

Autonomie im Internet der Dinge: Anforderungen an die Gestaltung autonomer Agenten in Cyber-physischen Systemen

Christian Janiesch¹ und Valentin Nentwich²

¹ Julius-Maximilians-Universität Würzburg, Würzburg, Germany
christian.janiesch@uni-wuerzburg.de

² Audi Planung GmbH, Neckarsulm, Germany
valentin.nentwich@audi-planung.de

Abstract. Durch die Digitalisierung in Gesellschaft und Industrie findet eine wachsende Autonomisierung bzw. Automatisierung der Wertschöpfungs-systeme über Sensoren und Effektoren, sogenannte Cyber-physische Systeme, statt. Diese Automatisierung erlaubt eine schnellere und häufig auch fehlerfreie Arbeit in vernetzten Systemen, bspw. im Bereich Industrie 4.0. Es ist derzeit allerdings weder möglich noch sinnvoll jedweder Maschine volle Autonomie zu gewähren, so dass es nötig ist, genau zu definieren, in welchem Rahmen und mit welchen Zielen eine Maschine selbstbestimmt handeln darf. Allerdings gibt es für die Gestaltung von autonomen Maschinen im CPS derzeit keinen allgemein anerkannten Ansatz oder Methodik, wie solch ein Autonomie-Konzept umgesetzt werden könnte. Basierend auf einer Literaturanalyse der IT-Autonomie-Forschung werden 12 Anforderungen für die Umsetzung von autonomen Agenten in CPS erarbeitet. Es werden die Eigenschaften eines Agenten sowie dessen Interaktion mit dem Menschen über Autonomiestufen detailliert aufgearbeitet. Das resultierende Meta-Modell kann als Basis für die Gestaltung von Agenten, bspw. über eine Modellierungssprache, dienen.

Keywords: *Internet der Dinge, Cyber-physische Systeme, IT-Autonomie, Modellierung, Agent*

1 Einleitung

Digitalisierung ist ein Leittrend in Gesellschaft und Industrie, der eine Kernherausforderung für die Wirtschaftsinformatik darstellt. In der Produktion wird in diesem Zusammenhang von Industrie 4.0 als dem Organisationsprinzip der Zukunft gesprochen. Allgemeiner spricht man von selbststeuernden Objekten, die Internet-ähnliche, vernetzte Strukturen zur Kommunikation nutzen, dem sogenannten Internet der Dinge [1]. So soll letztendlich eine wachsende Autonomisierung bzw. auch Automatisierung der Wertschöpfungssysteme über Sensoren und Effektoren, in sogenannten Cyber-physischen Systemen (CPS), stattfinden.

Das starke Wachstum an Sensoren und deren Vernetzung fördert dabei die Erzeugung und Verbreitung von Daten und wird folglich aufgrund der Menge und der Geschwindigkeit, in der Daten heute entstehen, zu einer großen Herausforderung für Entscheidungsträger. Diese sind letztendlich durch die Vielzahl an Informationen und die Beschränktheit ihrer kognitiven Fähigkeiten nicht in der Lage, adäquate und zeitnahe Entscheidungen für alle durchgeführten Analysen zu treffen, ohne dabei Optimierungspotentiale zu verschenken und ohne Fehler zu machen [2].

Ein möglicher Lösungsansatz besteht darin, den Grad der Autonomie von Maschinen zu erhöhen. Mit anderen Worten ausgedrückt bedeutet dies, eine Maschine zu befähigen Entscheidungen als selbstständiger Agent zu treffen und dadurch die Kontrolle über Situationen und Handlungen zu übernehmen. Allerdings gibt es für die Gestaltung von autonomen Maschinen im CPS derzeit keinen allgemein anerkannten Ansatz oder keine Methodik, wie ein Autonomie-Konzept umgesetzt werden könnte.

Der Schwerpunkt dieses Artikels liegt auf der Systematisierung von Forschungsansätzen der IT-Autonomie. Basierend auf einer Literaturanalyse wird ein Konzept für die Umsetzung von Autonomiestufen und autonomen Agenten in CPS erarbeitet und die Basis für deren Spezifikation, bspw. über eine grafische Modellierungssprache, vorbereitet, welche die Modellierung von Autonomie ermöglichen würde. Das bereitgestellte Konzept schließt die Lücke zwischen der abstrakten Idee von autonomen CPS und der praktischen Gestaltung solcher Systeme in der Realität.

