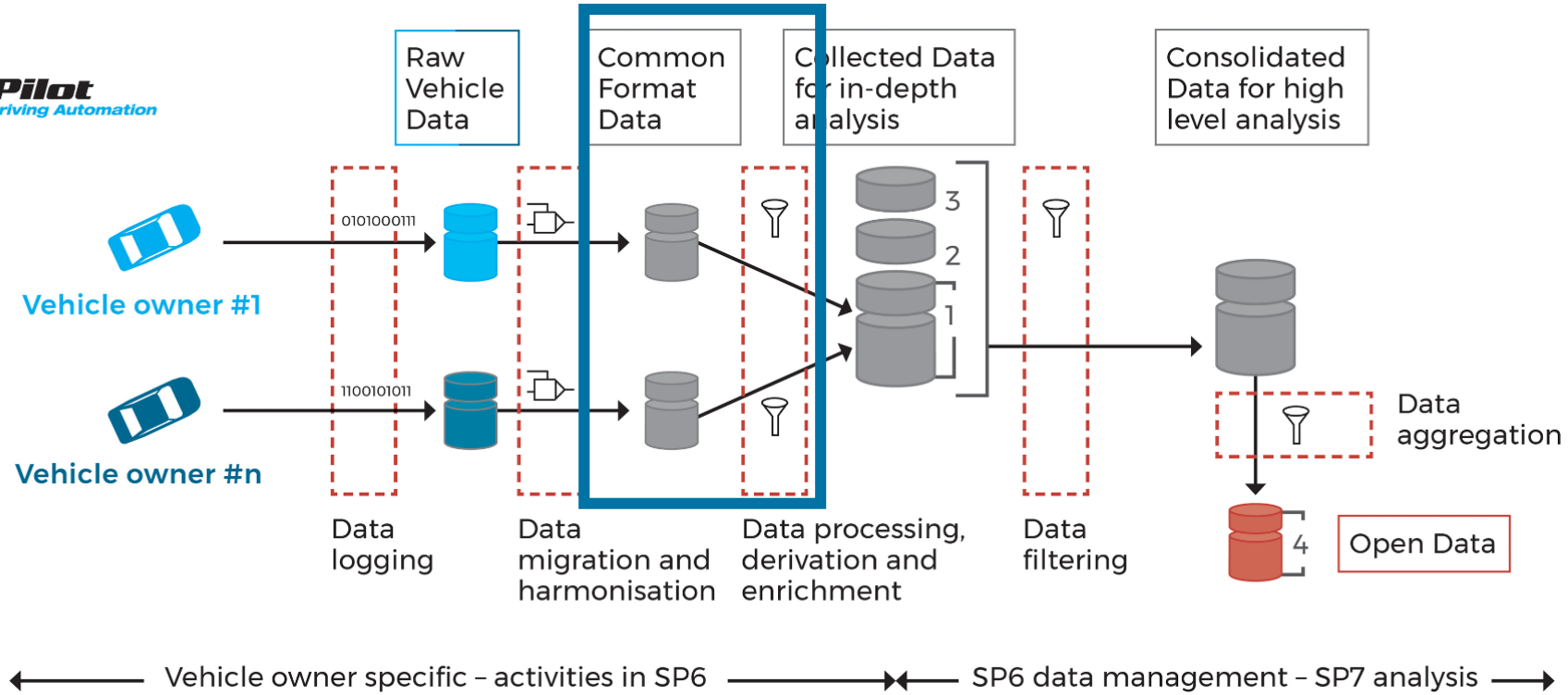
1 Derived Vehicle Data (CAN, GPS, PIs, video, and/or video annotations)
2 Subjective Data (interviews, questionnaires, simulator studies)
3 External Data (weather, map, …)
4 Open Data (aggregated data)

Tools provided by SP5

© L3Pilot

L3Pilot Common Data Format

Data Flow



Categories of data:

1 Derived Vehicle Data (CAN, GPS, PIs, 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

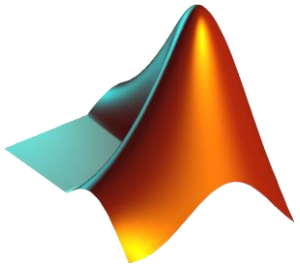
L3Pilot Common Data Format
Requirements

- 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]

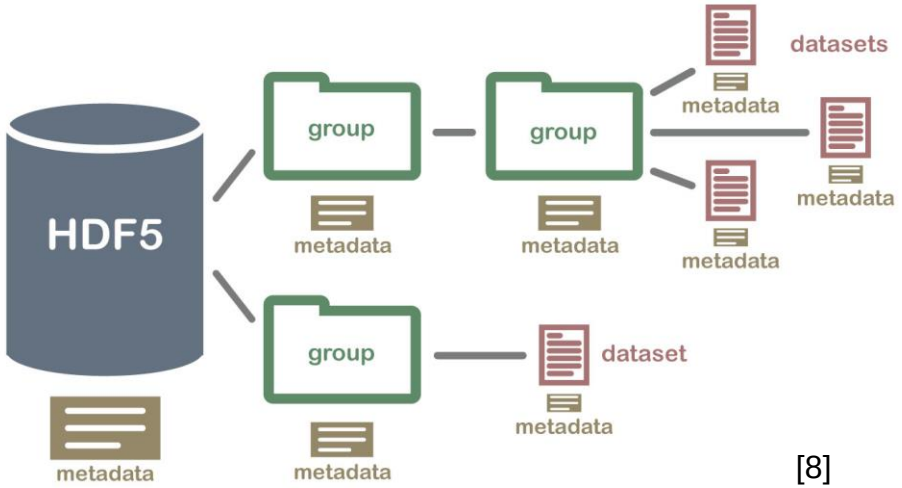
[5]

[6]

[7]

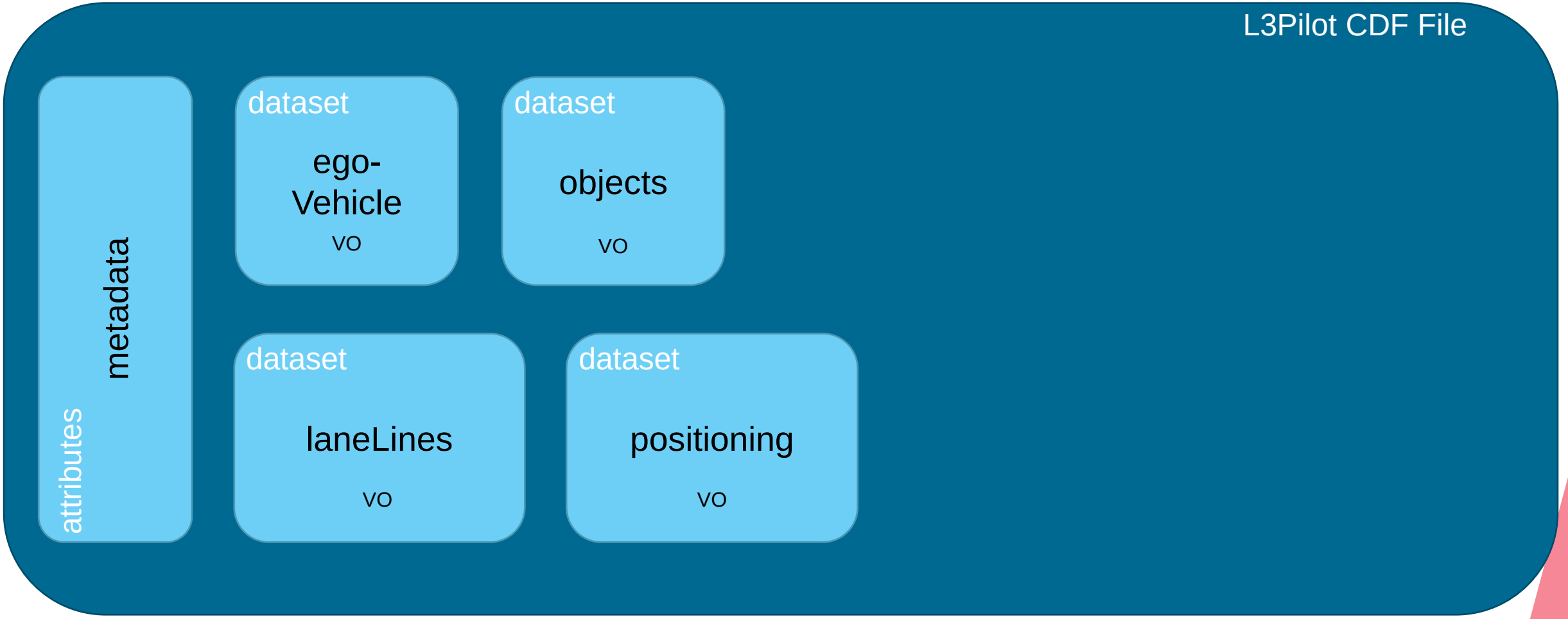
L3Pilot Common Data Format
HDF5

- 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, …



[8]

