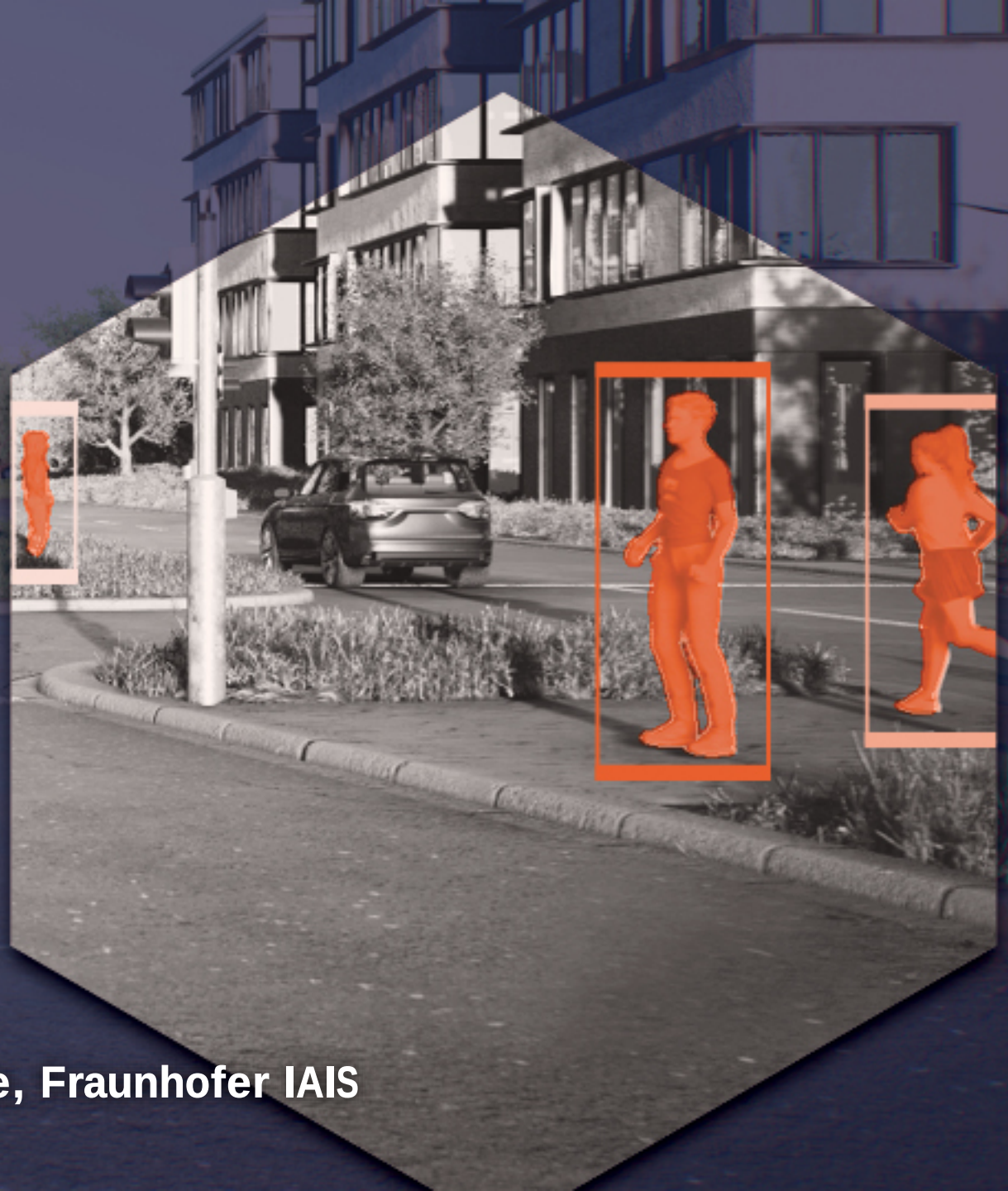
KI ABSICHERUNG

Safe AI for Automated Driving



Datum, Ort, Anlass

KI Absicherung Projektvorstellung

Vorname Name, Volkswagen Vorname Name, Fraunhofer IAIS

KI Absicherung – Safe AI for Automated Driving

Konsortialleitung: Volkswagen AG

Stellv. Konsortialleitung: Fraunhofer IAIS
Wiss. Koordination:

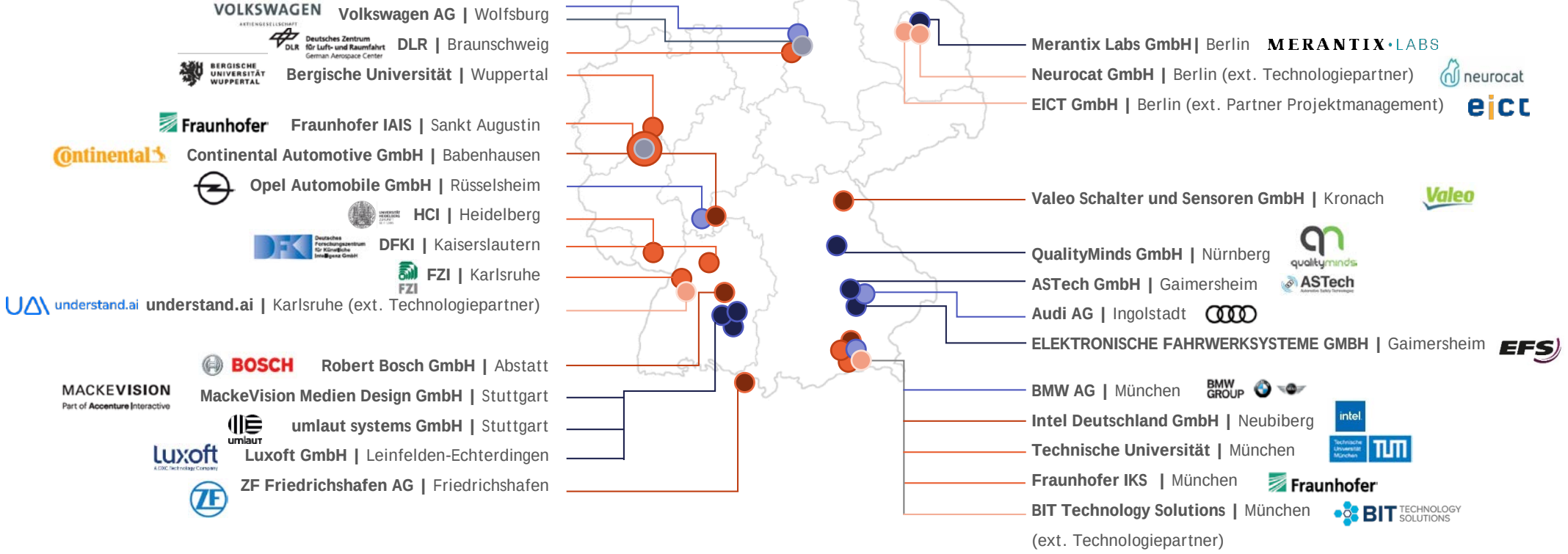
Budget: 41 Mio. €

Förderung: 19,2 Mio. €

Laufzeit: 36 Monate

01.07.2019 – 20.06.2022

24 Partner



Konsortialleitung · OEMs · Zulieferer · Technologieprovider · Forschung · Externe Partner

Gefördert durch:

Bundesministerium für Wirtschaft und Energie

aufgrund eines Beschlusses des Deutschen Bundestages

1

Vision und Ziele des Projekts

Vision

KI ABSICHERUNG
Safe AI for Automated Driving

KI Absicherung macht die Sicherheit KI-basierter Funktionsmodule für das hochautomatisierte Fahren nachweisbar.

Zentrale Ziele in KI Absicherung

1. Trainings- und Testmethoden für KI-basierte Funktionen

KI Absicherung entwickelt und untersucht Methoden und Maßnahmen für die Absicherung KI-basierter Funktionen für das hochautomatisierte Fahren.

2. Absicherungsargumentation

Am Use Case Fußgängererkennung erarbeitet das Projekt eine beispielgebende Argumentations- und Prozesskette zur Absicherung einer komplexen KI-Funktion.

3. Kommunikation mit Standardisierungsgremien zur KI-Zertifizierung

Für die Entwicklung eines Industriekonsenses zur Absicherung von KI-Funktionsmodulen werden die Projektergebnisse in den Dialog mit Standardisierungsgremien eingebracht.

