Datum, Ort, Anlass

KI Data Tooling Projektpräsentation

Erika Muster, BMW AG, Max Mustermann, Bosch

1

Motivation und Zielsetzung

Motivation

- Methoden der Künstlichen Intelligenz (KI) sind der Schlüsselfaktor für die Entwicklung und Umsetzung von hochautomatisierten Fahrfunktionen.
- Bevor KI-Funktionen in Fahrzeugen implementiert werden können, müssen sie umfassend trainiert und validiert werden.

Für das Training und die Validierung von KI-Funktionen ist eine hochwertige und umfassende Datenbasis erforderlich.

Ausgangslage

Es existiert aktuell keine umfassende Datenbasis von ausreichender Qualität für das Training und die Validierung von KI-Funktionen für das hochautomatisierte Fahren.

- Bisher sind lediglich Insellösungen für einzelne Datentypen, Sensortypen, eingeschränkte Szenarien oder Teile der Trainings- und Validierungskette verfügbar.
- Bisher wurden keine integrierten Betrachtungen der Stärken und Schwächen verschiedener Datengenerierungsmethoden vorgenommen.

Ausgangslage: Datentypen und ihre Eigenschaften

	+	−
Reale Daten	• Exakte Abbildung der Realität mit maximaler Sensorqualität • Direkte Abbildung von Sonderfällen wie Regen, Schnee, etc., die teils schwierig zu modellieren sind	• Erfassung und Aufbereitung zeit- und kostenintensiv • Labeling derzeit manuell mit hohen Kosten und Aufwand verbunden • Verkehrsvielfalt kaum vollständig abbildbar, insbesondere kritische Situationen, z.B. Unfälle • Rechtliche Einschränkungen, z.B. DSGVO-Regelungen
Synthetische Daten	• Systematische Generierung von vielfältigen Verkehrsszenarien • Gefahrlose Modellierung kritischer Situationen • Automatisierte Erzeugung von Ground-Truth • Keine Einschränkungen durch DSGVO	• Eingeschränkte Realitätsnähe • Eingeschränkte Übertragbarkeit trainierter KI-Modelle auf die Realität • Synthese von Radardaten bisher nicht implementiert • Kosten für Modellierung, Weiterentwicklung von Toolchains und Rechenleistung
Augmentierte Daten	• Kombination realer Daten mit der strukturierten Erzeugung synthetischer Daten	• Unzureichend erforscht

Zielsetzung des Projektes

ZIELGERICHTETE DATENBASIS

METHODEN UND TOOLS

OPTIMIERTE KI-TRAININGSSTRATEGIE FÜR AUTONOME FAHRZEUGE

KI Data Tooling entwickelt die Datenkomplettlösung für das Training und die Validierung von KI-basierten Funktionen für das hochautomatisierte Fahren.

Zielsetzung des Projektes

ZIELGERICHTETE DATENBASIS

Kamera, Lidar, Radar

Realdaten

Synthetische Daten

Augmentierte Daten

Standards zur Beschreibung von Sensordaten

METHODEN UND TOOLS

Aufbereitung

Augmentierung

Datenkompression

Automatisiertes Realdatenlabeling

Erzeugung und Validierung synthetischer Daten

Corner Cases: Detektion und Synthese

Kontextidentifizierung und -nutzbarmachung

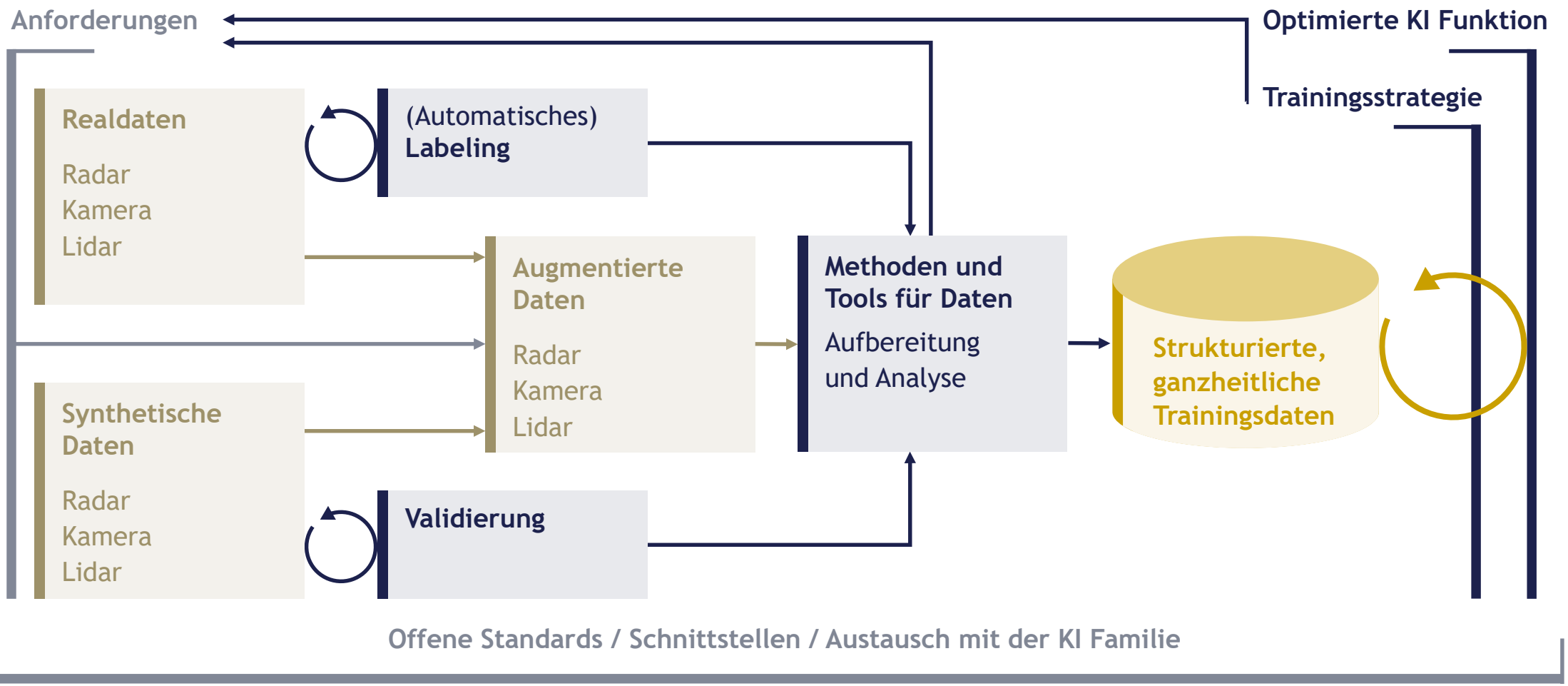
OPTIMIERTE KI-TRAININGSSTRATEGIE FÜR AUTONOME FAHRZEUGE