Im Folgenden werden zunächst Grundlagen zum Internet der Dinge und CPS vorgestellt. Darauf folgend werden die Kernaspekte hinsichtlich IT-Autonomie aufgezeigt. Diese allgemeingültigen Erkenntnisse bilden das Fundament der abschließend vorgestellten Grundlagen für eine Spezifikation des autonomen Verhaltens von Agenten. Sie gliedern sich zunächst in mehrere Anforderungen und werden anschließend in einem Meta-Modell umgesetzt und diskutiert. Das Papier schließt mit einem Ausblick auf weitere notwendige Arbeiten zur Gestaltung der Autonomie in CPS.

2 Grundlagen

2.1 Cyber-physische Systeme und Internet der Dinge

In der Literatur finden sich zahlreiche Erläuterungen des Begriffs Cyber-physisches System, eine anerkannte Definition liefern diese jedoch nicht. [3, 4] prägten zuerst den

Begriff des Internet der Dinge für Computer, die nicht nur eigenständig Informationen sammeln, sondern dazu befähigt werden, aus den gesammelten Informationen zu lernen und die Welt durch ihr stetig kumuliertes Wissen selbst wahrzunehmen und über ein Netzwerk, bspw. das Internet, kommunizieren zu können.

[5] differenziert darauf aufbauend CPS als physische, biologische und/ oder bautechnische Systeme, die über eine Recheneinheit integriert, überwacht oder gesteuert werden. Die Recheneinheit ist dabei ein verteiltes, eingebettetes System. [6] hebt zudem die Interdependenzen von Recheneinheiten und physischen Prozessen in Feedbackschleifen hervor. [7] ergänzen die Notwendigkeit einer oder mehrerer multi-modaler Mensch-Maschine-Schnittstellen zur Kommunikation und Steuerung von CPS und die weltweite Nutzbarkeit von Daten und Diensten für lokale Operationen.

Zusammenfassend bezeichnen CPS also die Kopplung von informations- und softwaretechnischen Komponenten mit mechanischen bzw. elektronischen Komponenten, die über eine Kommunikationsinfrastruktur, wie zum Beispiel das Internet, in Echtzeit miteinander kommunizieren. Diese Komponenten können in unterschiedlichem Grade automatisiert und autonom handeln, woraus Konsequenzen für die Mensch-Maschine-Interaktion entstehen [2].

2.2 (IT-)Autonomie

Der Begriff Autonomie besitzt im Allgemeinen die zwei Bedeutungen Unabhängigkeit und Selbstständigkeit. Selbstständigkeit beschreibt die Fähigkeit einer Entität, bzw. eines Agenten, für sich selbst zu sorgen, während Unabhängigkeit für die Freiheit einer Entität vor der Kontrolle einer von außen einwirkenden Instanz steht [8].

Im Kontext der IT kann Autonomie als eine nichtfunktionale Systemeigenschaft gesehen werden, die sich auf einzelne Funktionen übertragen lässt. Autonomie wird hierbei als Spezialisierung der Adaptivität betrachtet, der Fähigkeit zu verstehen, eine Zielstellung in Bezug auf eine Eingabemenge zu erfüllen und dabei die eigenen Strukturen an die Änderungen in der Umwelt anzupassen. Es wird eine Bindung von Autonomie an eine Funktion erreicht, indem die autonome Funktionserfüllung als Basiseigenschaft aufgefasst wird, mehrere Systeme miteinander in Bezug zu setzen [9].

Autonomie steht in einem engen Zusammenhang mit Automatisierung, welche eine maschinelle Handlung ohne menschliche Unterstützung darstellt. Es ist demnach ein Mittel für die Ablösung von Aufgaben eines Menschen [10]. Durch Automatisierung kann eine höhere Geschwindigkeit sowie eine bessere Effizienz bei der Aufgabenbewältigung erreicht werden. Weiterhin kann der physische und geistige Aufwand für einen Menschen reduziert werden.

IT-Autonomie kann den Menschen insbesondere bei Koordinationsaufgaben entlasten und so die Geschwindigkeit oder Sicherheit erhöhen. Sie lässt sich nicht nur gänzlich auf einzelne funktionale Aspekte eines Agenten übertragen, sondern durch eine graduelle Abstufung unterschiedlicher Ebenen der Interaktion zwischen Menschen und Maschinen kennzeichnen [11]. Diese Abstufung kann als Kontinuum zwischen der vollständigen Autonomie, bei der Maschinen Entscheidungen selbstständig gemäß ihrer eigenen Zielstellungen treffen, und dem größten kognitiven sowie physischen, menschlichen Einsatz auf der anderen Seite verstanden werden [2].