Herausforderung Sicherheitsnachweis KI-basierter Funktionsmodule

Vor KI Absicherung

Künstliche Intelligenz / Maschinelles Lernen / Tiefe neuronale Netze

Unabhängig voneinander agierende Experten

Etablierte Absicherungsprozesse nicht übertragbar.

Funktionale und Gebrauchssicherheit

In KI Absicherung

Künstliche Intelligenz / Maschinelles Lernen / Tiefe neuronale Netze

Kooperation der Industrie- und Wissenschaftsexperten

Entwicklung gemeinsamer Absicherungsargumentation

Funktionale und Gebrauchssicherheit

Die KI Familie und ihre Projekte

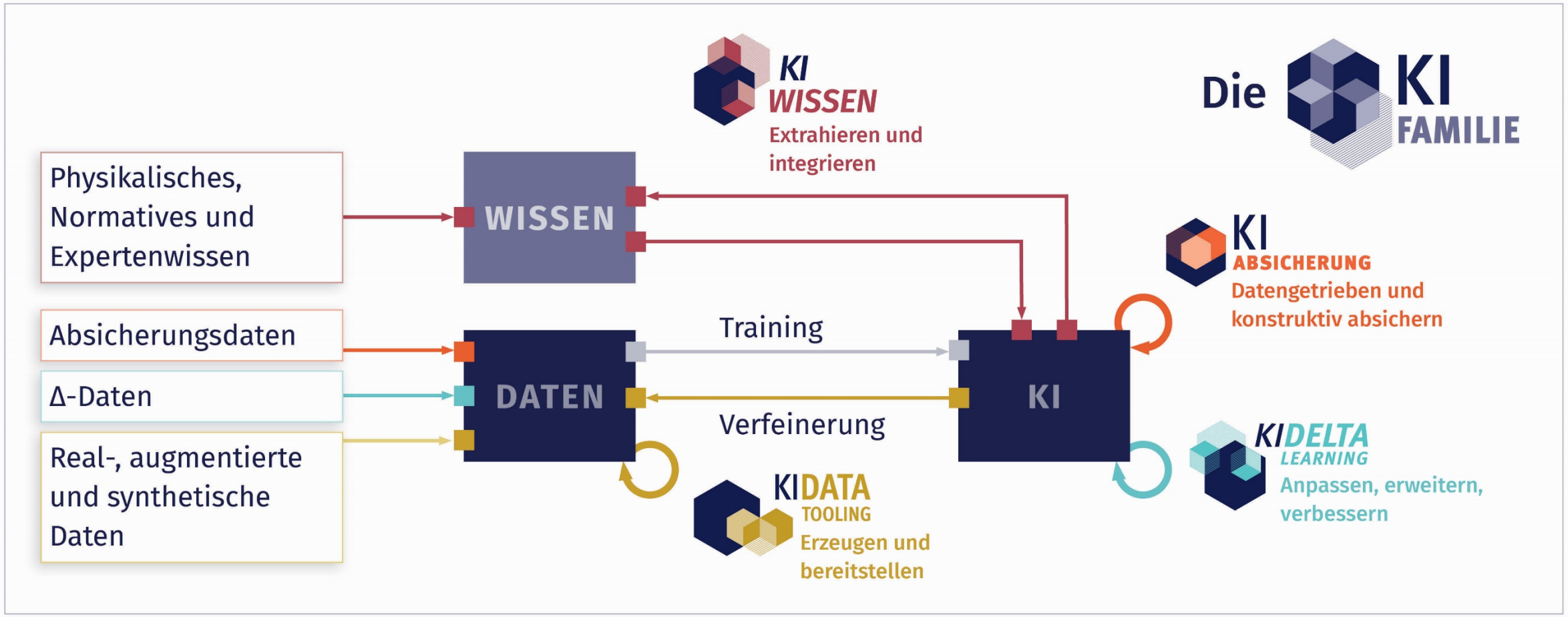
KI WISSEN Entwicklung von Methoden für die Einbindung von Wissen in maschinelles Lernen

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 Absicherung im Zusammenhang der KI Familie



Zusammenfassung der wichtigsten Projektergebnisse KI Absicherung

Neue Algorithmen für die Entwicklung und Absicherung tiefer neuronaler Netze:

- Wirksamkeitsbewertete Methoden und Maßnahmen zur Bestimmung und Reduktion systematischer Unzulänglichkeiten einer KI-Funktion.
- Hinsichtlich ihrer Erkennungsleistung und Absicherbarkeit weiterentwickelte KI-basierte Algorithmen zur Fußgängererkennung.