Vorgehensmodell zur Datenerzeugung und Methoden, um die Entwicklung von KI-Funktionen zu automatisieren und beschleunigen

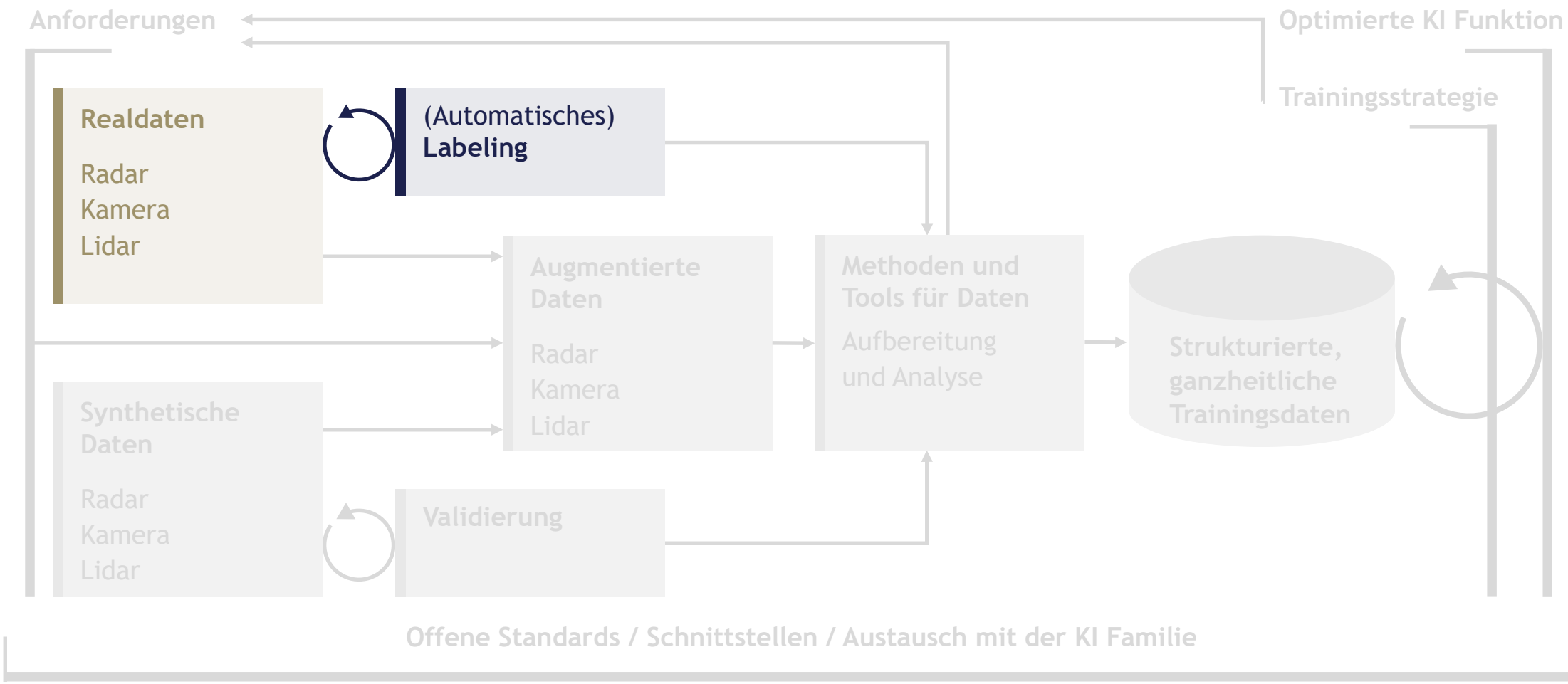
2

Methodischer Ansatz

Projektkonzept



Realdaten


Anforderungen
Optimierte KI Funktion
Trainingsstrategie
Realdaten
Radar
Kamera
Lidar
(Automatisches)
Labeling
Augmentierte Daten
Radar
Kamera
Lidar
Methoden und Tools für Daten
Aufbereitung und Analyse
Strukturierte, ganzheitliche Trainingsdaten
Synthetische Daten
Radar
Kamera
Lidar
Validierung
Offene Standards / Schnittstellen / Austausch mit der KI Familie

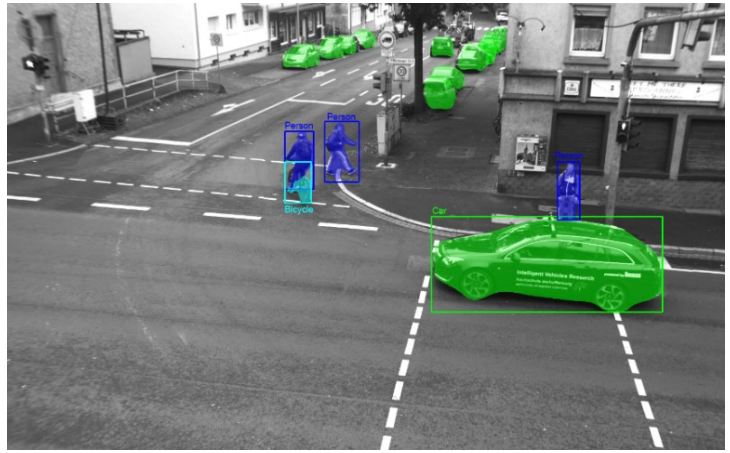
Realdaten im Projekt: Datenbasis

Im Projekt wird eine umfangreiche Realdatenbasis aus Kamera-, Lidar- und Radardaten geschaffen.

