

# Autonome Systeme

*Fähigkeiten und Anforderungen*

Werner Damm · Ralf Kalmar

## Mehrwert durch Autonomie

Wie viel Autonomie können, wie viel Autonomie sollen technische Systeme besitzen?

Heutzutage scheinen der technischen Realisierbarkeit immer höherer Grade an Autonomie (siehe nachfolgender Abschnitt) keine Grenzen gesetzt zu sein. Wir stehen vor der Schwelle der Markteinführung technischer Systeme, die extrem komplexe Umweltsituationen auf der Basis komplexer Verarbeitungsketten von Rohdatenaufbereitung, automatischer Objektidentifikation, Sensorfusion auf verschiedenen Ebenen bis hin zur Schaffung eines für die jeweilige Mission der Systeme genügend genauen digitalen Repräsentation der Wirklichkeit vollständig automatisiert vornehmen. Wir stehen vor der Markteinführung technischer Systeme, welche auf dieser Basis eine Situationsanalyse der möglichen Weiterentwicklungen dieser Wirklichkeit in einem Maße vornehmen können, welche die menschlichen Analysefähigkeiten weit übersteigt. Wir stehen vor der Markteinführung technischer Systeme, die für die Planung und Durchführung einer Mission keinerlei menschlicher Unterstützung bedarf, dessen kognitive Fähigkeiten dem Menschen für diesen Aufgabenkontext überlegen sind, also vollkommen autonom agieren können.

Der gerade veröffentlichte Bericht zur Hightechstrategie der Bundesregierung [3] zeigt zahlreiche Anwendungsfelder solcher hochautonomer Systeme auf, darunter alle Klassen von „Smart Systems“ wie Smart Mobility, Smart Health, Smart Production, Smart Energy, Smart Mobility – in denen also die „Intelligenz“ der Lösungsansätze durch die oben dargestellte Fähigkeit entsteht, zahlreiche Informationsquellen in Echtzeit zu solchen digitalen Weltbildern zusammenzufügen, und eine

Orchestrierung der Millionen von Subsystemen so vorzunehmen, dass übergeordnete Optimierungsziele wie etwa eine optimale Ressourcennutzung gewährleistet werden. Die daraus resultierenden Mehrwerte reichen von der Sicherung einer flächendeckenden Gesundheitsversorgung zu nachhaltiger Mobilität für alle, von drastischen Reduktionen des Energieverbrauches zur Steigerung der Produktivität und Produktqualität, von Katastrophenmanagement zur automatischen Kollisionsvermeidung unterschiedlichster Klassen von Verkehrsträgern. So erwartet etwa Philips durch die Integration von tragbarer Sensorik in der postoperativen Patientenüberwachung eine Reduktion von postoperativem Herzstillstand um 86 %, und durch eine „smarte“ Überwachung der kritischen Gesundheitsparameter in der ambulanten Pflege eine Reduktion der Pflegekosten um 34 % [9].

Bereits seit Jahren sind viele Systeme über regelungstechnische Lösungen automatisiert. Die Automatisierung ermöglicht, technische Systeme ohne die Notwendigkeit manueller Eingriffe effizient und effektiv zu betreiben. Die typischen Aufgaben zur automatischen Überwachung, Steuerung und Regelung von klassischen technischen Anlagen stellen oft physikalisch-technische Steuerketten dar und können gut mathematisch beschrieben und zur Entwicklungszeit mit Software- und Elektroniklösungen realisiert werden.

DOI 10.1007/s00287-017-1063-0  
© Springer-Verlag Berlin Heidelberg 2017

Werner Damm  
OFFIS e. V., Oldenburg  
E-Mail: damm@offis.de

Ralf Kalmar  
Fraunhofer IESE, Kaiserslautern  
E-Mail: ralf.kalmar@iese.fraunhofer.de

## Zusammenfassung

Getrieben durch aktuelle Diskussionen im Automobilbereich sind autonome Systeme in aller Munde. Tatsächlich sind automatisierte Systeme unterschiedlichen Autonomiegrads Bestandteil aktueller Roadmaps und Projektionen in vielen Branchen. In diesem Artikel werden die verschiedenen branchenspezifischen Taxonomien und Standards zusammengefasst und im Hinblick auf ihre funktionalen Fähigkeiten und Anforderungen an Methoden, Prozesse und Werkzeuge aus Sicht der Softwaretechnik charakterisiert.

Aktuelle Entwicklungen cyberphysischer Systeme gehen weit darüber hinaus: Die Verbindung von IT-Systemen mit eingebetteten Systemen, sowie die dynamische Vernetzung von Systemen untereinander schafft komplexe Wirkketten mit heterogenen Datenschnittstellen. Ähnlich wie bei der Automatisierung in den 80er-Jahren verspricht der autonome Betrieb eine höhere Effizienz, Effektivität und Qualität.

Diese Entwicklung vollzieht sich in vielen Anwendungsdomänen technischer Systeme. Wenn gleich sich die verschiedenen Anwendungsszenarien auf den ersten Blick unterscheiden, lassen sich viele enthaltene softwaretechnische Problemstellungen auf allgemeine Lösungsansätze abbilden, die im letzten Abschnitt dieses Artikels dargestellt werden. Besondere Herausforderungen ergeben sich dabei durch den Einsatz selbstlernender Systeme. Mit dem Einsatz selbstlernender Systeme sind den potentiellen Fähigkeiten autonomer Systeme scheinbar keine Grenzen gesetzt, sind sie dann doch in der Lage, auch bisher unbekannte, die Mission eines Systems beeinflussende, Artefakte der Systemumgebung zu erkennen und etwa dazugehörige Dynamikmodelle zu erlernen. Damit erschließen sich auch neue Klassen von Anwendungen wie etwa autonome Unterwasserfahrzeuge zur Exploration des Meeresbodens unter der Arktis.

Die durch diese Fähigkeiten sich eröffnenden Marktpotentiale liegen insgesamt im Bereich mehrerer 100 Billionen USD. Die folgenden Schätzungen sind der Studie des EU-Projektes Platforms4CPS entnommen [13]:

- Bis 2035 werden 10 % aller Verkäufe autonome Fahrzeuge sein – dies entspricht etwa 12 Millionen Fahrzeugen und einem Marktvolumen von 39 Milliarden USD;
- der Markt für Flugverkehrsmanagement wird Schätzungen zufolge von 50,01 Mrd. USD im Jahr 2016 auf einen Wert von 97,30 Mrd. USD bis 2022 wachsen bei einer durchschnittlichen jährlichen Wachstumsrate (CAGR) von 11,73 %;
- der Weltmarkt für Luftfahrtrobotik erwartet eine durchschnittliche jährliche Wachstumsrate von ca. 17,7 % innerhalb des nächsten Jahrzehnts, um bis 2025 einen Wert von ungefähr 7,9 Mrd. USD zu erreichen (nach „Markets and Markets“);
- das Marktvolumen für Drohnen wird auf 13,22 Mrd. USD 2016 geschätzt und soll bis 2022 28,27 Mrd. USD bei einer jährlichen Wachstumsrate von 13,51 % erreichen;
- der Markt für führerlose Oberflächenfahrzeuge wird auf ein Wachstum von einem Wert von 437,57 Millionen USD im Jahr 2016 auf 861,37 Millionen USD bis 2021 bei einer jährlichen Wachstumsrate von 14,51 % geschätzt;
- das Marktvolumen für autonome Unterwasserfahrzeuge wird von 2,29 Mrd. USD im Jahr 2015 auf 4,00 Mrd. USD bis 2020 klettern bei einer geschätzten jährlichen Wachstumsrate von 11,90 %;
- für den Markt des industriellen Internet der Dinge wird ein Wachstum erwartet von 110 Mrd. USD 2020 auf 123 Mrd. USD bis 2021;
- der Technologiemarkt für tragbare Geräte hat 2016 einen Marktwert von 28,7 Mrd. USD. „Gartner“ prognostiziert, dass der Markt für die Gesamtheit an tragbaren Gerätschaften mit einer durchschnittlichen jährlichen Wachstumsrate von 17,9 % zwischen 2015 und 2017 wachsen wird; dabei ist das Segment der am Handgelenk zu tragenden mobilen Geräte mit einer jährlichen Wachstumsrate von 30 % das am schnellsten Wachsende;
- der Markt der Microgrids wurde 2015 mit 16,58 Mrd. USD bewertet. „Markets and Markets“ erwartet ein Volumen von 38,99 Mrd. USD im Jahr 2022 bei einer jährlichen Wachstumsrate von 12,45 %.

Die aktuellen Entwicklungen am Markt zeigen, dass die technologischen Möglichkeiten ausgereizt werden, um durch höhere Automatisierungsgrade neue Klassen von Produkten oder Anwendungen zu erschließen. Welche Entwicklungen tatsächlich gesellschaftlich sinnvoll sind, oder welche gesell-

schaftlichen Auswirkungen diese Entwicklungen haben, diese Frage rückt erst langsam in das Zentrum der Diskussion. Mit dem zunehmenden Grad der Autonomie ändert sich offensichtlich die Qualität der Mensch-Technik-Interaktion. Während heutzutage der Mensch nicht nur als Endanwender, sondern in vielen Anwendungen Teil der Regelungsstrecke ist (Human-in-the-Loop), führt die zunehmende Autonomie zu einem Trend, der in der jetzt laufenden Entwicklung die Interaktion zwischen Mensch und Technik auf eine höhere Abstraktionsstufe stellt: Das autonome System vermittelt in geeigneten Abstraktionen etwa unter Verwendung von Virtual-Reality-Techniken dem Menschen einen Einblick in den Teil „seines“ digitalen Weltbildes, der für die aktuelle Entscheidungsfindung zentral ist. Umgekehrt kann der Mensch komplexe Abläufe durch intuitiv verständliche Mensch-Maschine-Schnittstellen leicht beeinflussen. Einher mit dem zunehmenden Abstraktionsgrad dieser Kommunikation gehen Anforderungen an Qualifikation und Schulung. Gleichzeitig entfallen Arbeitsplätze für niedrig qualifiziertes Personal. Die Vielzahl der beobachtenden Datenquellen stellt signifikante Risiken für die Verletzung des Datenschutzes dar. Die hochgradig vernetzte Systemarchitektur stellt extrem hohe Anforderungen an den Schutz des Systems, um den Ausfall ganzer Teilsysteme durch Cyberattacken mit entsprechend katastrophalen Auswirkungen zu vermeiden. Mit wachsender Autonomie stellt sich die Frage, welche Wertevorstellungen den zugrunde liegenden Entscheidungsverfahren unterliegen, und ob diese mit unseren Wertevorstellungen konform sind. So fordert das Europäische Parlament in seiner Resolution vom 16. Februar dieses Jahres [7]:

- das Prinzip der Transparenz, welches besagt, dass es immer möglich sein sollte, das Grundprinzip und die Argumentation hinter jeder mit Hilfe von KI getroffenen Entscheidung nachzuvollziehen, die mit einer substantiellen Auswirkung auf Menschenleben verbunden ist;
- dass es immer möglich sein muss, die auf KI basierende Berechnungsgrundlage eines Systems in einer für Menschen verständlichen Form darzustellen;
- dass fortschrittliche Roboter mit einer sogenannten „Blackbox“ ausgestattet sein sollten, welche alle

Daten bei jeder durch die Maschine ausgeführten Transaktion aufzeichnet, inklusive der Logik, die zu ihren Entscheidungen beigetragen hat.

Schließlich ist nicht zuletzt getrieben durch die bevorstehende Markteinführung hochautonomer Fahrzeuge die Frage des Haftungsrechtes neu zu regeln.

Dieses Themenheft geht somit bewusst über die rein fachliche Dimension hinaus. Wie hochautonome Systeme so entworfen werden können, dass sie nicht nur marktwirtschaftlich relevant sind, sondern auch eine hohe gesellschaftliche Akzeptanz haben, dieser Herausforderung muss sich die Informatik stellen. Es ist Zeit für ein Neudenken der Entwurfsprozesse, in denen die gesamtgesellschaftlichen Auswirkungen der von uns geschaffenen autonomen Systeme fest integrierter Bestandteil sind.

### Autonomiegrade am Beispiel verschiedener Branchen

Das in der Öffentlichkeit bekannteste Beispiel ist das der hochautomatisierten und autonomen Fahrzeuge in der **Automobilindustrie**. Verschiedene Hersteller haben angekündigt, entsprechende Fahrzeuge bereits in drei bis vier Jahren auf den Markt zu bringen. Tatsächlich leisten bereits heute verfügbare Assistenzsysteme schon Erstaunliches. Allerdings ist der Weg vom teilautomatisierten – manche Hersteller sprechen auch vom „pilotierten“ – Fahren hin zur Autonomie noch weit: Beim teilautonomen Fahren (SAE-Automationsstufe 2, vgl. [14]) ist der Mensch in der Verantwortung und muss kurzfristig selbst eingreifen können. Darüber hinaus sind die Systeme auf stark normierte Umgebungen beschränkt (z. B. Autobahnfahrt). In der Stufe des hochautomatisierten Fahrens (SAE-Automationsstufe 3) darf der menschliche Fahrer sich auch anderen Dingen zuwenden, d. h. die Software muss das Fahrzeug in jedem Fall sicher steuern (bzw. im Fall eines Fehlers in einen sicheren Zustand, z. B. Halt am Straßenrand, überführen). Fahrzeuge, die vollautomatisiert, d. h. ohne menschlichen Fahrer auskommen (SAE-Automationsstufe 4), stellen die höchste Stufe dar, wobei hier keine Vorgaben zu Geschwindigkeiten oder Umgebungseigenschaften gemacht werden.

Maßgeblich getrieben wird die Entwicklung bei PKW dabei primär nicht durch den Wunsch der

Menschen, das Steuer aus der Hand legen zu wollen, sondern vielmehr durch neue Mobilitätsanbieter und entsprechende Dienstleistungen, die neue Kundensegmente erschließen, oder z. B. in Städten einen effizienteren, schnelleren Nahverkehrsdienst anbieten möchten. Bei Nutzfahrzeugen ermöglicht die Automatisierung eine Entlastung des Fahrers, der sich dann auch anderen Aufgaben widmen kann und insgesamt produktiver ist.

In der **Bahntechnik** werden insbesondere im Bereich der Untergrundbahnen bereits etliche Abläufe automatisiert. Vereinfachend wirkt, dass man eine relativ homogene Systemlandschaft vorfindet, die Querregelung der Fahrzeuge nicht notwendig ist und viele Strecken seitlich abgesperrt sind. Allerdings kommt die übergeordnete Steuerung und Koordinierung der Abläufe hinzu, weshalb der Internationale Verband für Öffentliches Verkehrswesen (UITP) in seinen Automatisierungsgraden (engl. grades of automation level, GoA, vgl. [15]), die Überwachung und Steuerung des Gesamtsystems mit einbezieht. Ein automatisiertes Zugsystem (ATS) besteht dabei aus den folgenden drei Komponenten: der automatischen Zugsicherung (ATP), der automatischen Zugsteuerung (ATO) und der Zugüberwachung (ATS). Die Zugsicherung kontrolliert die Einhaltung von Abständen im Gleisabschnitt (Blocksicherung) sowie die Geschwindigkeit. Die Zugsteuerung fährt den Zug gemäß Fahrplan und bedient z. B. Türen. Die übergeordnete Zugüberwachung schließlich überwacht alle Fahrten und die Infrastruktur und gibt entsprechende Meldungen an die Leitstelle weiter.

Eine Umsetzung solcher Systeme im Metrobereich ist aufgrund des homogenen Zugmaterials und der isolierten Infrastruktur am einfachsten möglich. Die Konzepte lassen sich aber auch in andere Bereiche übertragen, bis hin zum Rangierbahnhof. Hier ergeben sich jedoch Herausforderungen im länderübergreifenden Betrieb oder auch in der Komplexität der Umgebung, z. B. im S-Bahn-Betrieb an unterschiedlichsten Bahnhöfen.

Treiber für voranschreitende Automatisierung in der Bahntechnik ist die größere Wirtschaftlichkeit solcher Lösungen. Beispielsweise durch Energieeinsparungen bei abgestimmten Brems- und Beschleunigungsvorgängen im gleichen Netz.

Im **Luftfahrtbereich** werden schon lange automatische Flugsteuerungen eingesetzt. Insbesondere

bei den unbemannten Luftfahrzeugen, welche überwiegend im militärischen Kontext eingesetzt werden, hat man die Automationsgrade im Vergleich zu anderen Domänen noch weiter in Richtung Aufgabenplanung und Missionssteuerung erweitert: die zehn ALFUS-Stufen (Autonomy Levels for Unmanned Systems, Autonomiestufen für unbemannte Systeme [11]) nutzen dabei drei Dimensionen, um die Fähigkeiten des Systems zu charakterisieren: Unabhängigkeit von menschlichen Eingriffen (HI), Aufgabenkomplexität (MC) und Umgebungskomplexität (EC). Sie charakterisieren gemeinsam die Fähigkeiten für autonome Operationen.

Bei den technologischen Lösungen kommen deshalb bei den höchsten Stufen auch Themen wie Schwarmverhalten, adaptive Kommunikation zwischen Fahrzeugen und eigenständiges Lernen hinzu, die andere Taxonomien so noch nicht enthalten.

*Definition Autonomie:* “A UMS’s own ability of integrated sensing, perceiving, analyzing, communicating, planning, decision-making, and acting, to achieve its goals as assigned by its human operator(s) through designed human-robot interface (HRI)”.

*Definition Contextual Autonomous Capability (CAC):* “An unmanned system’s contextual autonomous capability is characterized by the missions that the system is capable of performing, the environments within which the missions are performed, and human independence that can be allowed in the performance of the missions.” [11]

In der **Produktionstechnik** und Prozesstechnik sind automatisierte Abläufe bereits seit Einführung der Speicherprogrammierbaren Steuerung (SPS) in den 80er-Jahren Standard. Allerdings sind diese Abläufe sehr starr und auf die Massenproduktion ausgelegt. Individualisierte Produkte oder durch den Markt bedingte Änderungen im Portfolio führen zu teuren Umrüstungen und Umbauten. Im Zuge der Digitalisierung und im Sinne einer Industrie 4.0 wird die individualisierte Produktion bei gleicher Effizienz und Qualität wie bei einer Massenproduktion angestrebt. Dabei soll sich die Produktion automatisch ändernden Rahmenbedingungen bzw. Produktionszielen anpassen. Die Fraunhofer-



Tabelle 1

### Autonomiegrade von vier Anwendungsdomänen

Autonomiegrad/ Stufe	Automobil	Bahntechnik	Luftfahrt (UAV)	Produktionstechnik
0	SAE 0 – keine Automatisierung	GoA 0 – „Fahrt auf Sicht“		Datenerfassung und -verarbeitung
1	SAE 1 – Fahrerassistenten			Assistenzsysteme
2	SAE 2 – teilweise automatisiert	GoA 1 – Automatische Zug- sicherung mit Zugführer (ATP)	ALFUS1 – gering angeleite	Vernetzung und Integration
3	SAE 3 – bedingte Auto- matisierung	GoA 2 – Automatische Zug- sicherung und Betrieb mit Zug- führer (ATP+ATO)	ALFUS2 – Echtzeit Gesundheit/ Diagnose	Dezentralisierung, Dienstorientierung und Transforma- tionsfähigkeit
4	SAE 4 – Hochautomation	GoA 3 – Führerloser Betrieb (ATO)	ALFUS3 – Anpassung bei Aus- fällen und Fehlern/ Flugbedingungen	
5	SAE 5 – Vollständige Auto- matisierung (Autonomie)	GoA 4 – Führer- und begleitloser Betrieb (ATC)	ALFUS4 – bordseitige Routen- umplanung	Selbstorganisie- rung und Autonomie
6			ALFUS5 – Gruppengleichstellung ... ALFUS10 – Autonomes Schwarm- verhalten in jeder Umgebung	

Gesellschaft schlägt fünf Evolutionsschritte vor, die diese Entwicklung begleiten (vgl. [12]): zunächst die Betriebsdatenerfassung und -verarbeitung als Informationsgrundlage gefolgt von Assistenzsystemen für Entscheider und Werker. Auf der dritten Stufe soll die datentechnische Vernetzung und Integration von Produktionsschritten die Voraussetzung für eine Optimierung des Gesamtsystems liefern. Zur Erhöhung von Flexibilität wird auf der vierten Stufe die Fähigkeit des Produktionssystems zur Transformation bzw. Rekonfiguration gefordert. Schließlich soll sich das Produktionssystem auto-

nom selbst organisieren können (Stufe 5). Heutige Produktionssysteme sind auf den Stufen 1 (Betriebsdatenerfassung) bis 3 (vernetzte Produktion z. B. in der Automobilproduktion) anzusiedeln. Die Erreichung der nächsten Stufe erfordert in der Regel eine Restrukturierung der Architektur und Anpassung der Schnittstellen – und ist damit entsprechend aufwändig.

Die Autonomiegrade der vier Anwendungsdomänen sind in Tab. 1 noch einmal gegenübergestellt, dabei wurde versucht, ähnliche Autonomiegrade auf einer Ebene darzustellen.



Anhand der vier Beispiele kann man bereits viele Gemeinsamkeiten der Klassifikationen in den Stufen und in den Zielen erkennen. Eine generische Klassifikation, welche die unterschiedlichen Aspekte gut zusammenfasst, wurde im Rahmen einer technischen Roadmap hochautomatisierter Systeme von SafeTRANS [4] erarbeitet und veröffentlicht. Darin werden die wesentlichen Aspekte der Automatisierung in vier Systemklassen zusammengefasst:

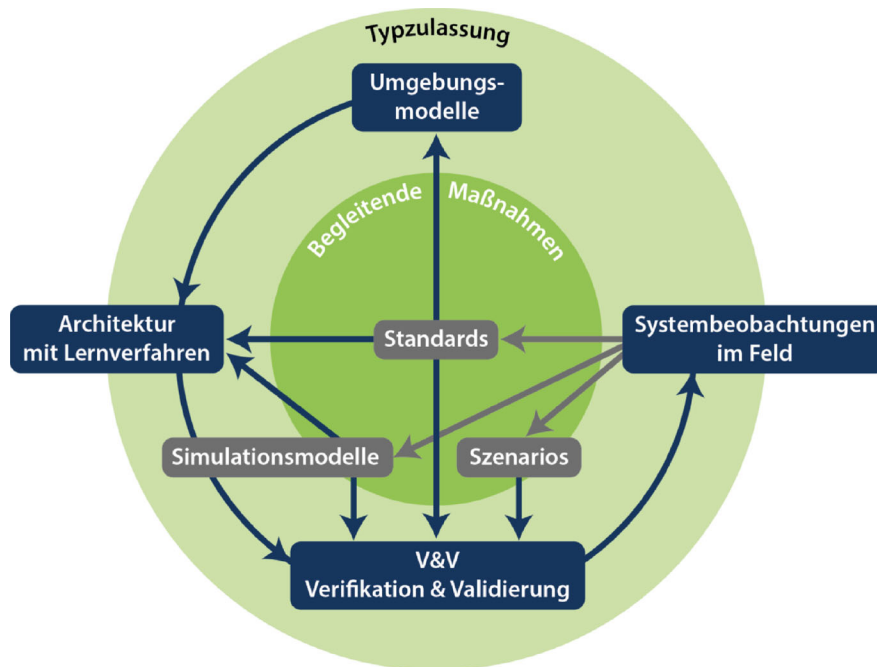
1. **Funktionale automatisierte Systeme** können begrenzte, klar definierte Aufgaben autonom erfüllen, wie z. B. automatisches Einparken, automatisches Landen oder die vollautomatische Produktion eines bestimmten Produkts. Diese Systeme können während des Betriebs nicht lernen; die Kooperation mit anderen Systemen ist auf den Austausch von Kontextinformationen beschränkt, d. h. es findet keine Kollaboration statt.
2. **Missionsorientierte Systeme** haben die Aufgabe, situationsabhängig eine ungeplante Kette beherrschbarer und bekannter Situationen zu durchlaufen, wie z. B. eine autonome Autobahnfahrt oder der automatische Zugbetrieb auf einer U-Bahn-Linie. Dabei können verschiedene Optimierungskriterien wie die Minimierung des Zeit- oder Ressourcenbedarfs in der Ausführung eine Rolle spielen. Planungs- und Optimierungsberechnungen werden vom System dynamisch zur Laufzeit durchgeführt. Auch diese Systeme können während des Betriebs nicht lernen; die Kooperation mit anderen Systemen ist auf den Austausch von Informationen über den Kontext und über das System selbst beschränkt.
3. **Kollaborative Systeme** sind Systeme wie Roboter, Fahrzeuge oder Schwärme von Flugzeugdrohnen, die z. B. einfädeln lassen oder die zur Kollisionsvermeidung miteinander kooperieren. Solche Systeme sind zur Erfüllung ihrer Mission in der Lage, mit anderen Systemen und Menschen zu kooperieren und ihre Wahrnehmungen, Interpretationen, Ziele, Pläne und Aktionen miteinander dynamisch abzustimmen. Die Systeme tauschen mit ihren Kooperationspartnern relevante Kontextinformationen aus, sind jedoch nicht lernfähig.
4. **Autopoietische Systeme** sind Systeme, die ihre Perzeption, ihre Interpretationen, ihre Aktionen

und ihre Kooperationsmöglichkeiten selbstständig erweitern und sich mit anderen Systemen darüber austauschen können (inklusive der Weitergabe von erlerntem Verhalten). Diese Systeme zeigen somit menschenähnliches Verhalten und sind in dieser Form heutzutage noch nicht realisiert. Die Fähigkeit des nichtüberwachten Lernens ist das wesentliche Charakteristikum dieser Systemklasse.

Wesentliches neues Element in Bezug auf die vorgenannten, bestehenden Klassifikationen ist das des Lernens bei autopoietischen Systemen. Diese Eigenschaft ist cyber-physischen Systemen heute nicht gegeben, sie ist in Bezug auf derzeitige normative Anforderungen an die funktionale Sicherheit auch nicht umsetzbar, da Sicherheit und Zuverlässigkeit nach einer Änderung kaum nachweisbar sind. Die jüngsten technologischen Durchbrüche im Bereich des Deep Learning und die hohe Qualität der Ergebnisse, z. B. bei der Bild- oder Mustererkennung, zeigen jedoch auf, dass diese Entwicklung und Möglichkeiten des maschinellen Lernens rasch voranschreiten. Allerdings gibt es hierbei noch große Hürden und viel Forschungsbedarf: Neuronale Netze können sich auch in unerwünschte Richtungen entwickeln bzw. aus Daten Muster extrahieren, die so nicht verallgemeinert werden dürfen [17]. Aktuelle Studien zeigen z. B., dass automatische Lernverfahren auch ethisch unerwünschte Muster aus historischen Daten „lernen“, wie z. B. geschlechtsspezifische Bevorzugung oder rassistisches Verhalten [7]. Eine entsprechende Überwachung nach ethischen und rechtlichen Maßstäben ist somit in jedem Fall notwendig.

Deshalb werden derzeit die neuronalen Netze nach der Lernphase nicht verändert. Das Problem des nichtdeterministischen Verhaltens aufgrund der großen Menge an möglichen Eingaben und der Komplexität der neuronalen Netze bleibt jedoch.

Speziell im asiatischen Raum wird die Rolle autonomer Systeme immer auch in Bezug auf den Menschen betrachtet. Die autonomen Systeme sollen Menschen unterstützen und entlasten. Der Mensch wird hierzu an bestimmten Stellen in die Regelschleife mit eingebunden (human-in-the-loop), sodass man auch von kooperativer Intelligenz sprechen kann, wie z. B. bei einer Mensch-Roboter-Interaktion bei der kooperativen Bearbeitung einer



**Abb. 1 Metamodell eines Lernprozesses für autonome Systeme**

Aufgabe. Solche zukünftigen autopoietischen Systeme teilen Erfahrungen zwischen Maschinen und Menschen und verhalten sich entsprechend. Spätestens an diesem Punkt werden ethische Fragestellungen relevant. Das Gebiet der Maschinenethik beschäftigt sich genau damit: mit dem Verhalten von Maschinen gegenüber Menschen. Aber auch einfachere autonome Systeme benötigen Schnittstellen zu Menschen. Entsprechende Benutzerschnittstellen müssen Nutzer verständlich informieren und ggf. situativ unterschiedlich Informationen und Dienste bereitstellen und dabei idealerweise das individuelle Verhalten von Mensch und Tier in ihren Handlungen mit einbeziehen.

### Anforderungen und Fähigkeiten aus Sicht der Informatik

Dieser Abschnitt gibt die zentralen Empfehlungen des SafeTRANS AK Hochautomatisierte Systeme zu Sicherheit, Testen und Entwicklung hochautonomer Systeme wieder [4]. Dabei wird zum einen die hohe Komplexität der durch das autonome System zu erkennenden Systemumgebung als zentrale Herausforderung gesehen.

1. Die hohe Komplexität der Umwelt macht die Durchführung einer für eine Zulassung ausreichend großen Anzahl von Feldtests für hochautonome Systeme impraktikabel. Daher

empfiehlt SafeTRANS ein System der kontinuierlichen Überwachung solcher Systeme zu implementieren sowie von Daten aus Feldbeobachtungen zu lernen. Abbildung 1 zeigt einen solchen Lernprozess auf der Metaebene, bei dem die Systeme im Feld beobachtet und diese Felddaten nach Beurteilung durch eine unabhängige Stelle die Basis für den Lernprozess sind. Diese Beurteilung liefert Richtlinien oder Empfehlungen für neue Features und/oder neue Funktionalitäten für den Entwicklungs- und Validierungsprozess mit dem zweifachen Ziel (1) der Verbesserung der Wahrnehmungsfähigkeiten des Systems und (2) der Sicherstellung von situationsangepasstem und sicherem Verhalten der Systeme. Ein solcher Lernprozess ermöglicht eine kontinuierliche Evolution autonomer Systeme auch dank der Möglichkeit der virtuellen Freigabe neuer Features und Funktionen im Rahmen eines modellzentrierten Entwicklungsprozesses. Zentrale Bausteine dieses Prozesses sind die Systemarchitektur und Algorithmen, Umweltmodelle, Verfahren der Verifikation und Validierung – sowie die systematische Erfassung von Betriebsdaten aus dem Feld. Diese grundlegenden Elemente müssen entsprechend standardisiert sein, benötigen eine gemeinsame offene Simulationsumgebung



## Handlungsempfehlungen

### Handlungsbereich

### Maßnahmen

1. Umweltmodelle	<ul style="list-style-type: none"> <li>I. Entwicklung eines offenen europäischen durch die Industrie getriebenen Standards für Umweltmodelle in den einzelnen Anwendungsfeldern, angepasst an die einzelnen Ausbaustufen und mit davon abhängigen Komplexitätsgraden.</li> <li>II. Aufbau eines durch die öffentliche Hand getriebenen Prozesses und entsprechender Infrastruktur zur Etablierung virtueller Systemvalidierung. Dazu nötig sind: <ul style="list-style-type: none"> <li>a. akkreditierte Einrichtungen</li> <li>b. eine öffentlich zugängliche Validierungsumgebung</li> <li>c. weitere Spezifikationen für Validierungen im Feld</li> </ul> </li> <li>III. Erstellung einer durch Zulassungsstellen und Gesellschaft akzeptierten Argumentationskette für den Sicherheitsnachweis hochautomatisierter Systeme, bestehend aus einer Kombination aus virtueller Freigabe und Brauchbarkeitstests im Feld.</li> </ul>
2. Lernende Community	<ul style="list-style-type: none"> <li>I. Aufbau eines durch die öffentliche Hand getriebenen Prozesses zum Lernen aus Feldbeobachtungen. Dazu sind nötig: <ul style="list-style-type: none"> <li>a. durch die öffentliche Hand akkreditierte Trust Center</li> <li>b. Selbstverpflichtung der Industrie, die dazu relevanten Daten an durch die Industrie akzeptierte Trust Center anonymisiert zur Verfügung zu stellen</li> <li>c. Rückführung der Analyseergebnisse der Trust Center in den Validierungsprozess</li> </ul> </li> </ul>
3. Architektur	<ul style="list-style-type: none"> <li>I. Eine durch die Industrie getriebene Standardisierung der Repräsentation der auszutauschenden Informationen zu Objekten und Situationen, um die Kooperation zwischen Systemen zu ermöglichen</li> <li>II. Eine durch die Industrie getriebene standardisierte funktionale Systemarchitektur für automatisierte Systeme und ihre Komponenten, die kompositionale Sicherheitsnachweise erlaubt und sichere Mindestfunktionalität in degradierten Modi unterstützt</li> <li>III. Ein öffentlich abgestimmter Entwicklungsprozess für hochautomatisierte Systeme, inklusive sicherer Upgradefähigkeit</li> <li>IV. Ein Industriegetriebener Standard für online Zertifizierung/Validierung der Kompatibilität von Upgrades mit der existierenden E/E-Architektur.</li> <li>V. Sichere, standardisierte Degradationsstufen mit garantierter Mindestfunktionalität.</li> </ul>
4. Absicherung der Interoperabilität autonomer Fahrzeuge	<ul style="list-style-type: none"> <li>I. International abgestimmte Klassifikation von Ausbaustufen der Architektur von hochautomatisierten Systemen und ihrer Interoperabilität</li> <li>II. Einführung von Zertifikaten für die Übereinstimmung von Architekturen mit dieser Klassifikation, die durch die öffentliche Hand benannte Stellen vergeben werden</li> <li>III. International abgestimmte Release-Prozesse für neue Ausbaustufen</li> </ul>
5. Framework	<ul style="list-style-type: none"> <li>I. Bereitstellung einer Plattform mit Basisdiensten für autonomes Fahren für die unterschiedlichen Ausbaustufen</li> <li>II. Etablierung eines anwendungsspezifischen Industriestandards für Frameworks, der von durch die öffentliche Hand benannten Stellen zertifiziert ist</li> <li>III. Bereitstellung von Representation Engines zur Aktualisierung der jeweils wahrgenommenen Umgebungssituation, der Darstellung möglicher Zukünfte sowie zur Ableitung von daraus resultierenden Handlungsempfehlungen</li> </ul>



und konkrete, akzeptierte (Test-)Szenarien für die Freigabe.

2. SafeTRANS gibt die folgenden Handlungsempfehlungen (s. Tab. 2).

## Literatur

1. acatech (Hrsg) (2017) Autonome Systeme – Chancen und Risiken für Wirtschaft, Wissenschaft und Gesellschaft, Abschlussbericht, März 2017
2. acatech (Hrsg) (2015) Neue autoMobilität, Automatisierter Straßenverkehr der Zukunft, Acatech POSITION, September 2015
3. Bundesministerium für Bildung und Forschung (BMBF) (2017) Fortschritt durch Forschung und Innovation, Bericht zur Umsetzung der Hightech-Strategie, März 2017
4. Damm W, Heidl P (Hrsg) (2017) Positionspapier SafeTRANS Arbeitskreis „Hochautomatisierte Systeme: Testen, Safety und Entwicklungsprozesse“, SafeTRANS e. V. [http://www.safetrans-de.org/de/Aktuelles/?we\\_objectID=2](http://www.safetrans-de.org/de/Aktuelles/?we_objectID=2), letzter Zugriff: 10.8.2017
5. Damm W, Sztipanovits J (Hrsg) (2016) cps summit Action Plan, Towards a Cross-Cutting Science of Cyber-Physical Systems for Mastering all-Important Engineering Challenges, 10.4.2016
6. Engell S (Hrsg) (2016) Proposal of a European research and innovation agenda on cyber-physical systems of systems 2016–2025, CPSoS Project, TU Dortmund
7. European Parliament resolution of 16.2.2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))
8. Fachforum Autonome Systeme im Hightech-Forum: Autonome Systeme – Chancen und Risiken für Wirtschaft, Wissenschaft und Gesellschaft. Langversion, Abschlussbericht, Berlin, April 2017
9. von Houten H (2017) CTO Philips, Presentation at DIF Amsterdam, Mai 2017
10. Larson J, Angwin J, Parris T Jr (2016) How Machines Learn to Be Racist by, ProPublica, 19.10.2016. <https://www.propublica.org/article/breaking-the-black-box-how-machines-learn-to-be-racist>, last access: 4.4.2017
11. National Institute of Standards and Technology (2004) Autonomy Levels for Unmanned Systems (ALFUS) Framework, Version 1.1. [http://ws680.nist.gov/publication/get\\_pdf.cfm?pub\\_id=823618](http://ws680.nist.gov/publication/get_pdf.cfm?pub_id=823618)
12. Neugebauer R, Hippmann S, Leis M, Landherr M (2016) Industrie 4.0 – From the perspective of applied research. In: 49th CIRP Conference on Manufacturing Systems (CIRP-CMS 2016). <http://www.sciencedirect.com/science/article/pii/S2212827116311556>
13. Platforms4CPS (2017) D1.1 Market Segmentation for CPS Technology, Version 0.5
14. SAE standard J3016 (2014) Automated Driving. SAE international
15. UITP (Union Internationale des Transports Publics) (2012) Metro Automation, Facts, Figures, Trends
16. Wahlster W, Kirchner F (2015) Autonome Systeme – Technisch-wissenschaftliche Herausforderungen und Anwendungspotentiale. DFKI Documents, Issue D-15-04
17. Wolfangel E (2017) „Dunkle Algorithmen“. Süddeutsche Zeitung vom 25.3.2017, S. 35