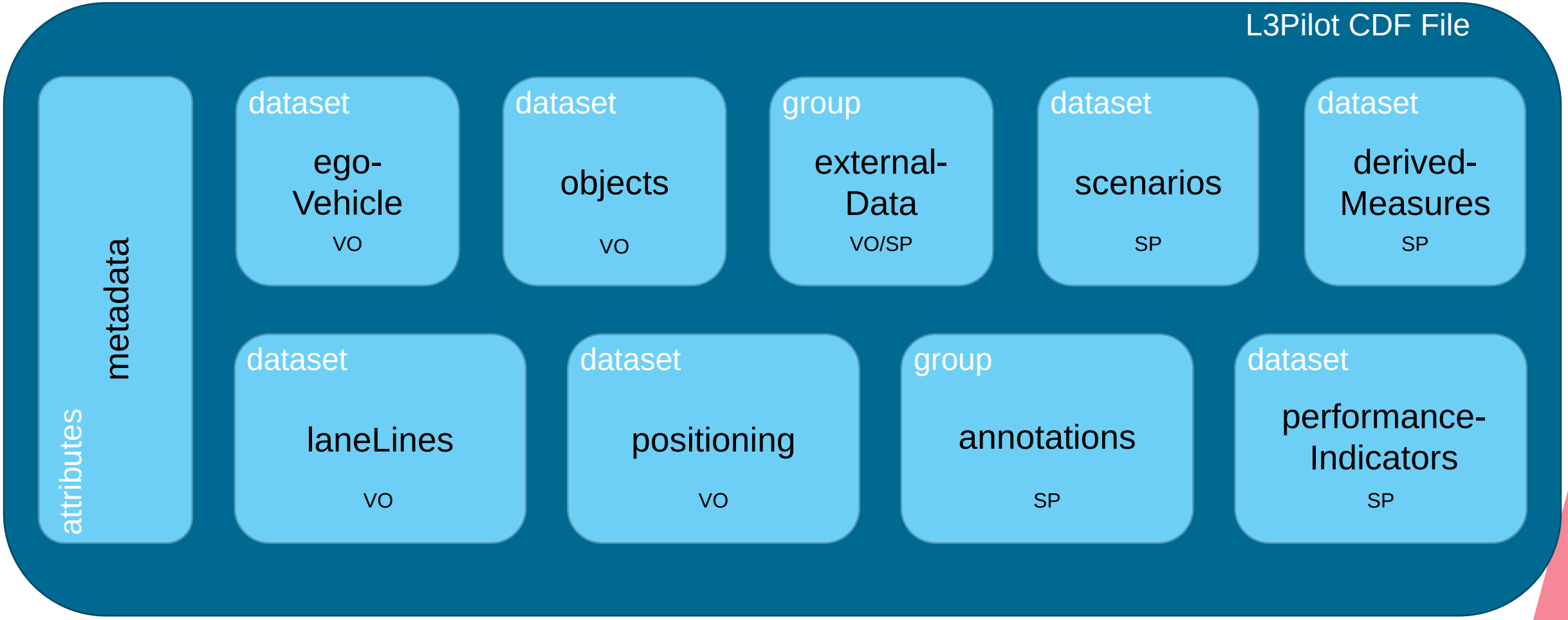
L3Pilot Common Data Format
Structure with Vehicle Signals



VO – vehicle owner

L3Pilot Common Data Format

Structure with Vehicle Signals



VO – vehicle owner

SP –selected partner

- CDF

Vehicle owners

Collect **raw data** from vehicle

Quality check on raw data

Conversion of raw data to L3Pilot **Common Data Format** (CDF)

Quality check on converted data

Vehicle owner transfers converted data to **selected partner**

Selected partners

Quality check on converted data

Run **Derived measures** (DM) and **scenario scripts**

Video annotation based on generated DMs and Scenarios

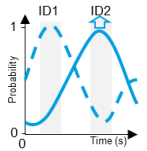
Run **Performance Indicators** (PI) **scripts**

Aggregate enriched data

Upload aggregated data to **consolidated database** (CDB)