- Mehrere Versuchsträger nehmen deutschlandweit Daten diverser Szenarien auf.

- Zusätzlich werden Realdaten an den Forschungskreuzungen in Braunschweig und Aschaffenburg erhoben, um die Validierung synthetischer Daten zu unterstützen.

Versuchsträger des Projekts



Forschungskreuzung in Aschaffenburg

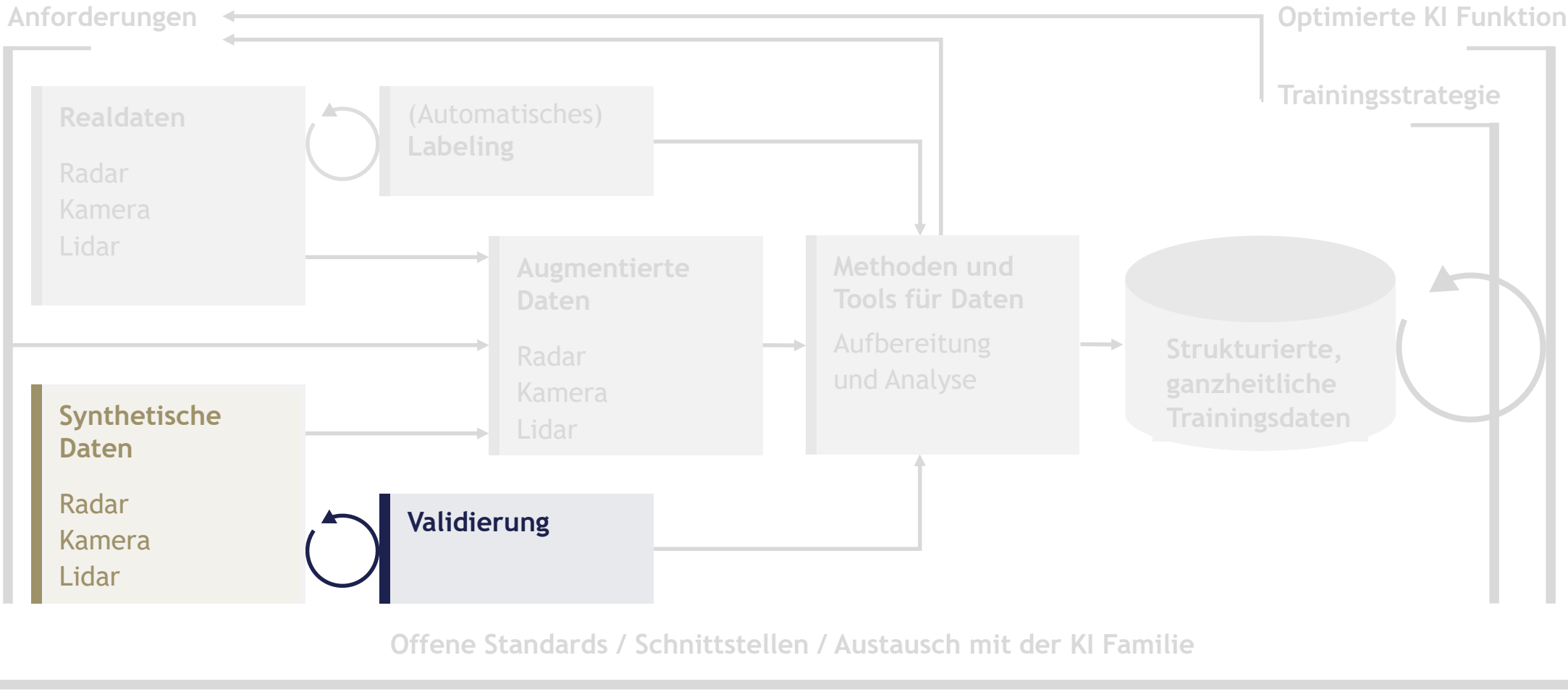
Realdaten im Projekt: Methoden und Tools

Anhand der erhobenen Realdaten werden Methoden und Tools entwickelt für:

- Die effiziente Aufbereitung und Veredelung der Realdaten.
- Das teil- und vollautomatisierte Labeling der Realdaten.
- Die Erfassung und Nutzbarmachung von Kontextinformationen für die Strukturierung, Analyse und Bearbeitung der Realdaten.

Beispiele für Realdaten: Lidar- und Kameraaufnahmen eines Versuchsträgers

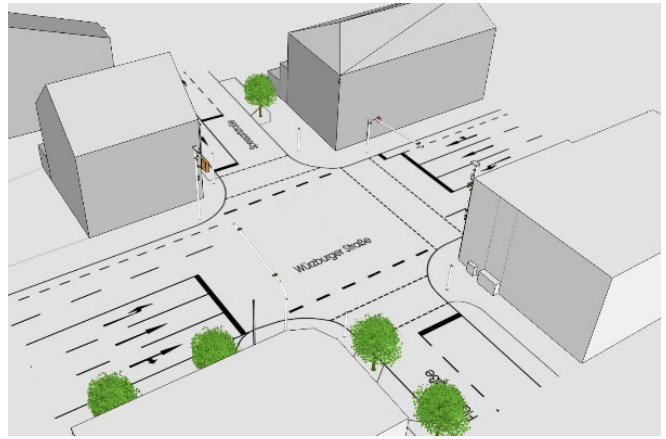
Synthetische Daten



Synthetische Daten im Projekt: Datenbasis

Im Projekt werden synthetische Kamera-, Lidar- und Radardaten generiert.

- Die Grundlage für die Datengenerierung sind die synthetischen Abbilde der beiden Forschungskreuzungen und ein im Projekt erarbeiteter Szenenkatalog.
- Für den Generierungsprozess wird im Projekt eine Toolchain-Architektur entwickelt und implementiert.