Neue Verfahren und Werkzeuge zum Testen und Nachweisen der Eigenschaften tiefer neuronaler Netze:

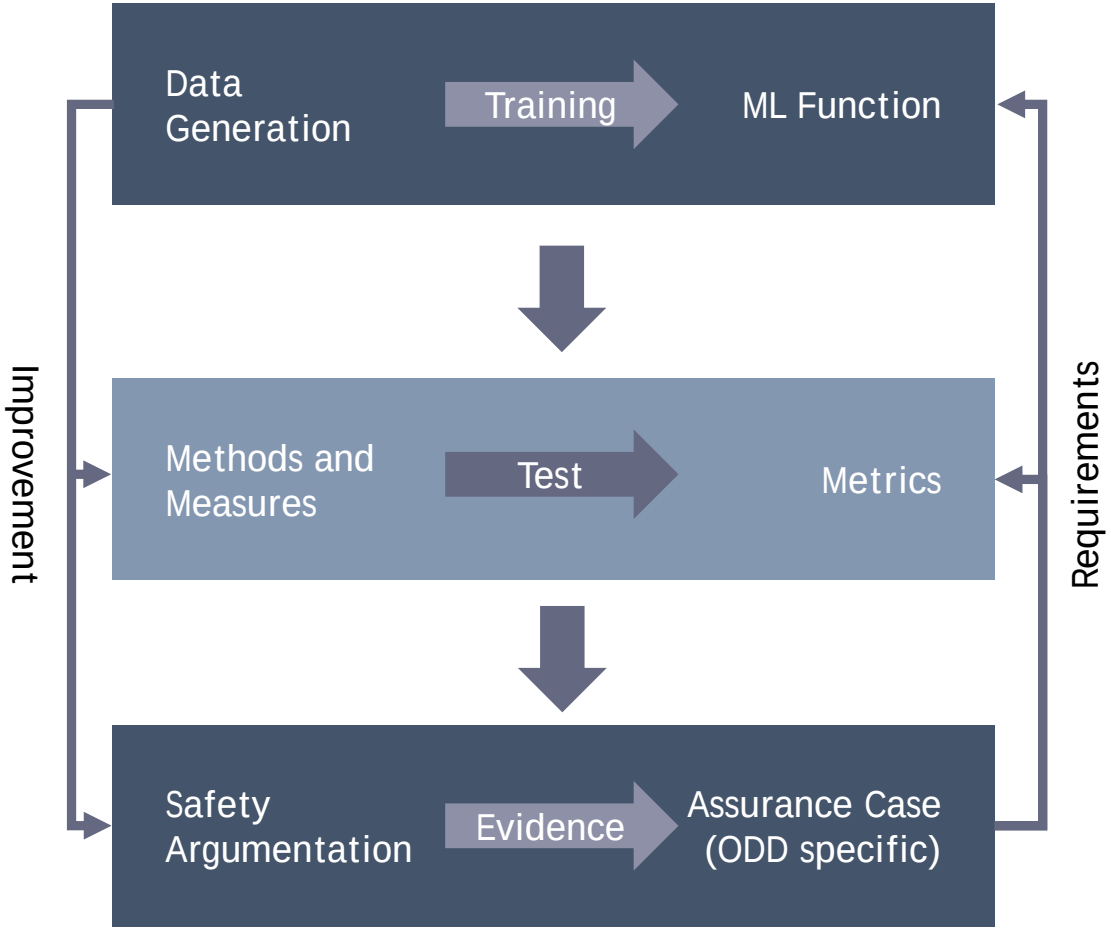
- Testmethodik und Prozesskette zum Nachweis der Absicherbarkeit einer datengetriebenen KI-Funktion.
- Prozess und Schnittstellen zur systematischen Erstellung von synthetischen Trainings- und Testdatensätzen zur Analyse und Bewertung systematischer Unzulänglichkeiten KI-basierter Verfahren.

2

Methodischer Ansatz und konzeptionelles Vorgehen

Von der datengetriebenen KI-Funktion hin zum Assurance Case an der Beispielapplikation Fußgängererkennung

- Prozessbegleitende Generierung von synthetischen Lern-, Test- und Absicherungsdaten.
- Entwicklung von Methoden und Maßnahmen, die die KI-Funktion bzgl. eines breiten Spektrums von Metriken verbessern.
- Entwicklung und Validierung von Testmethoden für diese Metriken.
- Stringente Argumentationskette für die KI-Funktion und ihre Operational Design Domain (ODD).



Konzeptionelles Vorgehen

1. Bereitstellen der KI-Funktion zur Fußgängererkennung.
2. Generieren von synthetischen Lern-, Test- und Absicherungsdaten.
3. Entwickeln und Bewerten von Methoden und Maßnahmen zur Absicherung der KI-Funktion.
4. Aufbauen einer gesamtheitlichen Absicherungsstrategie für die KI-Funktion.
5. Definieren und exemplarisches Umsetzen eines Assurance Case.

1. Bereitstellen der KI-Funktion zur Fußgängerdetektion

KI Absicherung entwickelt Algorithmen zur KI-basierten Fußgängererkennung basierend auf Bild- und Tiefendaten:

- Detektion in 2D/3D, Posenschätzung, semantische Segmentierung
- Fusionsansätze für Daten von Kamera- und Tiefensensoren

Erwartete Ergebnisse:

➔ State-of-the-Art-Analysen.

➔ Neuronale Netzarchitekturen.

➔ Trainierte Modelle.

➔ Qualitätsmetriken zur Beurteilung der Absicherbarkeit.

Synthetisch erzeugte Straßenszenen und ihre semantische Segmentierung

2. Generieren von synthetischen Lern-, Test- und Absicherungsdaten

Zur einfachen Kontrolle und Variation von Kontextdimensionen und Einflussfaktoren wird die systematische Entwicklung und Beurteilung von Absicherungsmethoden und -maßnahmen mit synthetisch erzeugten Trainings-, Test- und Validierungsdaten umgesetzt.

Erwartete Ergebnisse:

- ➔ Prozess zur methodischen Ableitung und Erzeugung von Corner Cases.
- ➔ Parallele Datengenerierung und Werkzeugentwicklung basierend auf Sensormodellen und korrektem Rendering.
- ➔ Methoden zur Bewertung synthetischer Daten.

Synthetisch erzeugte Daten: Szenenvariationen bei gleichbleibender Sensorposition

3. Entwickeln und Bewerten von Methoden und Maßnahmen zur Absicherung der KI-Funktion

Methoden und Maßnahmen zur Bestimmung und Reduktion systematischer Unzulänglichkeiten der KI-Funktion werden entwickelt, kombiniert und bewertet.

Erwartete Ergebnisse:

- Werkzeugkasten mit hinsichtlich ihrer Wirksamkeit auf die Sicherheit bewerteten Methoden und Maßnahmen für die Absicherung der KI-Funktion.
- Liste inhärenter und systematischer Unzulänglichkeiten von tiefen neuronalen Netzen.
- Wirksamkeits- und Sicherheitsmetriken für KI-Algorithmen und Maßnahmen.

Heatmap zur Plausibilisierung der KI-Funktion

4. Aufbauen einer gesamtheitlichen Absicherungsstrategie für die KI-Funktion

Der Systemkontext wird durch eine gemeinsame Beschreibungssprache und Ontologie festgelegt.

Die KI-spezifischen Unzulänglichkeiten und Mitigationsmaßnahmen werden analysiert und bewertet.

Erwartete Ergebnisse:

- ➔ Nachweis der hinreichenden Mitigation der systematischen Unzulänglichkeiten eines beispielhaften KI-Moduls zur Fußgängererkennung.
- ➔ Testverfahren zum Nachweis der sicherheitsrelevanten Wirksamkeit von Maßnahmen.
- ➔ Erzielung eines Konsens zum stringenten Aufbau einer Prozesskette und Testmethodik zum Nachweis der Absicherbarkeit einer datengetriebenen KI-Funktion zur Fußgängererkennung.

5. Assurance Case und Gesamtargumentation

Definition und exemplarische Umsetzung eines systematischen und gesamtheitlichen Vorgehens zur Absicherung einer spezifischen KI-Funktion zur Fußgängererkennung.

Eingaberaumstrukturierung und -formalisierung

- Definition Grund- und Nutzungskontext
- Domänenanalyse, Beschreibungssprache, Ontologie

Testmethoden und Testen

- Methoden zur Identifikation fehlender Testdaten
- Teststrategie und -plan

- Trainierte KI-Funktionsalgorithmen
- Synthetische Daten und Corner Cases
- KPIs, Methoden und Maßnahmen zur Absicherung von KI- Funktionen

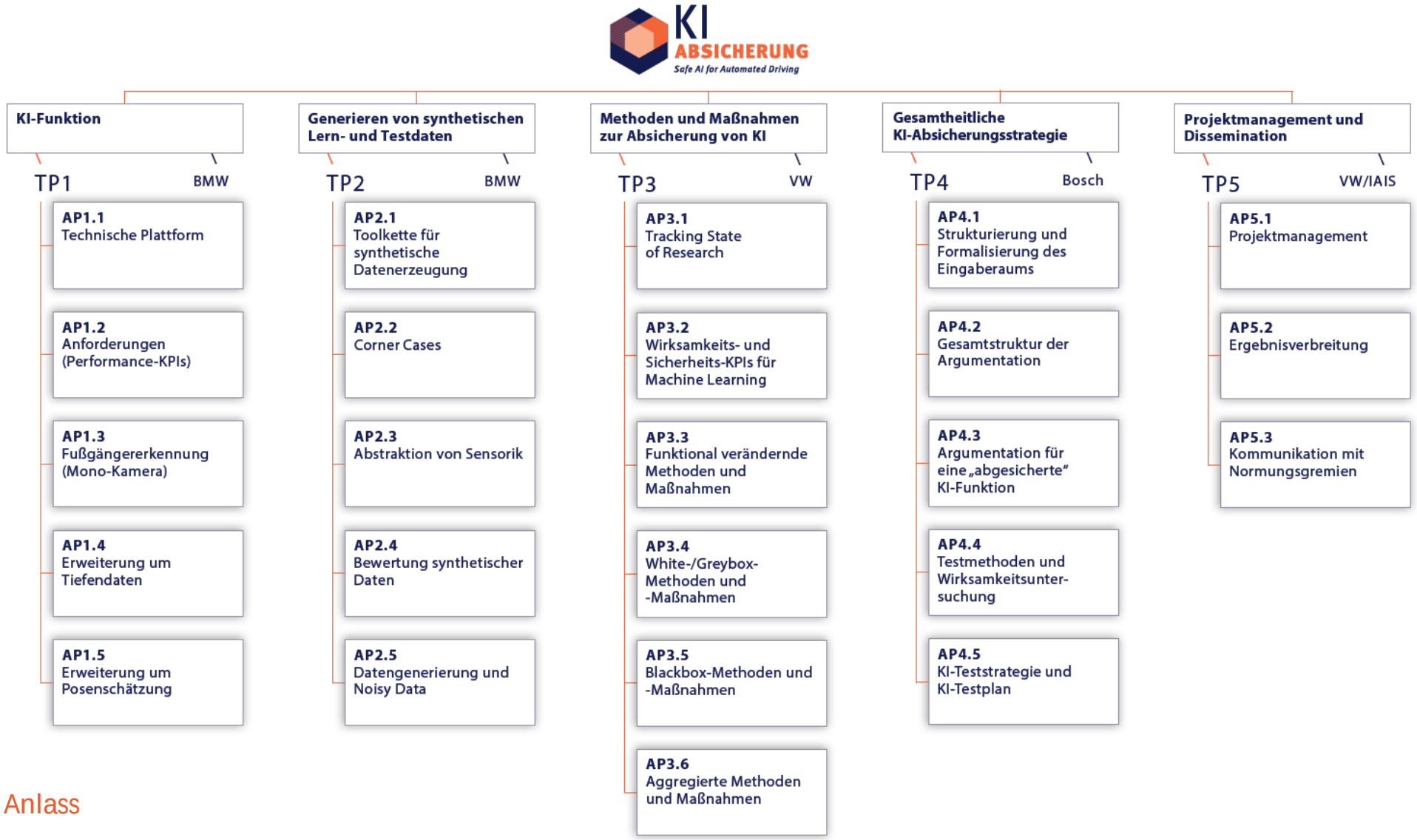
Anforderungen und Sicherheitsziele

- Gesamtfunktion und Systemarchitektur
- Safety Contracts inkl. Assumptions und Guarantees

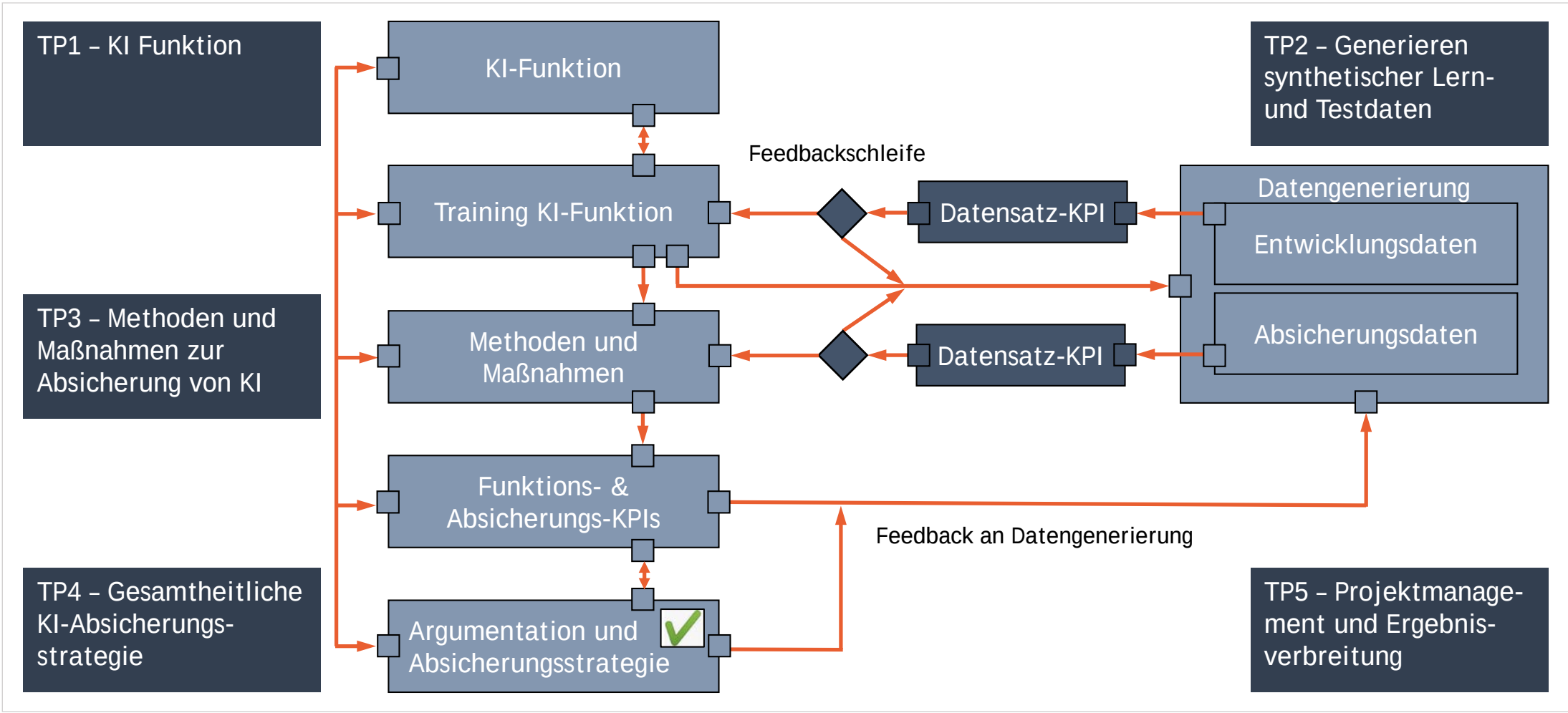
3

Projektstruktur

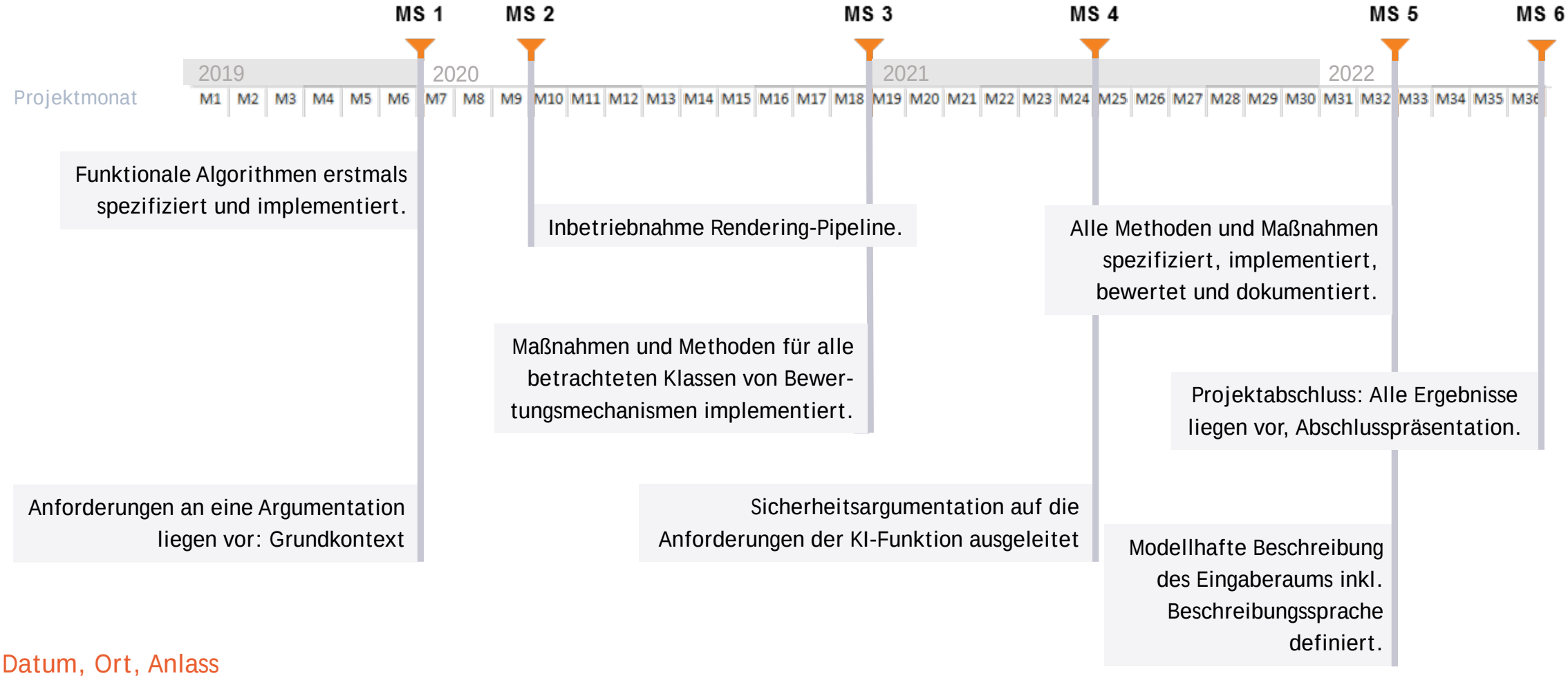
Projektstruktur mit Teilprojekten und Arbeitspaketen



Zusammenarbeit der Teilprojekte in KI Absicherung



Überblick Projektmeilensteine



Ansprechpartner

Projektkoordinator:
Dr. Stephan Scholz,
VOLKSWAGEN AG, Group | Autonomous Driving
Brieffach 011/1799/1
D-38436 Wolfsburg

Projektmanagement:
EICT GmbH,
EUREF Campus Haus 13, Torgauer Straße 12-15,
D-10829 Berlin
E-Mail: ki-absicherung-projektmanagement@eict.de

Stellv. Konsortialleiter/Wissenschaftlicher Koordinator:
PD Dr. Michael Mock
Fraunhofer IAIS, Knowledge Discovery,
D-53754 Sankt Augustin

E-Mail: ki-absicherung-konsortialfuehrung@eict.de

KI ABSICHERUNG

Safe AI for Automated Driving

KI Absicherung 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 Energie gefördert.

KI FAMILIE

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

www.ki-absicherung-projekt.de @KI_Familie KI Familie