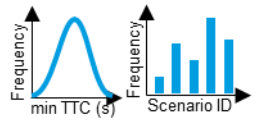
Query aggregated data from consolidated database (CDB)

Perform evaluation and **produce results** for technical and traffic evaluation



Generate input for **impact assessment**

Calculate impact



Pseudonymization Process

Confidentiality

Data Detail

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 OEM has access to

SHA256 driverID is only link between these

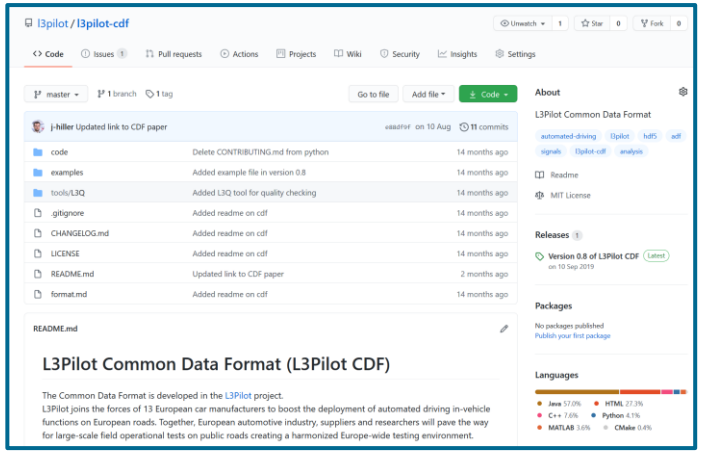
H5 file	Start timestamp	End timestamp	Driver ID	Trip ID (SHA256 of start timestamp + salt)	var
2019-04-23.h5	1530364592574	1530364652474	1a064a72	691ea24d…b79b5524	…
2019-04-24.h5	1530364651438	1530364677843	b2452fbb	4a8c43dd…1d228db6	…
2019-04-25.h5	1530364651778	1530364651234	339b0d9a	dc0d005d…4e26967a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

H5 data for selected partner

Consolidated database (Consortium has access)

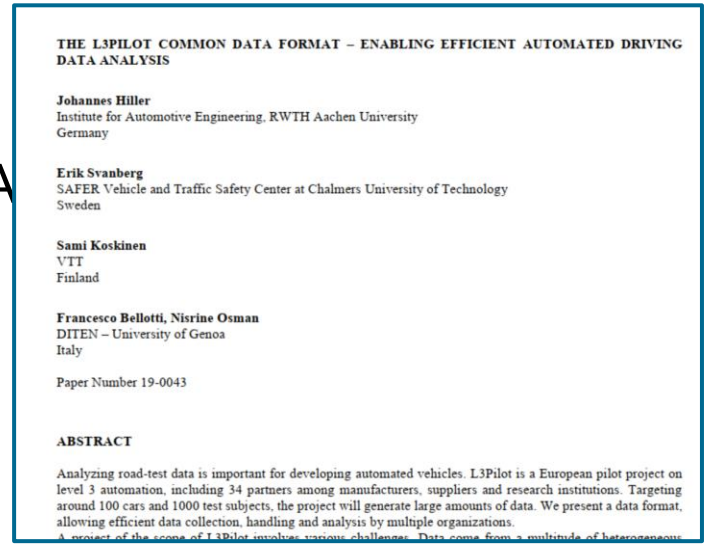
Trip ID	Max/min/avg var
691ea24d	…
4a8c43dd	…
dc0d005d	…

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”


THE L3PILOT COMMON DATA FORMAT – ENABLING EFFICIENT AUTOMATED DRIVING DATA ANALYSIS

Johannes Hiller
Institute for Automotive Engineering, RWTH Aachen University
Germany

Erik Svanberg
SAFER Vehicle and Traffic Safety Center at Chalmers University of Technology
Sweden

Sami Koskinen
VTT
Finland

Francesco Bellotti, Nisrine Osman
DITEN – University of Genoa
Italy

Paper Number 19-0043

ABSTRACT

Analyzing road-test data is important for developing automated vehicles. L3Pilot is a European pilot project on level 3 automation, including 34 partners among manufacturers, suppliers and research institutions. Targeting around 100 cars and 1000 test subjects, the project will generate large amounts of data. We present a data format, allowing efficient data collection, handling and analysis by multiple organizations.

Thank you for your kind attention.

Johannes Hiller
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.