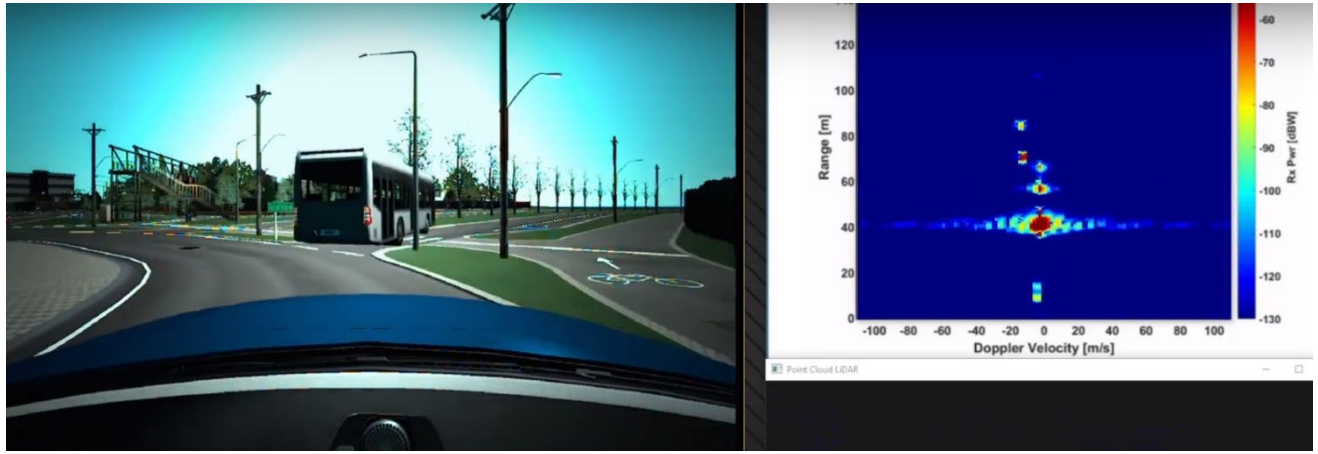


Digitales Modell der Forschungskreuzung Aschaffenburg

Synthetische Daten im Projekt: Methoden und Tools

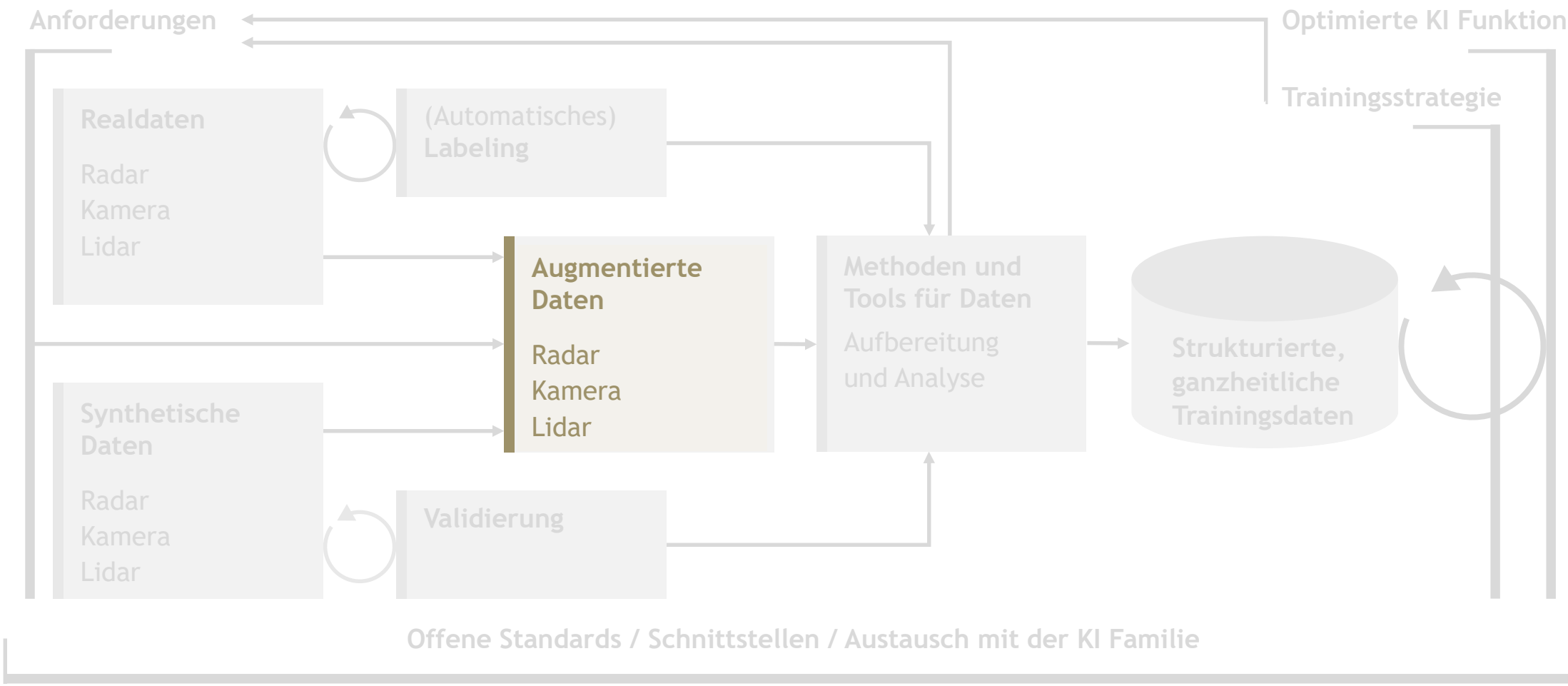
Im Kontext der synthetischen Datengenerierung werden folgende Methoden untersucht:

- Metriken zur Bewertung der Qualität synthetischer Kamera-, Lidar und Radardaten.
- Konzepte zur Validierung synthetischer Kamera-, Lidar und Radardaten.



Beispiele für synthetische Daten: Kamerabilder links, Radardaten rechts

Augmentierte Daten

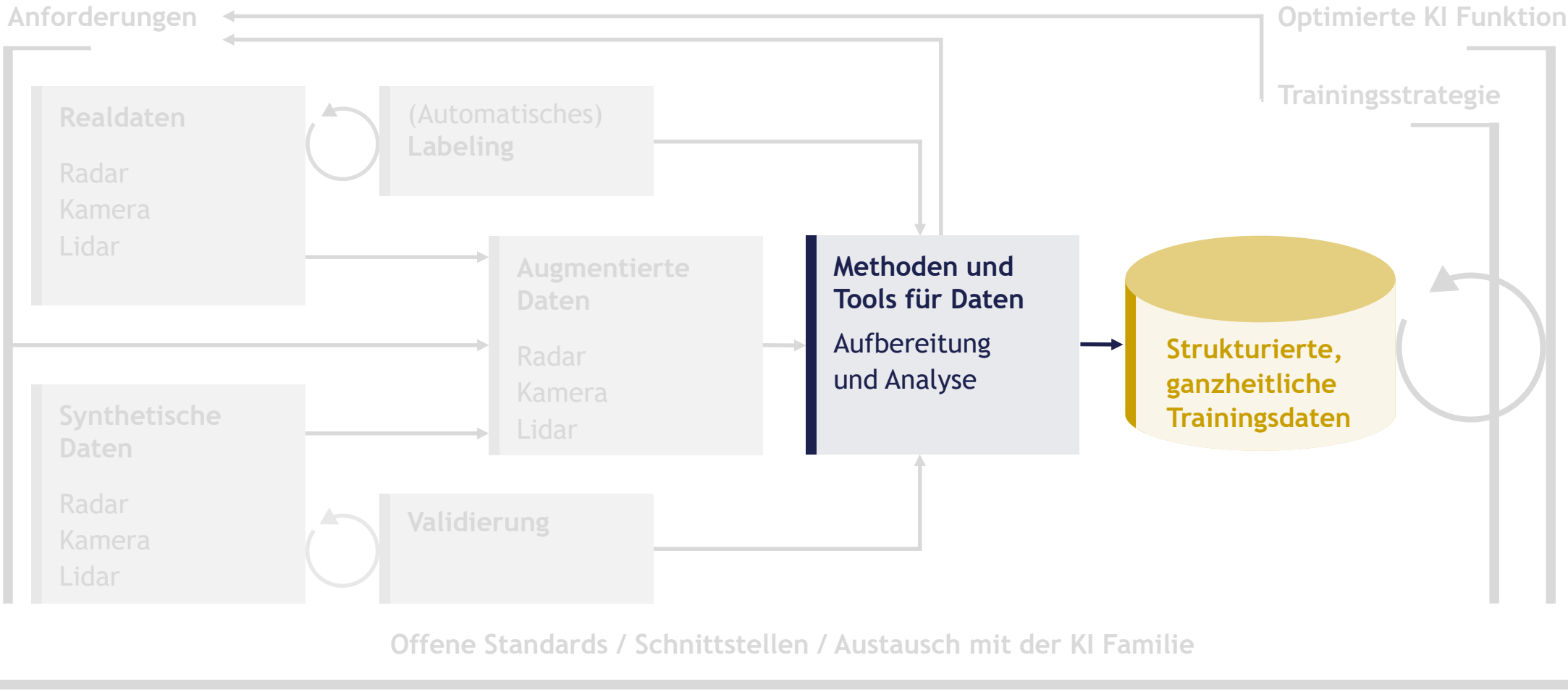


Augmentierte Daten im Projekt

Im Projekt werden augmentierte Kamera-, Lidar- und Radardaten generiert.

- Es werden Augmentierungsverfahren für Kamera-, Lidar- und Radardaten entwickelt.
- Mit Hilfe dieser Verfahren werden Realdaten mit synthetisch erzeugten Objekten augmentiert.
- Es werden Methoden und Metriken zur Bewertung der Güte dieser augmentierten Daten erarbeitet.

Effizienzpotentiale der Datenbereitstellung und -nutzung



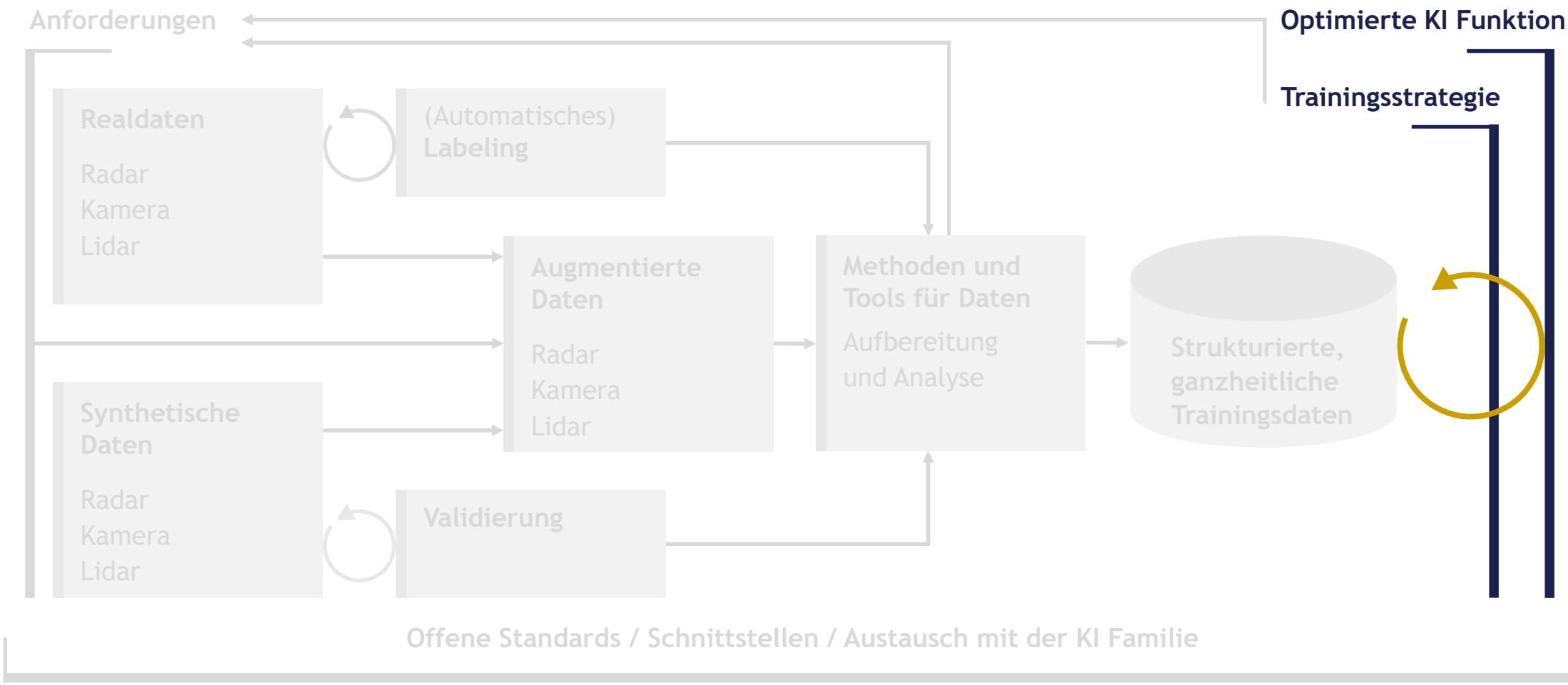
Effizienzpotentiale der Datenbereitstellung und -nutzung

Im Projekt werden verschiedene Aspekte der Effizienz bei der Datennutzung untersucht. Es werden Methoden und Tools entwickelt für:

- Die automatisierte Detektion und Synthetisierung von Corner Cases.
- Die effiziente Kompression und Speicherung von Daten.
- Die Abstraktion der Sensorik.

Kamera- und Lidardaten einer urbanen Straßenszene

Optimierte KI-Trainingsstrategie



Optimierte KI-Trainingsstrategie

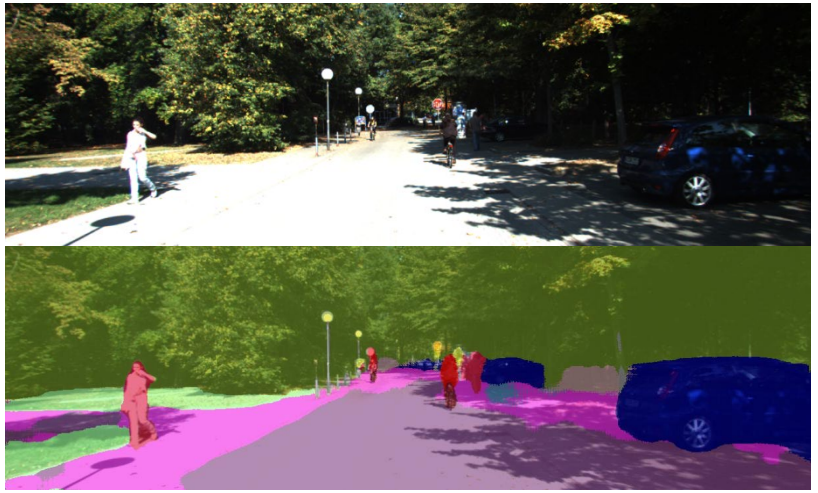
→ Optimal und effizient trainierte KI-Funktionen für das hochautomatisierte Fahren.

- Die im Projekt erschaffene Datenbasis dient als Grundlage für die Entwicklung von KI-Trainingsverfahren.
- Die Potentiale der Nutzung verschiedener Datenkombinationen (real, synthetisch, augmentiert) in verschiedenen Phasen des Trainings und der Validierung werden identifiziert und evaluiert.
- Die Übertragbarkeit des Gelernten der mit synthetischen und/oder augmentierten Daten trainierten Modelle auf die Realität wird überprüft.

Standardisierungsaktivitäten und Ergebnisdemonstration

- Das Projekt leistet Beiträge zu Standardisierungsaktivitäten in Bezug auf Daten und ihre Beschreibungen, Labeling-Spezifikationen, Schnittstellen und Szenenbeschreibungsformate.

- Für die Darstellung der Projektergebnisse werden Demonstratoren aufgebaut, die die Rolle der Daten und deren Funktionen in Fahrzeugen visualisieren.