Wir verstehen unter Autonomie keine reine Systemeigenschaft, welche sich anhand des Informationsgradienten und des emergenten Verhaltens feststellen lässt [9]. Vielmehr rücken Aspekte der Interaktion zwischen einem Menschen und einer Maschine sowie der dadurch festgelegte Grad vorhandener Autonomie in einem System in den Vordergrund (siehe Kapitel 3.3). Abgedeckt wird hierbei das Spektrum vollständiger Kontrolle der Zielerreichung durch sowohl den Menschen als auch der Maschinen sowie das Kontinuum an Möglichkeiten, welches zwischen beiden Extremen liegt.

3 Autonomie in Cyber-physischen Systemen

3.1 Eigenschaften autonomer Agenten

Agenten sind eine Entität, wie bspw. Maschinen aber auch der Mensch, die ihre Umwelt über Sensoren wahrnehmen und mittels Effektoren auf diese Umwelt einwirken können [12]. Autonome Agenten handeln dabei entsprechend ihrer eigenen Agenda [13]. Die wichtigsten Bestandteile eines Agenten sind folglich Sensoren für die Sinneswahrnehmung und Effektoren für das Ausführen von Aktionen auf eine Umwelt. Eine Umwelt bezeichnet dabei den Wirkungskreis eines Agenten. Nach dem Verständnis von [14] stellt eine Umwelt ein System dar, welches aus einer Vielzahl von Agenten besteht. Hierbei kann ein System selbst wiederum ein Agent sein, welcher ein Teil eines größeren Systems respektive Agenten ist.

Beispiele für Sensoren menschlicher Agenten sind deren Augen und Ohren. Als ihre Effektoren werden häufig Hände und Mund genannt. Für künstliche Agenten zählen bspw. Kameras, Temperaturfühler oder digitale Schnittstellen zu Sensoren. Effektoren können Teile der Robotik aber, auch akustische Signale oder elektronisch übertragene Nachrichten sein.

Die Implementierung der Logik, Zielstellung und Wertvorstellung in Bezug auf die Informationsverarbeitung eines Agenten lassen [12] offen. Sie unterscheiden für die konkrete Definition genannter Eigenschaften folgende vier Typen von Agenten:

- **Agent mit Reflexen:** Agiert entsprechend einer vorgegebenen Menge von Regeln. Befindet sich die wahrgenommene Welt in einem definierten Zustand, wird die dem Zustand zugeordnete Aktion ausgeführt.
- **Agent mit internen Zuständen:** Erweitert einen einfachen Agenten mit Reflexen, so dass nicht nur der aktuelle Umweltzustand, sondern auch die Vergangenheit und die verfolgte Entwicklung des Umweltzustandes relevant für die auszuwählende Aktion des Agenten sind.
- **Agent mit Zielen:** Erweitert einen Agenten mit internen Zuständen um die Fähigkeit anhand von Zielstellungen den Einfluss von Aktionen auf seine Umwelt abzuwägen. Eine Zielvorgabe beschreibt den Zustand der Umwelt, der durch einen Agenten erreicht werden soll. Hierbei wird nicht genauer spezifiziert, wie der Zielzustand erreicht werden kann [15].
- **Agent mit Nutzenfunktion:** Erweitert einen Agenten mit Zielen um die Fähigkeit, den Nutzen für seine möglichen Handlungsalternativen festzustellen und die entsprechend der Nutzenfunktion beste Alternative zu wählen.

Durch das Hinzufügen des Aspektes der Autonomie erhalten Agenten die Fähigkeit, selbstständig zu entscheiden, wann sie welche Aufgaben wie ausführen und wann sie mit wem, auf welche Art und Weise, zu diesem Zweck kommunizieren. Hierdurch sind Agenten fähig, auf vorhergesehene und unvorhergesehene Ereignisse aus ihrer Umwelt zu reagieren, erwartete und unerwartete Probleme zu lösen und Teil einer Organisation mit dynamischen Strukturen zu sein [15, 16]. [13] geben eine Übersicht von Eigenschaften, welche für die Klassifizierung von Agenten (siehe Tabelle 1) herangezogen werden kann. Ihrer Auffassung zufolge erfüllen autonome Agenten zumindest die ersten vier Eigenschaften dieser Übersicht: Reaktivität, