

Autonome Systeme in Wirtschaft und Gesellschaft

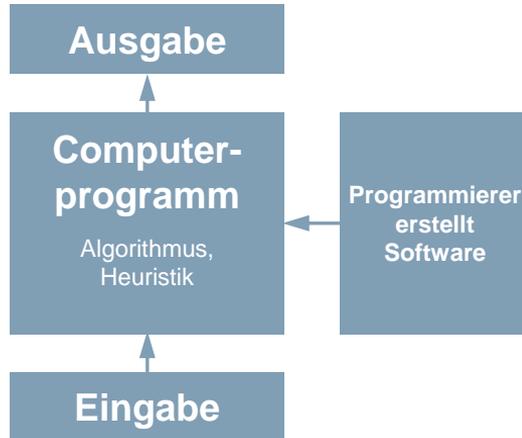
Prof. Dr. Henning Kagermann
Präsident acatech

Jahrestagung des Deutschen Ethikrates
Berlin, 21. Juni 2017



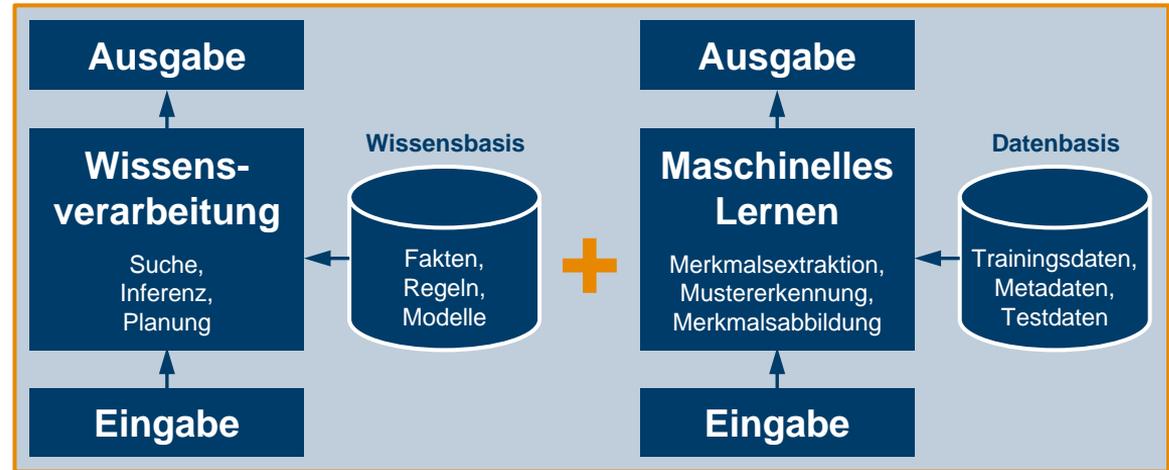
Renaissance der Künstlichen Intelligenz

Von programmierten zu selbstlernenden Systemen



Flaschenhals: Programmierer

- Hoher Entwicklungsaufwand
- Aufwändige Adaption
- Geringe Erklärungsfähigkeit



Flaschenhals: Wissensbasis

- Hoher Entwicklungsaufwand
- Hoher Pflegeaufwand
- Gute Erklärungsfähigkeit

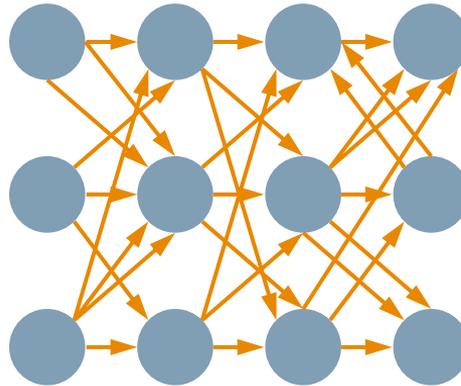
Flaschenhals: Trainingsdaten

- Geringer Entwicklungsaufwand
- Leichte Anpassbarkeit
- Schlechte Erklärungsfähigkeit



Maschinelles Lernen mit neuronalen Netzen

Deep Learning



Unfall
auf der
Landstraße

- **Verbindungsmuster** zwischen Neuronen ersetzt „Programm“
 - **Lernen durch Daten** ändert die Verbindungsmuster
- Ersetzt Programmierung

Eingabe

Bild-
elemente

Objekt-
teile

Objekte

Szenen-
und
Objekt-
modell

Ausgabe

Von automatisierten zu selbststeuernden Systemen

- **Automatisierte Systeme:** Vorgegebener Handlungsablauf wird selbständig ausgeführt
→ Ablauf ist vorhersehbar:
aber keine Anpassung an unvorhergesehene Situationen
- **Lernende Systeme:** Automatischer Aufbau der Wissensbasis durch Training (Test vor Inbetriebnahme)
- **Selbstlernende Systeme:** Wissensbasis wird im Betrieb fortlaufend erweitert
- **Selbststeuernde / selbstregulierende Systeme:** erreichen ein vorgegebenes Ziel selbstständig und situationsadaptiv
→ Ablauf wird angepasst, ist aber nur eingeschränkt vorhersehbar





Selbststeuerung als wesentliches Merkmal autonomer Systeme

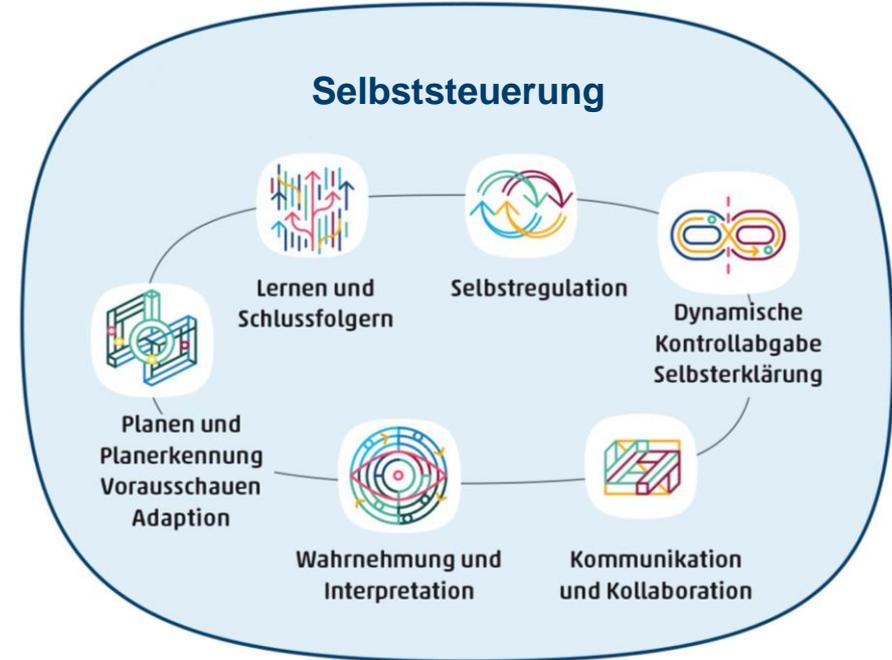
▪ Komponenten autonomer Systeme

- Sensorik
- Selbststeuerung
- Aktorik

▪ Merkmale:

- Selbstlernfähigkeit
- Selbsterklärungsfähigkeit
- Fehlertoleranz und Resilienz
- Kooperativität
- Proaktivität

▪ Kurz- vs. Langzeitautonomie

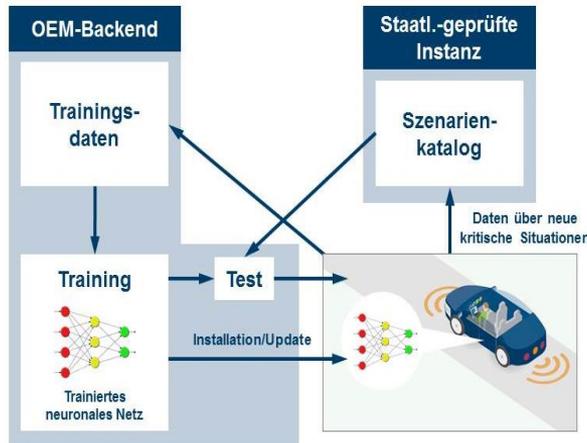


Quelle: Fachforum Autonome Systeme im Hightech Forum: Abschlussbericht 2017



Aufbau und Erweiterung der Wissensbasis

Beispiel automatisiertes Fahren



▪ Lernende Systeme:

- Test auf Basis des zentralen Szenarienkatalogs
- Freigabe pro Fahrzeugtyp
- Identisches neuronales Netz für Fahrzeuge eines Typs
- Permanente Verbesserung des Trainingsmaterials
- Periodisches Re-Training und Test

▪ Selbstlernende Systeme (entwicklungsoffen):

- Aktualisierung der eigenen Wissensbasis im Betrieb
- Unterschiedliche Wissensbasis individueller Fahrzeuge
- Noch nicht empfehlenswert für sicherheitskritische Funktionen
- Einstieg in Selbstverbesserung

→ Vielfalt und Qualität der Trainingsdaten ist erfolgskritisch.

Generierung von Trainingsdaten aus virtuellen Modellen



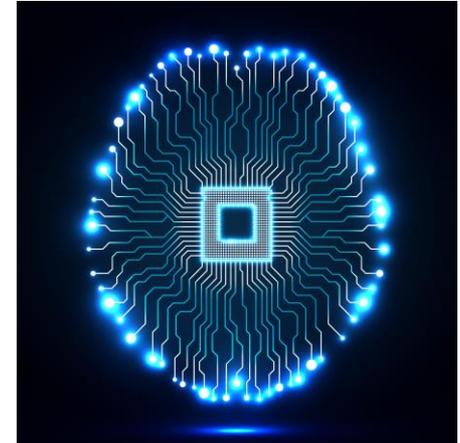
- **Deep-Learning-Verfahren**
liefern derzeit die besten Ergebnisse
- **Voraussetzung:**
Ausreichende Trainingsdaten
- **Ansatz: Virtuelle Modelle** der realen Welt
 - Automatische Generierung
von Trainingsdaten
 - Vorteil: enorme **Variationsmöglichkeiten**



Bildquelle: DFKI

Bislang ungelöste Probleme

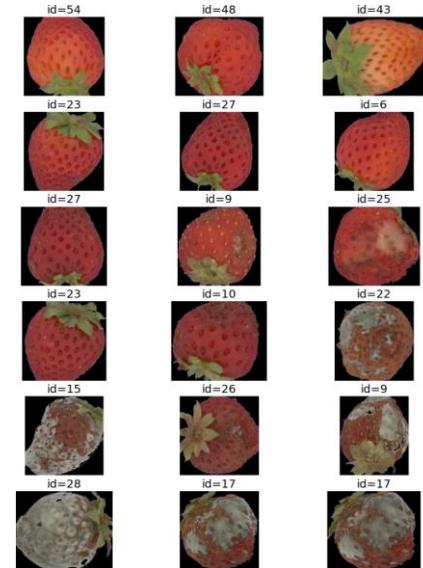
- Verständliche **Erklärung und Begründung** der Lernergebnisse für den Endanwender
- **Robustheit** gegenüber geringfügigen Transformationen der Trainingsdaten
- **Überanpassung** durch zu lange Trainingsdauer
- **Aufwandsreduktion** zur Datenannotation für das überwachte Lernen
- Erfolgreiches Lernen bei sehr **wenigen Trainingsdaten**
- Garantierte **Konfidenzniveaus**



Bildquelle: vladystock / Fotolia

Einige Anwendungsfelder

- **Spiele:** Schach, Jeopardy, Bridge, Go
- **Wohnumfeld:** Rasenmäher, Staubsauger, Smart Home-Anwendungen
- **Produktion:** Kollaborative Roboter, fahrerlose Transportsysteme, Gourmet-Burger zu Fast-Food-Preisen
- **Mobilität:** selbstständiges Einparken
- **Verwaltung:** Automatisches Durchsuchen von Verträgen
- **Qualitätssicherung:** Verhinderung von Maschinenausfall, Vorhersage der Haltbarkeit von Früchten, digitales Feldfruchtgedächtnis
- **Kundenpflege:** Voraussichtliche Vertragskündigung
- **Beratung:** Robo-Advisor für einfach Auskünfte
- **Anomalieerkennung:** z.B. Wirtschaftskriminalität
- **Kreative Arbeiten:** Berichte, Bilder, musikalische Kompositionen



Bildquelle: Ralf Herbrich (2017) – Machine Learning bei Amazon, Beispiel „Predicting Longevity“, Age Aligned Strawberries



Hightech Forum: Fachforum Autonome Systeme 2015 - 2017



„Wir konzentrieren uns auf autonome Systeme – es ist eine Kerntechnologie, die wir als sehr wichtig ansehen.“
Tim Cook, CEO Apple

- Themenauswahl entspricht den wichtigsten Anwendungsbereichen bis 2030:
 - **Produktion:** Industrie 4.0, Smart Services
 - **Mobilität:** Straßen- und Schienenverkehr
 - **Smart Home:** Sicherheit, Assistenz / Pflege, Energieeffizienz
 - **Menschengefährdende Umgebungen:** Sicherheit, Rettung, Rückbau von Kernkraftwerken

→ **Ergebnis: Zukunftsprojekt „Lernende Systeme“**

Autonome Systeme in der Produktion I

Chancen und Risiken

- **Chancen:**
 - Hoher **Individualisierungsgrad** von Produkten und Dienstleistungen
 - Agile, lernende, **wandlungsfähige Fabrik:** Selbstständige Anpassung an Nachfrageschwankungen
 - **Entlastung** der Arbeitnehmer durch Assistenzsysteme
- **Risiken:**
 - **Akzeptanz:** Wandel der Arbeitsplätze und Aufgabenzuschnitte / Arbeitsplatzverluste
 - Erhöhtes Risiko von **Cyberangriffen**
 - Mobilisierung der **KMU** (Finanzmittel, Fachkräfte)



Quelle: Fachforum Autonome Systeme im Hightech
Forum: Abschlussbericht 2017





Autonome Systeme in der Produktion II

Sichere Zusammenarbeit in hybriden Teams

- **Hybride Teams:**
 - Menschen (unterstützt durch Augmented Reality)
 - Autonome Roboter und Softbots als Partner
- **Arbeitsteilung** gemäß spezifischer Fähigkeiten, gemeinsame Problemlösung
- **Autonome Reorganisation** bei unvorhergesehenen Ereignissen

→ **Roboter müssen sich dem Verhalten der Menschen anpassen.**

Autonome Systeme in der Mobilität I

Chancen und Risiken

▪ Chancen:

- **Verringerung der Unfälle** durch menschliches Versagen
- Gesellschaftliche **Teilhabe**: verbesserter Zugang zu Mobilität
- Optimierung des **Verkehrsflusses** (weniger Staus)
- Reduktion der **Umweltbelastung** und des Flächenverbrauchs

▪ Risiken:

- **Akzeptanz** bei Unfällen durch maschinelles Fehlverhalten (s. Tesla-Unfall)
- **Missbrauch** von Bewegungsprofilen
- **Rebound-Effekte**

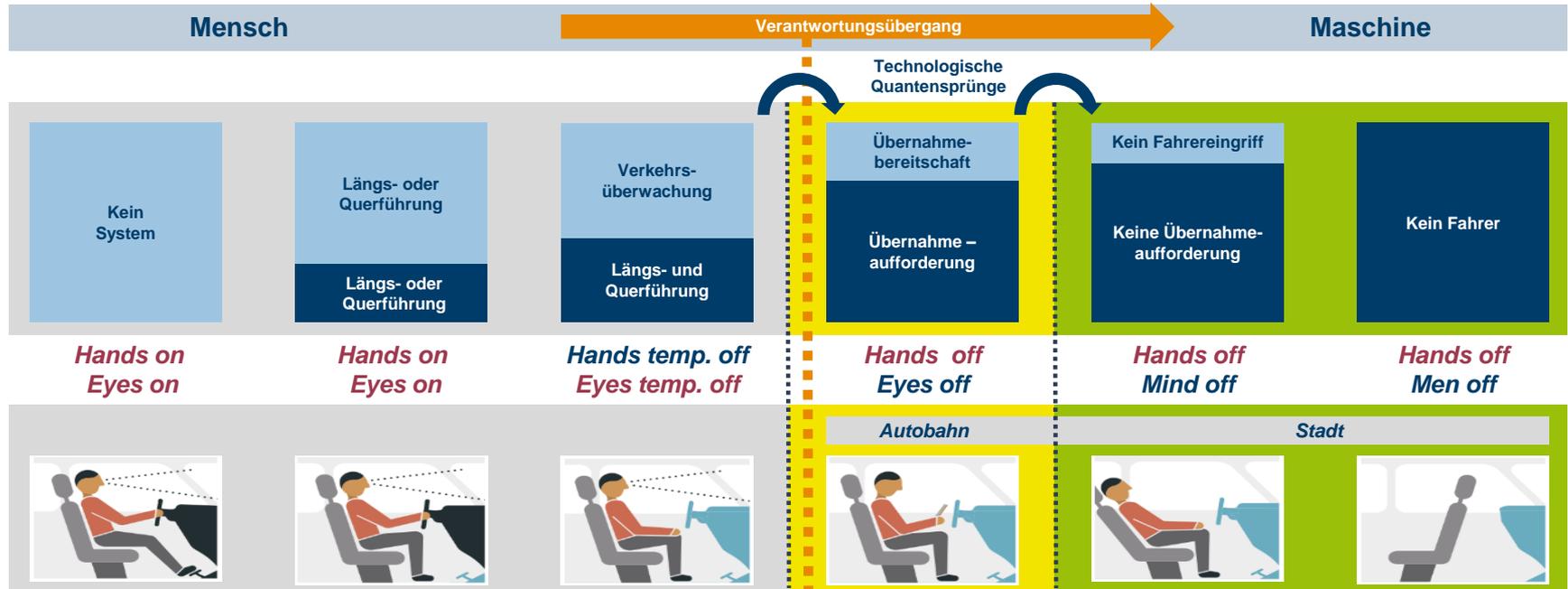


Quelle: Fachforum Autonome Systeme im Hightech Forum: Abschlussbericht 2017



Autonome Systeme in der Mobilität II

Automatisierungsstufen



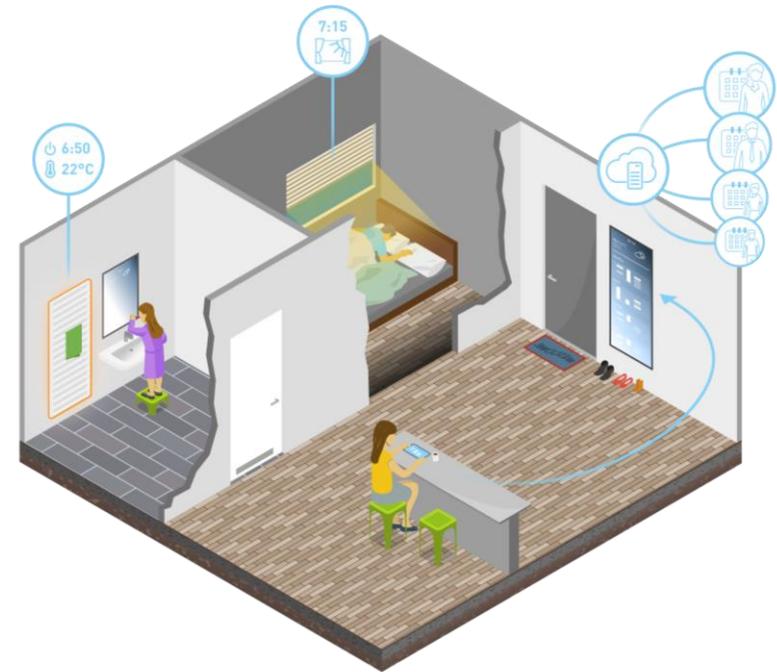
Quelle: Reinhard Stolle / BMW (2017): Machine Learning für autonomes Fahren



Autonome Systeme im Smart Home I

Automatisierungsstufen

- **Keine Automatisierung**
- **Assistiert:** einzelne Bedienelemente über feste Profile
- **Teilautomatisiert:** für einzelne Funktionen mit umfangreicher Programmierung
- **Hochautomatisiert:** temporäre Selbststeuerung mit häufiger Kontrollübergabe
- **Vollautomatisiert:** Selbststeuerung einzelner Funktionen, z.B. Energiemanagement
- **Vollständige Autonomie:** Vernetzung und Selbststeuerung aller Funktionen, volle Multi-Nutzer-Fähigkeit



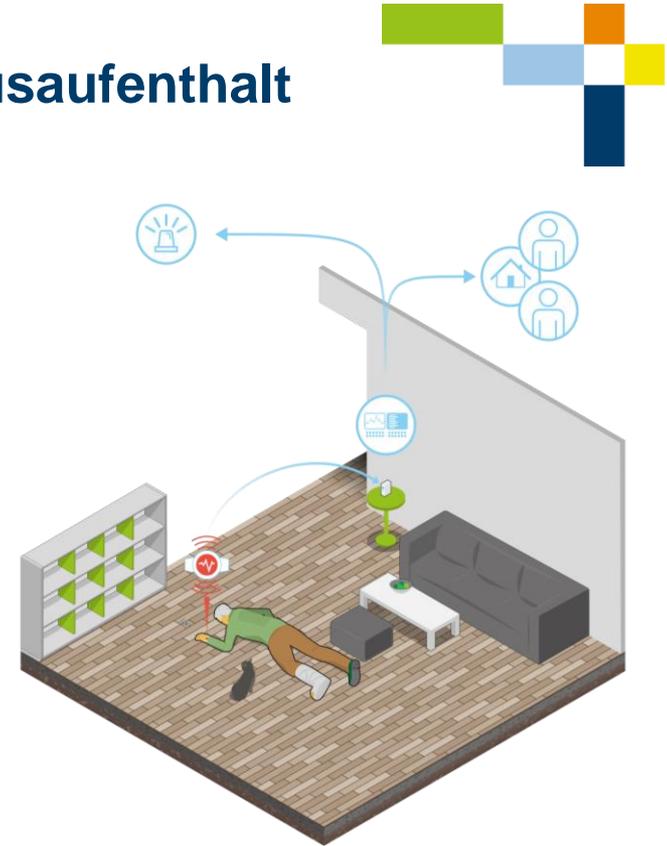
Quelle: Fachforum Autonome Systeme im Hightech Forum: Abschlussbericht 2017

Smart Home II: Assistenz nach Krankenhausaufenthalt

Einführung anhand von Use-Cases

- **Vernetzte E-Health-Systeme**
im Smart Home:
 - Unterstützung der Rehabilitation
 - Nachsorgende gesundheitliche Überwachung
 - Informationen für Ärzte
 - Situationsangepasste Alarmierungskette bei Unregelmäßigkeiten (Sturz, Vitalfunktionen)
- **Chance:** Selbstständiges Leben zu Hause
- **Risiko:** Komplexität der Systeme, Datendurchgängigkeit

Quelle: Fachforum Autonome Systeme im Hightech Forum: Abschlussbericht 2017





Anwendungsszenario

Menschengefährdende Umgebungen

▪ Beispiele:

- Einsatz im Katastrophenfall
- Sicherheit und Gefahrenabwehr

▪ Risiken:

- Dual-Use Potenzial
- Langzeitautonomie

▪ Chance:

- Weniger Gefahr für Menschen
(z.B. Minenräumung, Abtragen kontaminierter Böden)
- Hohe gesellschaftliche Akzeptanz
- Einstieg in Einführung autonomer Systeme

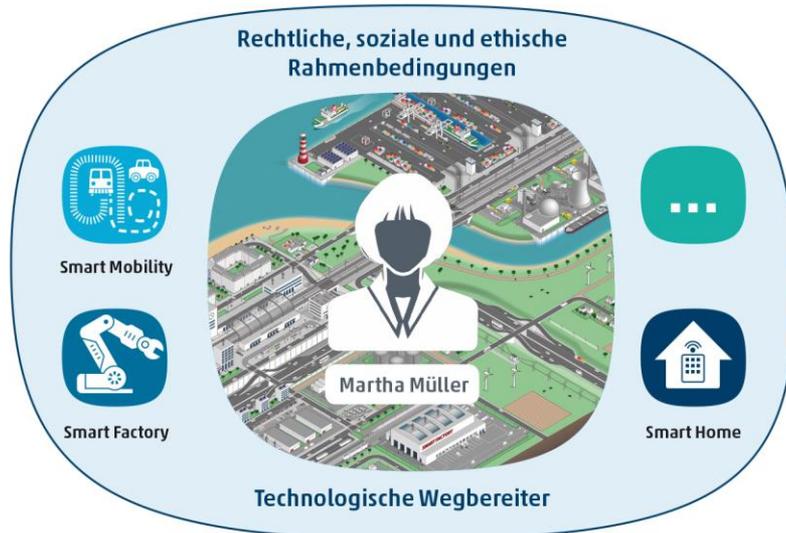


Quelle: Fachforum Autonome Systeme im
Hightech Forum: Abschlussbericht 2017



Eine ganzheitliche Betrachtungsweise ist notwendig

Ein Tag im Leben von Martha Müller



- Die Menschen werden autonomen Systeme **täglich** und **in vielen unterschiedlichen Lebenssituationen** begegnen.
- Anwendungsbereiche können **nicht mehr isoliert** betrachtet werden.
 - Konsistente **juristische / ethische Regeln**
 - Konsistente **Sicherheitsniveaus**
 - Konsistente **Mensch-Technik-Interaktion**

Quelle: Fachforum Autonome Systeme im Hightech Forum: Abschlussbericht 2017



Handlungsempfehlungen

- Autonome Systeme stufenweise und erfahrungsbasiert einführen
 - Beständiger und transparenter Lernprozess
 - Einführung über Nutzungsszenarien → Beteiligung aller gesellschaftlichen Akteure
- Frühzeitig langfristig angelegten **Dialog** etablieren, **Transparenz** über Chancen und Risiken herstellen
- **F&E-Programme** für technologische Wegbereiter etablieren
- **Kompetenzaufbau** zu maschinellem Lernen fördern
- **Normungsroadmap**: Interoperabilität und Portabilität sicherstellen
- Herausforderungen für **Arbeits-, Aus- und Weiterbildungspolitik** formulieren
- Regelungen zu **Datenschutz und IT-Sicherheit** anpassen



Ausblick

- Sorgen vor...
 - ... **Kontrollverlust:**
Wird es eine 'übermenschliche' Intelligenz mit der Fähigkeit zur Selbstverbesserung geben? Wann?
 - ... **gesellschaftlicher Spaltung:** Wer hat Zugang zu autonomen Systemen?
 - ... **Jobverlusten:** Können wir schnell genug neue Berufsbilder entwickeln?
Welche Tätigkeiten können Menschen besser verrichten?

- Chancen und Risiken kommunizieren, technologische Entwicklung kanalisieren
- Der Mensch im Mittelpunkt und als letzte Kontrollinstanz
- AS können Menschen unterstützen und Fähigkeiten ergänzen, aber nicht ersetzen

**Vielen Dank für
Ihre Aufmerksamkeit.**