Kameradaten und ihre semantische Segmentierung

3

Konsortium und Projektstruktur

KI Data Tooling – The Data Kit for Automotive AI

Konsortialleitung: ***BMW AG***

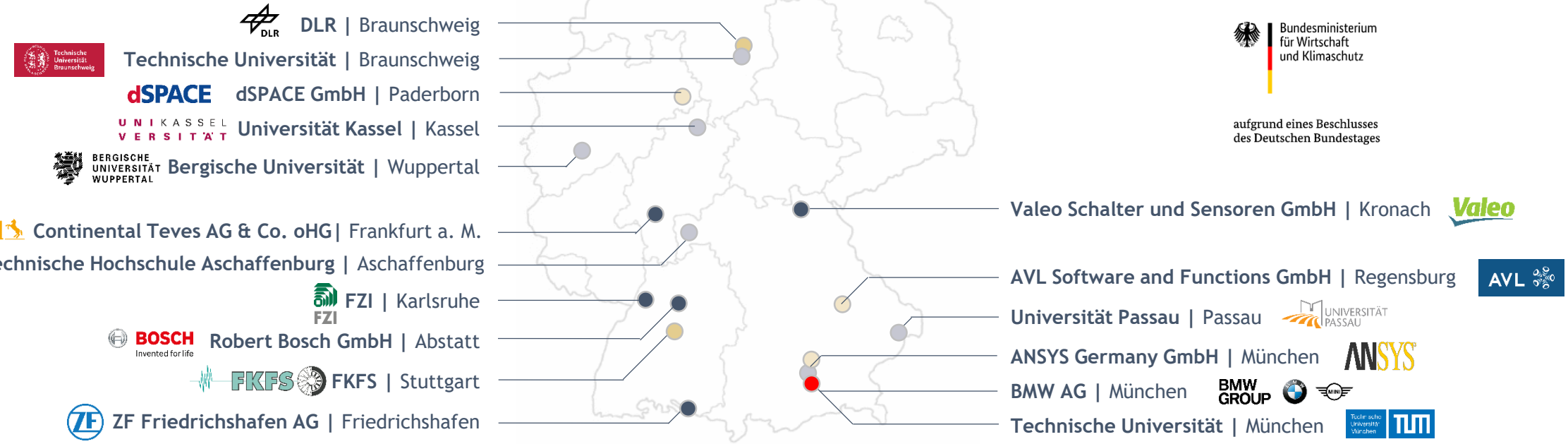
Budget: ***25,7 Mio. €***

Förderung: ***16,2 Mio. €***

Laufzeit: ***45 Monate***

01.04.2020 - 31.12.2023

17 Partner



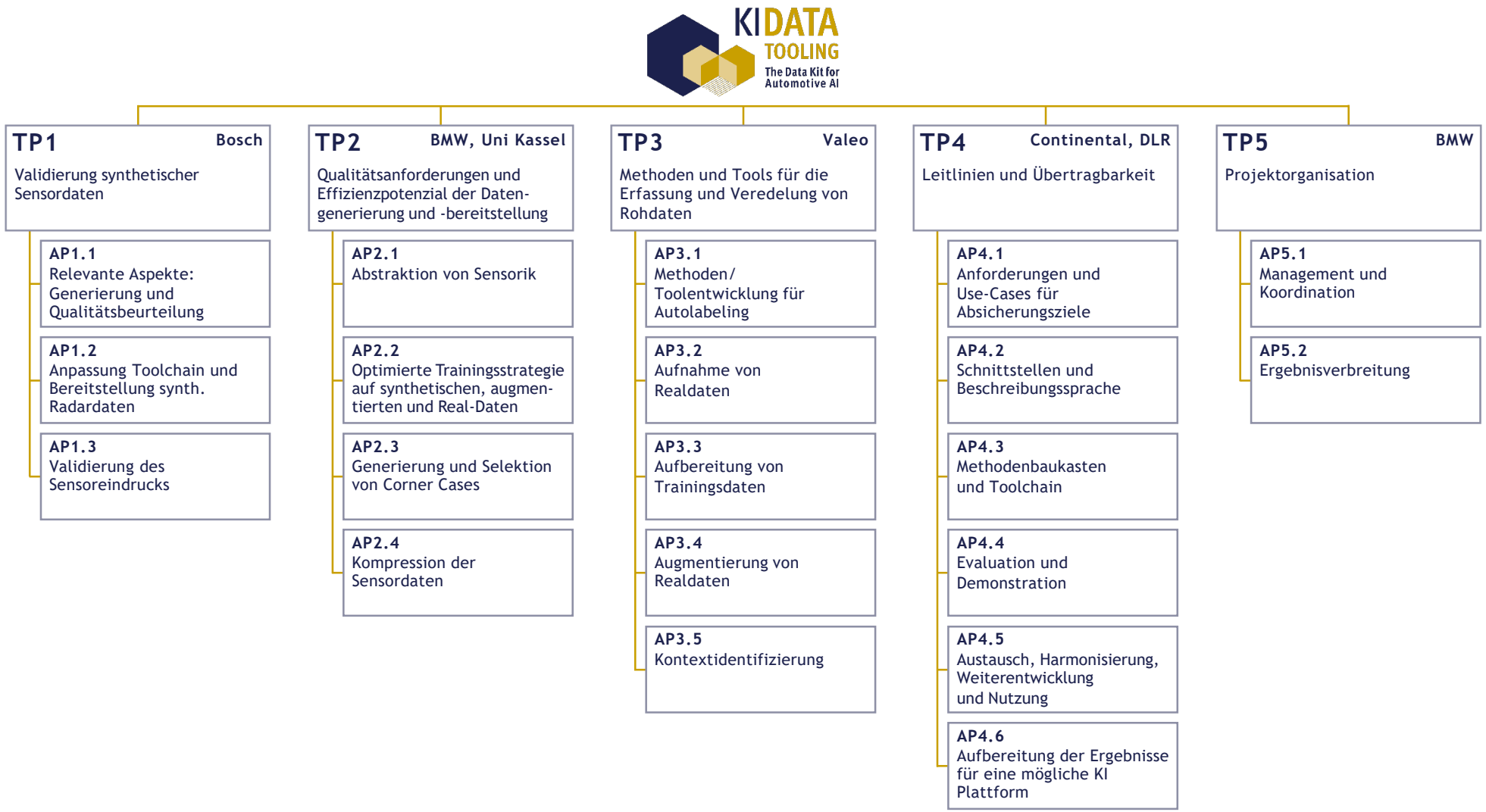
Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

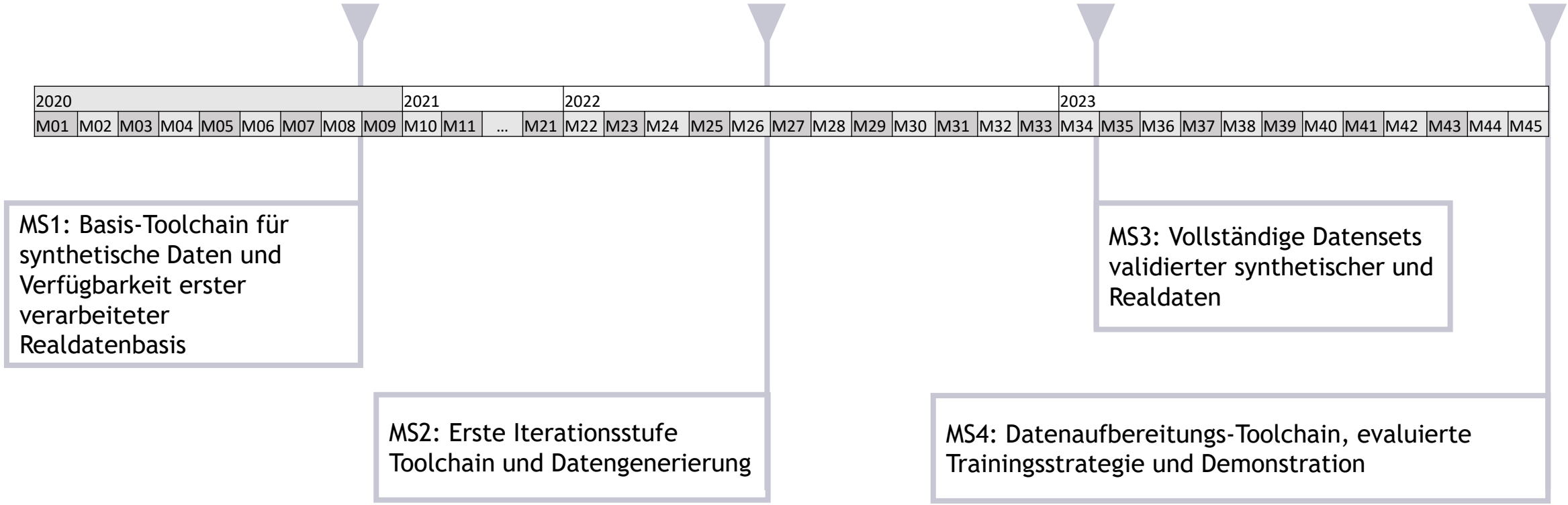
aufgrund eines Beschlusses des Deutschen Bundestages

● OEM ● Zulieferer ● Technologieprovider ● Forschungseinrichtung ● Universität

Projektstruktur



Übersicht über die Meilensteine im Projekt



4

KI Data Tooling als Teil der KI Familie

KI Data Tooling als Teil der KI Familie der VDA LI

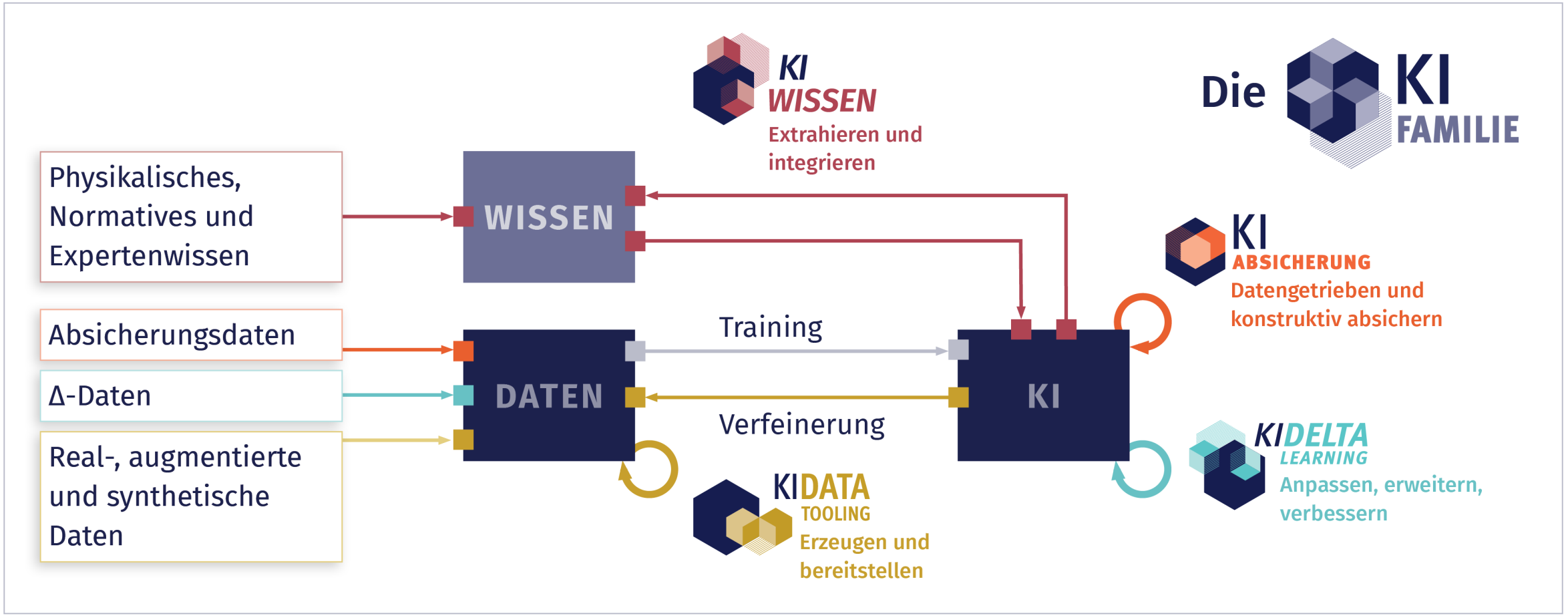
KI WISSEN Entwicklung von Methoden für die Einbindung von Wissen in maschinelles Lernen

KI DELTA LEARNING
Methoden und Werkzeuge zur Erweiterung und Transformation vorhandener KI-Module autonomer Fahrzeuge auf neue Domänen und komplexe Szenarien

KI ABSICHERUNG Methoden und Maßnahmen zur Absicherung von KI-basierten Wahrnehmungs-funktionen für das automatisierte Fahren

KI DATA TOOLING Methoden und Werkzeuge für das Generieren und Veredeln von Trainings-, Validierungs- und Absicherungsdaten für KI-Funktionen autonomer Fahrzeuge

KI Data Tooling im Zusammenhang der KI Familie



5

Ansprechpartner

Ansprechpartner

Projektkoordinator
Hans-Jörg Vögel
BMW Group
ki-dt-koordinator@eict.de

Projektmanagement
EICT GmbH
EUREF Campus Haus 13
Torgauer Straße 12-15
10829 Berlin
ki-dt-projektmanagement@eict.de

KI DATA TOOLING

The Data Kit for Automotive AI

KI Data Tooling ist ein Projekt der KI Familie. Es wurde aus der VDA Leitinitiative autonomes und vernetztes Fahren initiiert und entwickelt und wird vom Bundesministerium für Wirtschaft und Klimaschutz gefördert.

KI FAMILIE

Gefördert durch:

Bundesministerium für Wirtschaft und Klimaschutz

aufgrund eines Beschlusses des Deutschen Bundestages

www.ki-datatooling.de @KI_Familie KI Familie