Fraunhofer IESE

PRO OPT

Smart Data

Gefördert durch:
Bundesministerium für Wirtschaft und Energie
aufgrund eines Beschlusses des Deutschen Bundestages

Big Data meets Big Data

Wie die Integration von Big-Data-Lösungen über Unternehmensgrenzen gelingt

Torsten Lenhart
Fraunhofer IESE

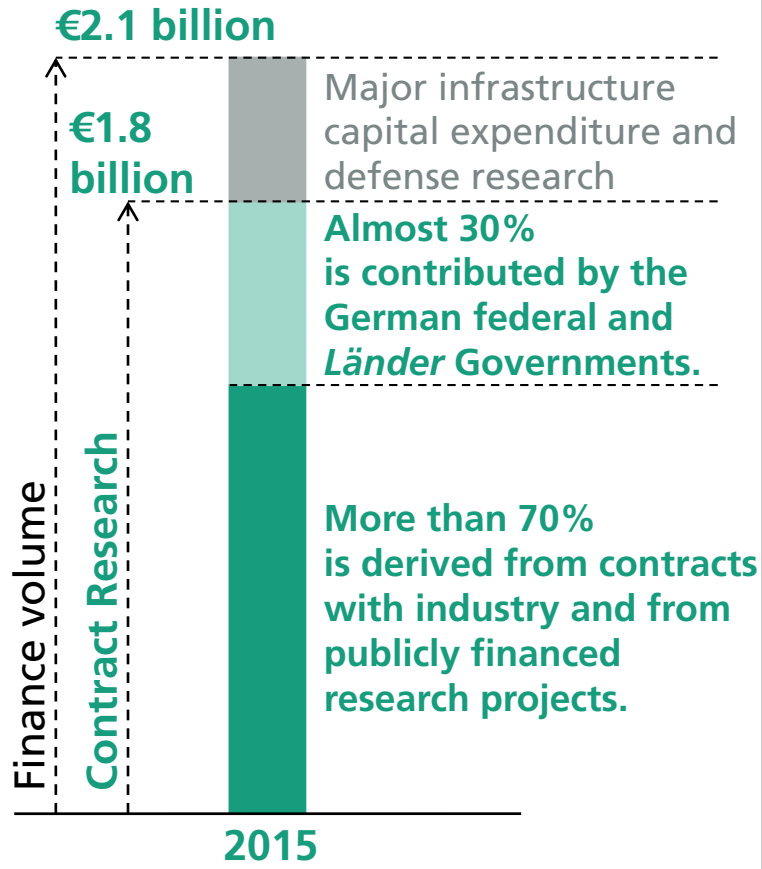
Architekturen 2016
Hildesheim, 23.06.2016

The Fraunhofer-Gesellschaft at a Glance

The Fraunhofer-Gesellschaft undertakes applied research of direct utility to private and public enterprise and of wide benefit to society.

24,000 staff

67 institutes and research units



© Fraunhofer IESE

Fraunhofer IESE

Fraunhofer IESE

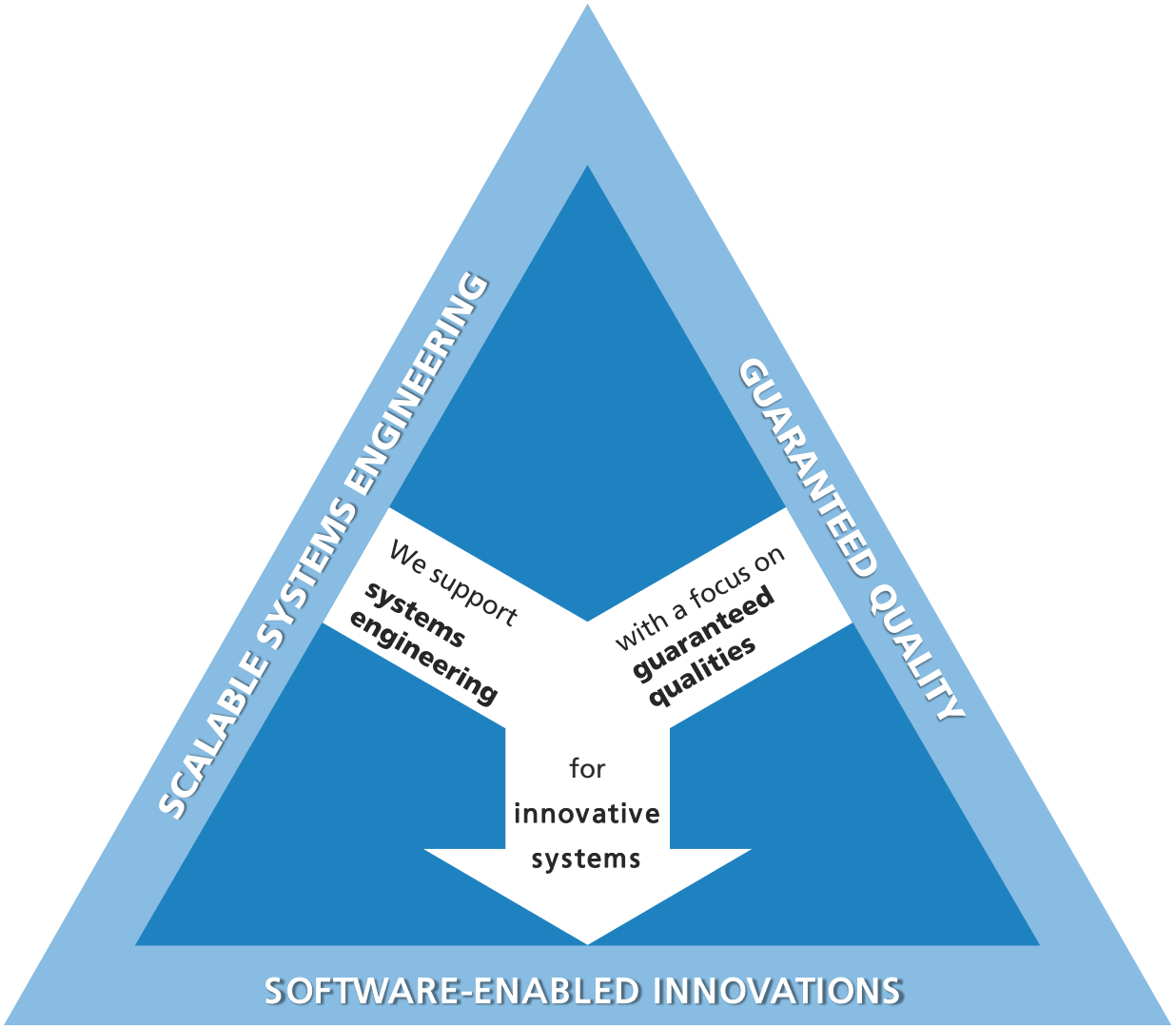
The institute for software and systems engineering methods

- Founded in 1996, headquartered in Kaiserslautern
- Over 155 full-time equivalents (FTEs)
- Our solutions can be scaled flexibly and are suitable for companies of any size

- Our most important business areas:
  - Automotive and Transportation Systems
  - Automation and Plant Engineering
  - Health Care
  - Information Systems
  - Energy Management
  - E-Government

© Fraunhofer IESE

Our Competencies – for Your Benefit



© Fraunhofer IESE

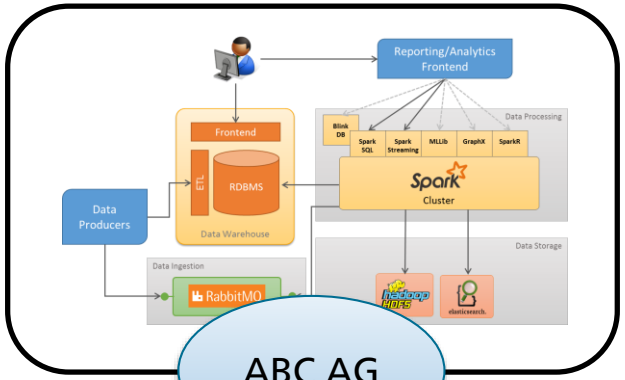
Fraunhofer IESE

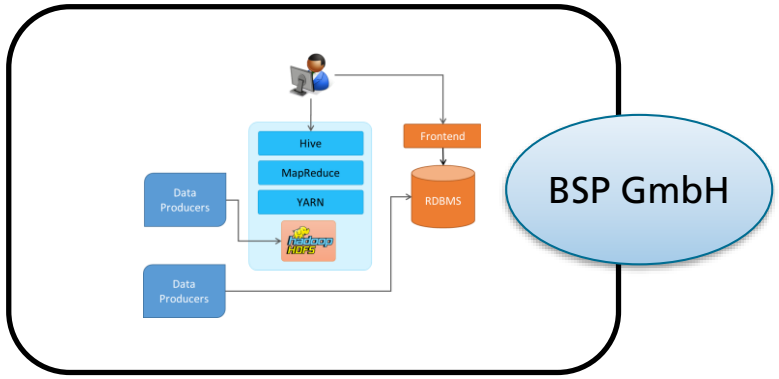
Our Competencies – for Your Benefit

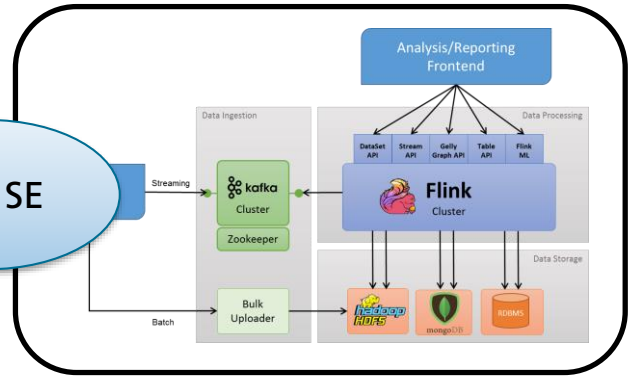
© Fraunhofer IESE

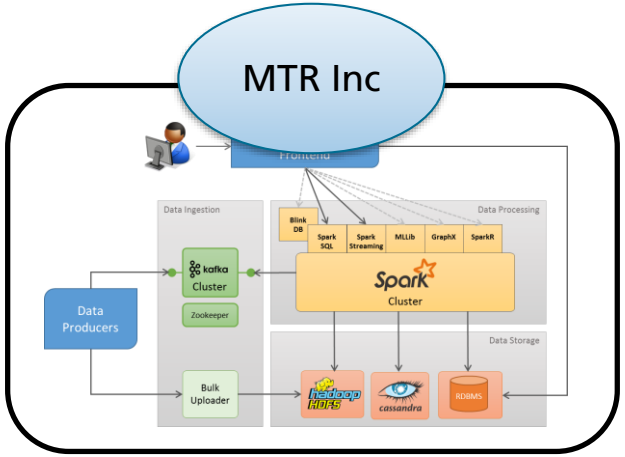
Fraunhofer IESE

Motivation



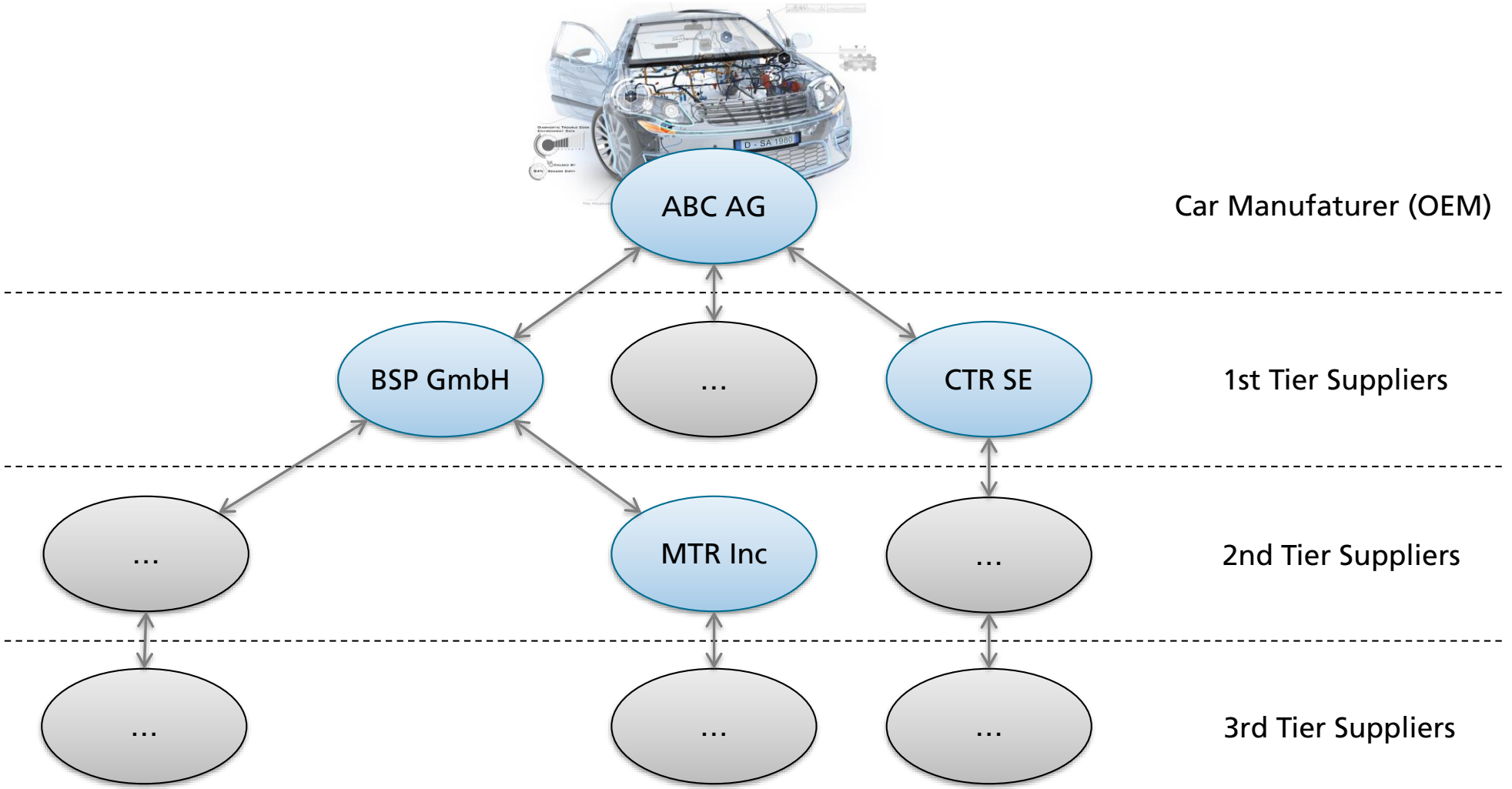






© Fraunhofer IESE

Motivation



© Fraunhofer IESE

The research project PRO-OPT

PRO-OPT aims at identifying valuable data and making it available for creating additional benefit for all members of a Smart Ecosystem.

Smart Data

Gefördert durch:

Bundesministerium für Wirtschaft und Energie

aufgrund eines Beschlusses des Deutschen Bundestages

DSA

Project Lead, Technology and Evaluation Partner for Automotive Diagnostics

Fraunhofer IESE

Technical Project Lead, Research Partner for Access Restriction, SW Architecture & Data Quality

camLine

agile • efficient • reliable

Technology and Evaluation Partner for Production Systems

Audi

Evaluation Partner and Data Supplier for Automotive Diagnostics

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH

Research Partner for Data Mining and Integration of System Components

Partners

Continental	Data Supplier and Evaluation Partner
smartFactory	Multiplicator
Technische Universität Kaiserslautern	Visualization

Project Duration: 01.01.2015 - 31.12.2017

© Fraunhofer IESE

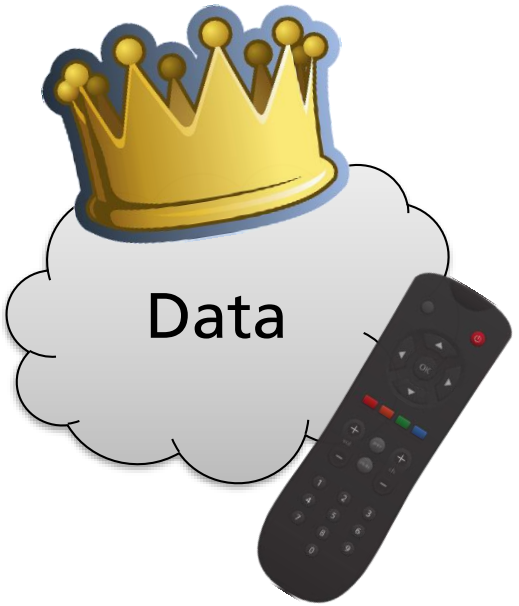
Main Challenge 1: Build a Technology Bridge

How can we bridge the **isolated data islands** of the partners in the ecosystem so that **analysis can be performed across these islands** while the **original owner of the data keeps control** of it?

Main Challenge 2: Data Usage Control

- Data is the DNA of a company
- There are common, but also conflicting interests in the ecosystem
- Data access is often perceived as a binary decision

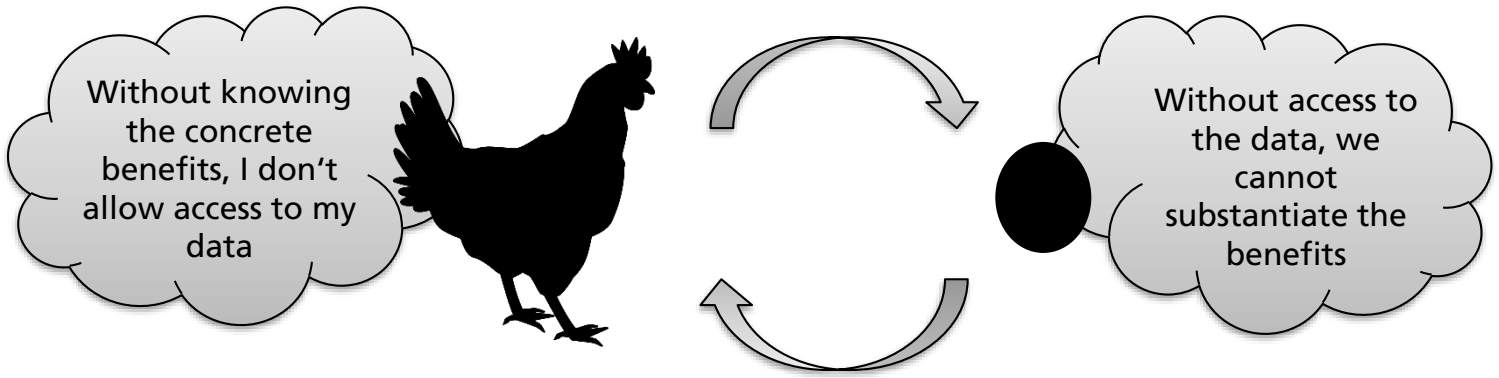
→ Fine granular access policies with additional control and protection mechanisms are a key success factor



© Fraunhofer IESE

Main Challenge 3: Substantiation of Benefits

- High-Level: Everyone agrees
- Concrete use cases are sometimes difficult to define
- But they are needed to justify investments & compromises

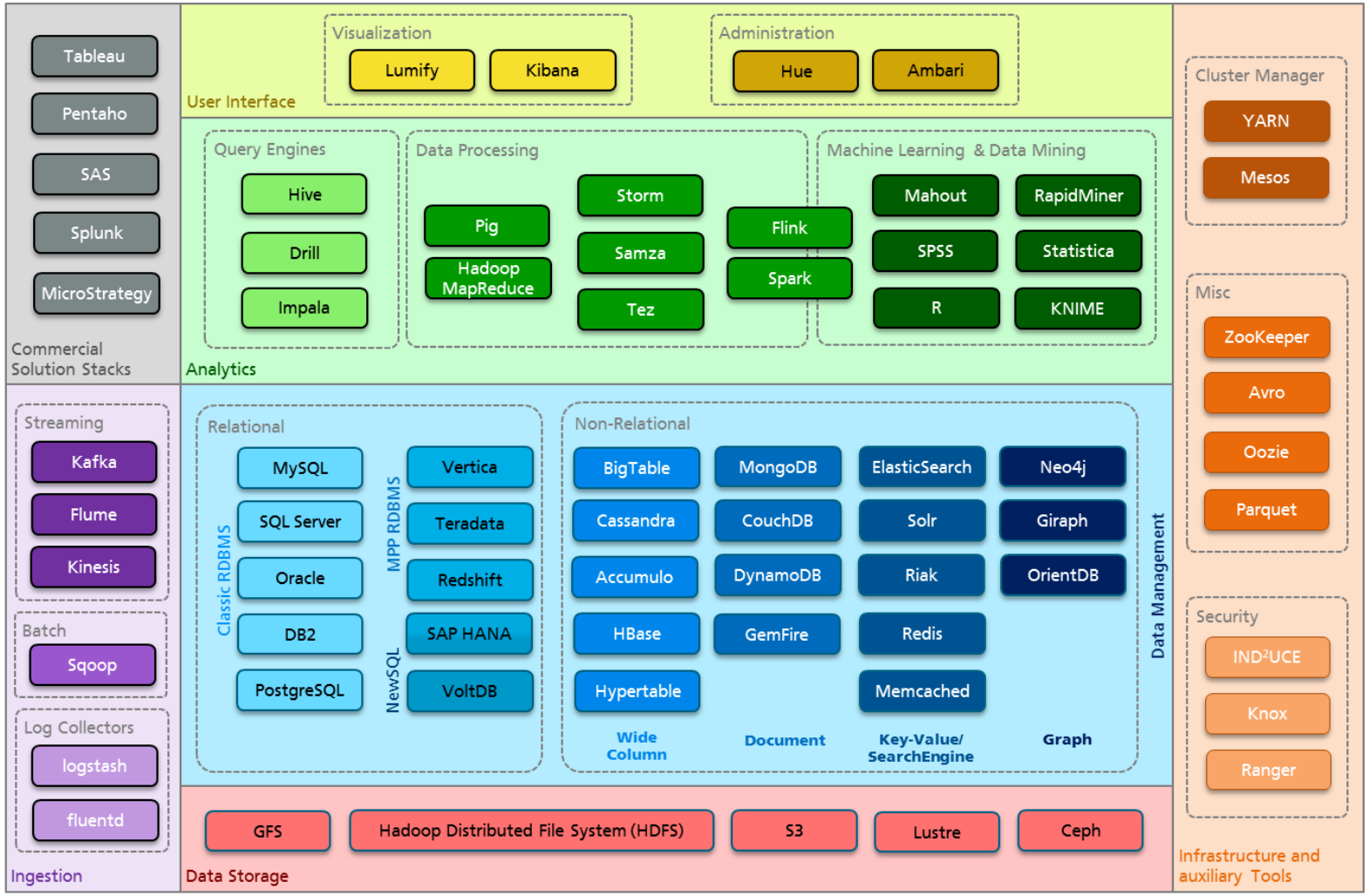


- A research project is well suited to resolve this deadlock

© Fraunhofer IESE

Step 0: Gathering an holistic overview

PRO-OPT Big Data Landscape



© Fraunhofer IESE

Fraunhofer IESE

Step 1:

Identifying potential pillars for the technology brigde

Observation

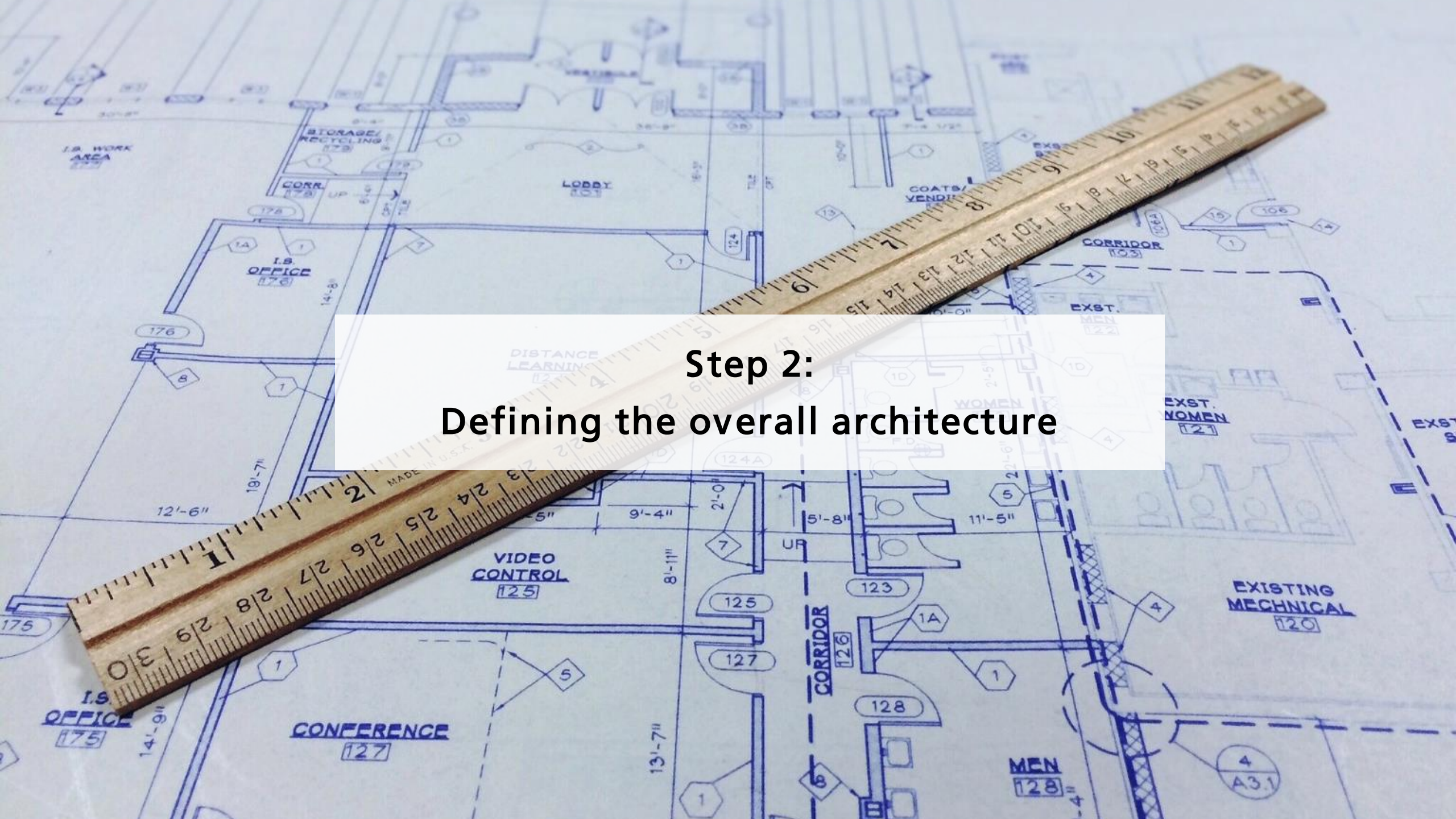
- Rise of multi-purpose data processing engines:

© Fraunhofer IESE

Spark vs Flink – High-Level Comparison

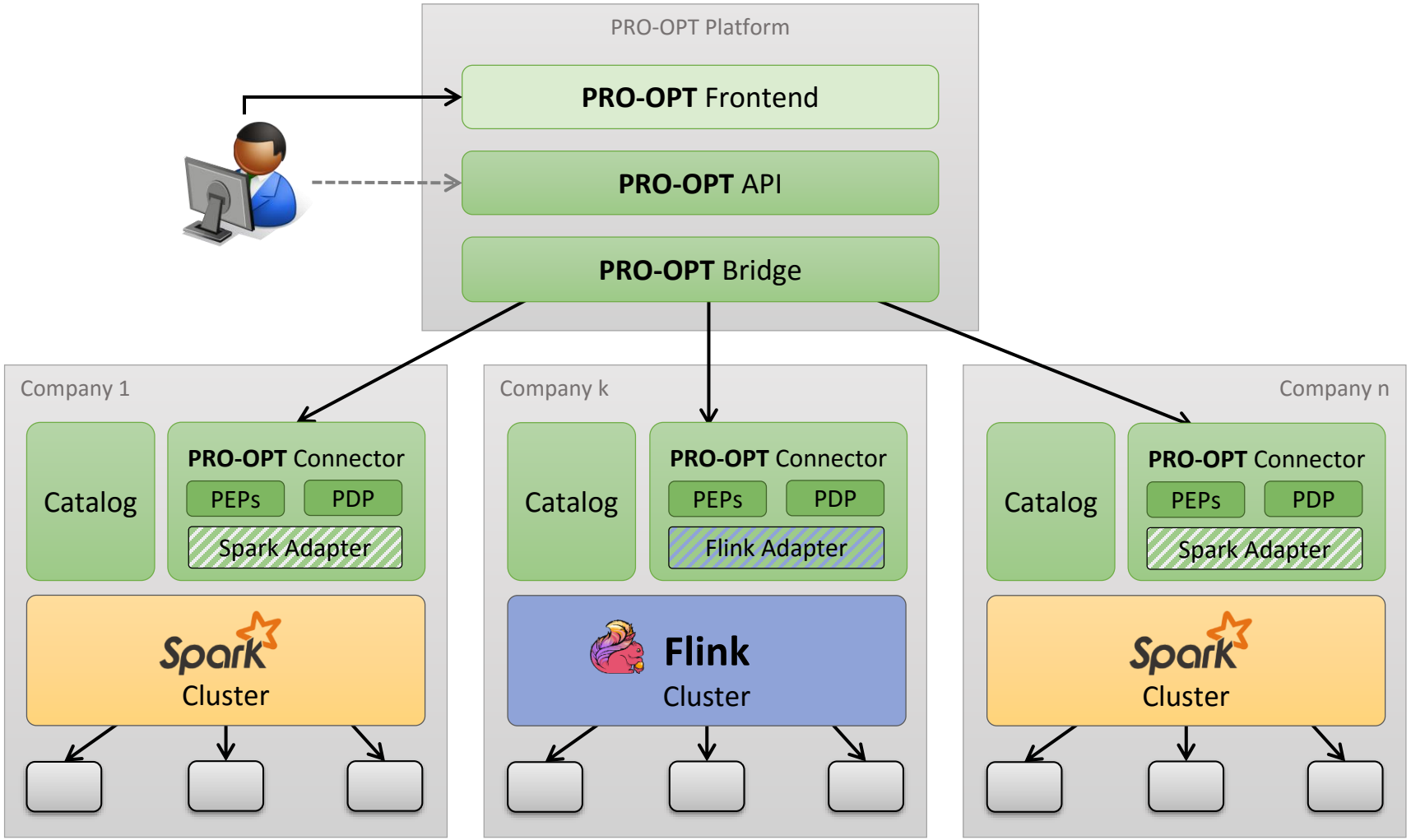
	Flink	Spark
Origin	TU Berlin	University of California, Berkeley
Execution Model	Directed Acyclic Graphs (DAGs)	Directed Acyclic Graphs (DAGs)
Streaming Support	Native	Micro-Batches
Latest Stable Release	1.0.3 (11.05.2016)	1.6.1 (09.03.2016)
Spreading	●●●	●●●●●●
Contributors	~ 200	~ 900
Sprachen	Scala, Java	Scala, Python, (Java)
Project Homepage	http://flink.apache.org	http://spark.apache.org
Backing Company	http://data-artisans.com	https://databricks.com

© Fraunhofer IESE



Step 2:
Defining the overall architecture

High-Level PRO-OPT Platform Architecture



© Fraunhofer IESE

PRO-OPT Components

PRO-OPT Frontend

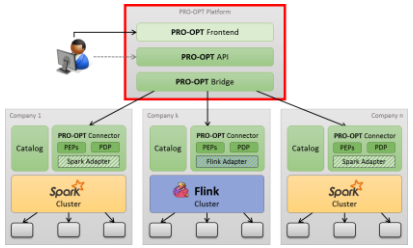
- UI for general analysis and reporting features
- Also used for catalog and platform management

PRO-OPT API

- Interface for defining PRO-OPT programs
- Inspired by the Spark and Flink API – but detailed structure is work in progress

PRO-OPT Bridge

- Acts as some kind of pre-compiler and dispatcher
- Conveys the PRO-OPT programs into one or multiple Spark and/or Flink programs
- Passes the Spark and/or Flink programs together with the identity of the originator to the respective PRO-OPT connectors
- Collects the results from the connectors and applies some additional processing to generate the final result

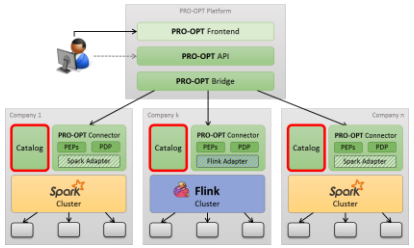


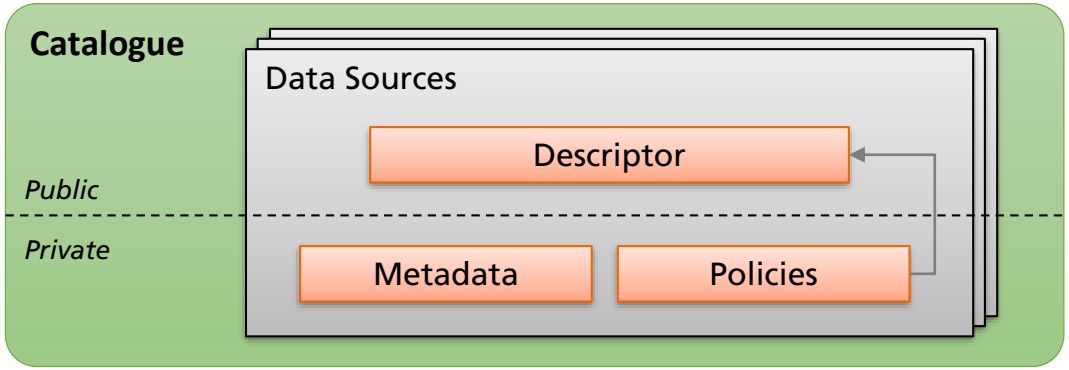
© Fraunhofer IESE

PRO-OPT Components

Catalogue

- Contains a list & description of all data sources of this ecosystem member
- Private: Policy for each data source with fine granular usage rules:
  - Who is allowed to use the data?
  - How often can it be used?
  - Wich parts can be used?
  - What parts can be actually returned?
  - Do additional measures have to be applied (e.g. pseudonymisation)?
  - ...





© Fraunhofer IESE

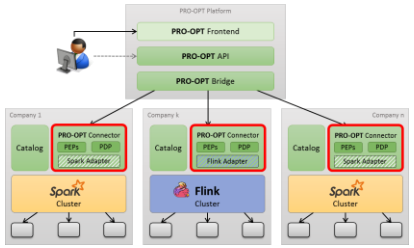
PRO-OPT Components

PRO-OPT Connector

- Receives Spark and/or Flink programs (dependent on the cluster that is installed at this particular ecosystem member)
- Performs a pre-processing step (e.g. replacing data source ids with actual addresses)
- Enforces the data usage control rules defined through the repective policies by applying Policy Enforcement Points (PEP) and Policy Decision Points (PDP)
- Data usage control is based on the Fraunhofer IESE IND²UCE Framework

IND²UCE SECURITY

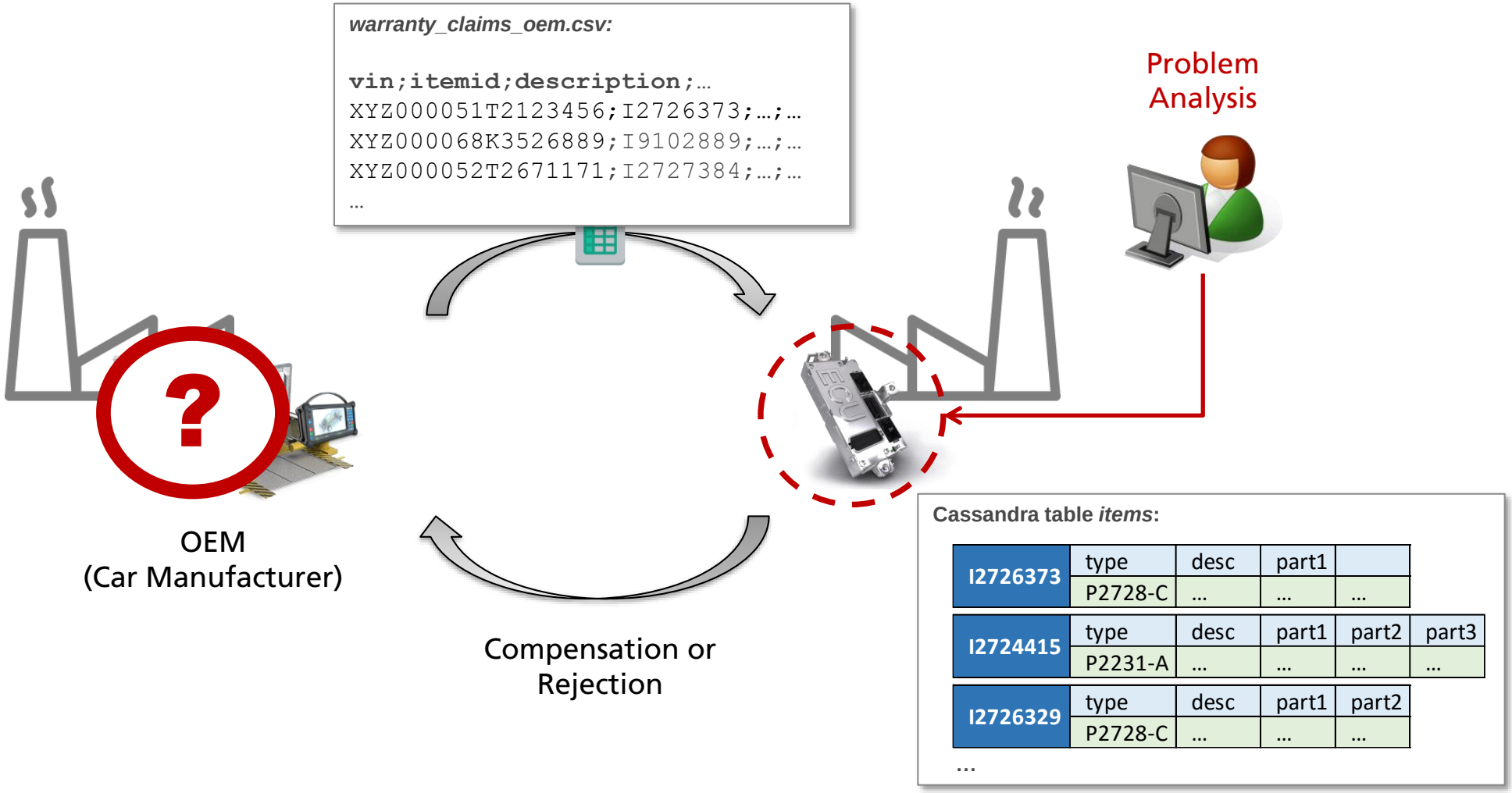
Integrated Distributed Data Usage Control Enforcement



© Fraunhofer IESE

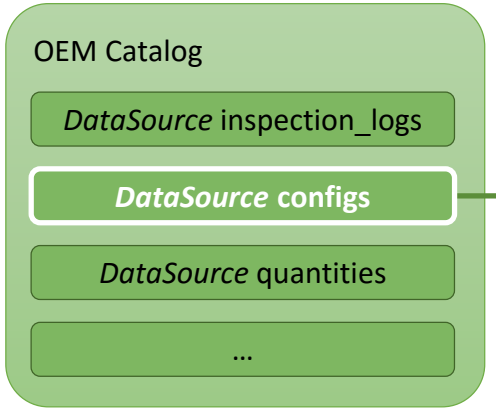
Fraunhofer IESE

Sample Scenario: Warranty Claim Process



© Fraunhofer IESE

OEM Catalog

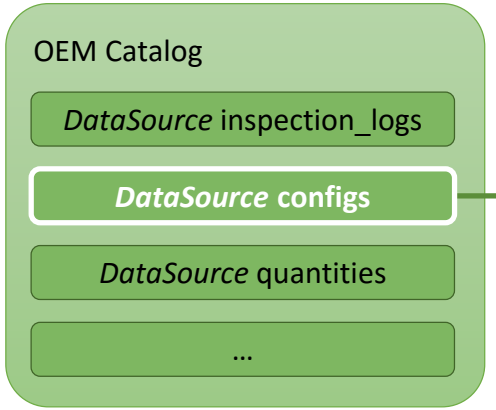



```
{
   "id": 126,
   "name":"configs",
   "description":"Information on car orders",
   "type":"structured",
   "columns":{
      "vin":{
         "description":"Vehicle Identification Number",
         "type":"String"
      },
      "dealer":{
         "description":"Id of the dealer",
         "type":"integer"
      },
      "enginetype":{
         "description":"Type of the engine (diesel, petrol, etc)",
         "type":"String"
      }, […]
   },
   "restrictions":{ […] }
}
```


© Fraunhofer IESE

Fraunhofer IESE

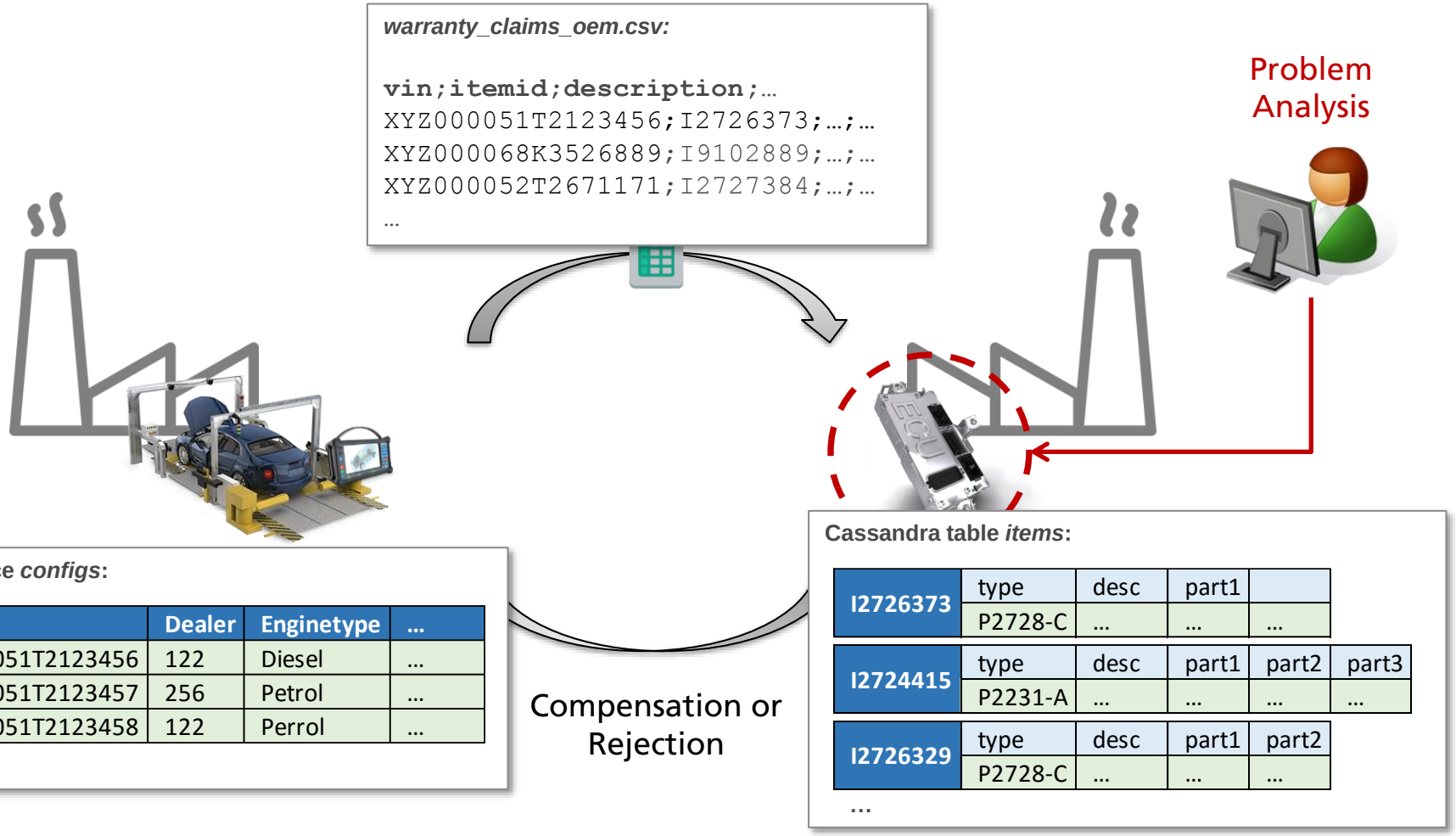
OEM Catalog




```
{
   "id": 126,
   […]
   "restrictions":{
      "limitrows": {
         "max": 100
      },
      "frequency": {
         "maxperday": 1,
         "maxpermonth": 10
      },
      „pseudonymize": {
         "column" : "dealer"
      }
   }
}
```

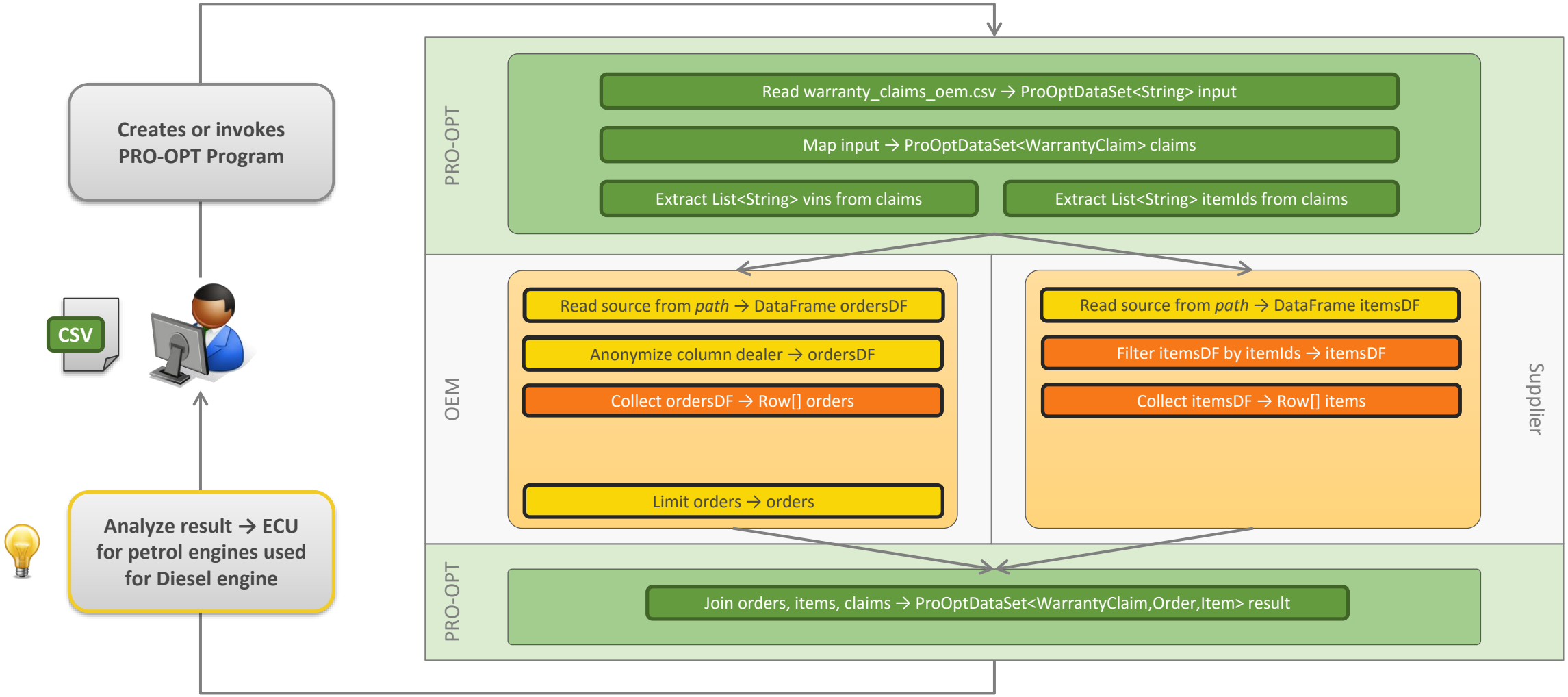
© Fraunhofer IESE

Sample Scenario: Warranty Claim Process



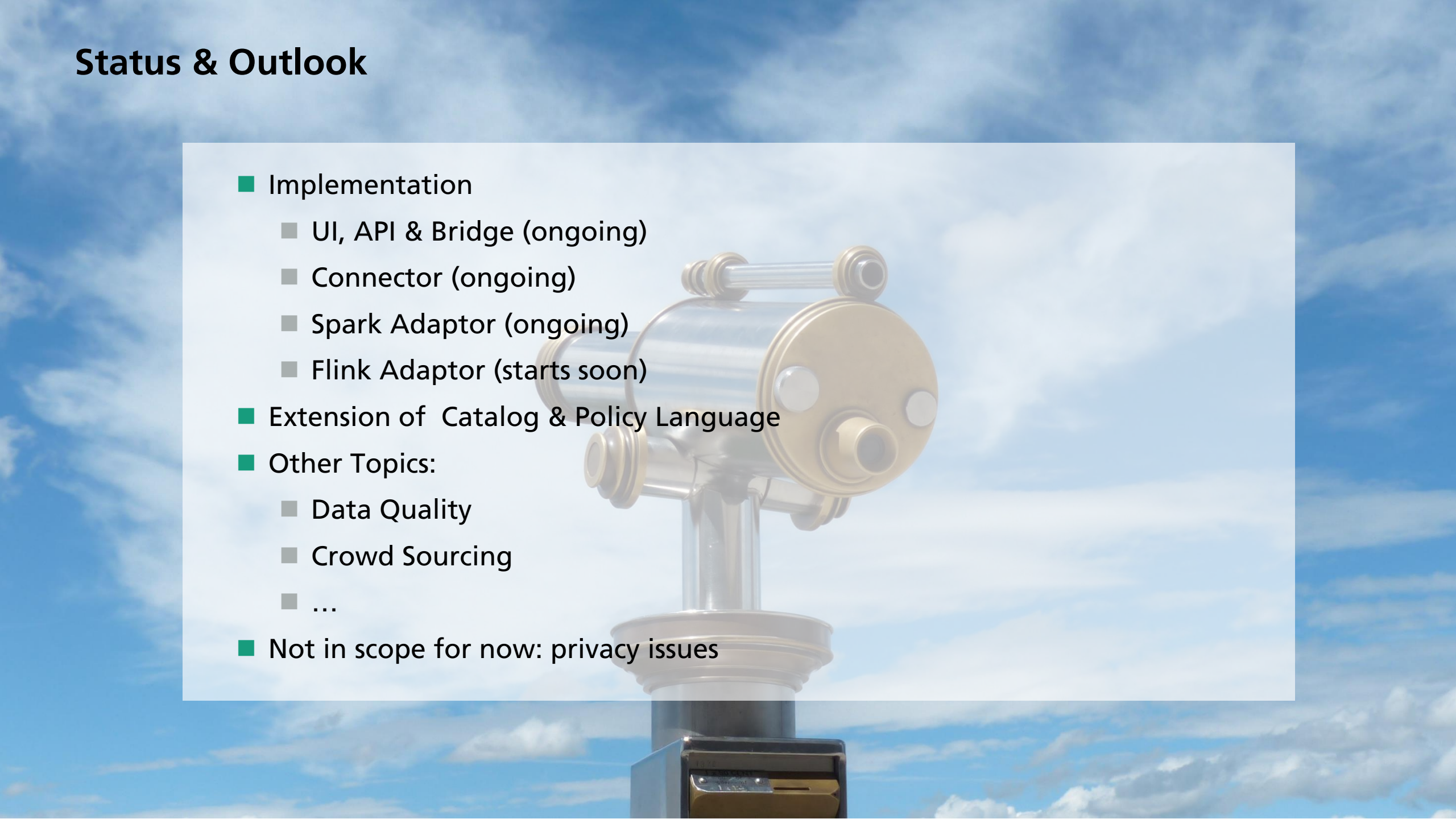
© Fraunhofer IESE

Sample Program



© Fraunhofer IESE

Status & Outlook



- Implementation
  - UI, API & Bridge (ongoing)
  - Connector (ongoing)
  - Spark Adaptor (ongoing)
  - Flink Adaptor (starts soon)
- Extension of Catalog & Policy Language
- Other Topics:
  - Data Quality
  - Crowd Sourcing
  - …
- Not in scope for now: privacy issues

Contact

Torsten Lenhart

torsten.lenhart@iese.fraunhofer.de

Fraunhofer IESE

Fraunhofer-Platz 1

67663 Kaiserslautern

Germany

Phone +49 631/6800-0

www.iese.fraunhofer.de

© Fraunhofer IESE

Image Credits

Slide	Description	Author	Link/Information	License
1	Data World Map	Geralt	https://pixabay.com/de/bin%C3%A4r-eins-null-kontinente-erde-1414315/	CC0 1.0
7	Car	DSA GmbH	© DSA GmbH – used with kind approval of DSA GmbH	Proprietary
9	Golden Gate	Unsplash	https://pixabay.com/de/golden-gate-br%C3%BCcke-san-francisco-388917/	CC0 1.0
10	Remote control	JJuni	https://pixabay.com/en/remote-control-one-trillion-kinds-1143461/	CC0 1.0
10	Crown	OpenClipartVectors	https://pixabay.com/de/krone-golden-gelb-kaiser-zubeh%C3%B6r-576226/	CC0 1.0
11	Chicken silhouette	ClkerFreeVectorImages	https://pixabay.com/de/hen-huhn-gefl%C3%BCgel-bauernhof-tier-311285/	CC0 1.0
12	Earth from space	Skeeze	https://pixabay.com/en/panorama-earth-canada-landscape-1241289/	CC0 1.0
14	Under the bridge	Unsplash	https://pixabay.com/en/bridge-river-under-cityscape-1149241/	CC0 1.0
17	Blueprint	Wokandapix	https://pixabay.com/en/blueprint-ruler-architecture-964630/	CC0 1.0
22	Sheet icon	Paomedia	http://www.iconarchive.com/show/small-n-flat-icons-by-paomedia.html	PD
22	Factory	IconsMind	https://www.iconsmind.com/	Custom
22	ECU	DSA GmbH	© DSA GmbH – used with kind approval of DSA GmbH	Proprietary
22	Test bed	DSA GmbH	© DSA GmbH – used with kind approval of DSA GmbH	Proprietary
26	Bulb	IconLeak	http://iconleak.com http://www.iconarchive.com/show/or-icons-by-iconleak/light-bulb-icon.html	Custom
27	Fernrohr	Hans	https://pixabay.com/de/fernrohr-durchblick-aussicht-blick-122960/	CC0 1.0

© Fraunhofer IESE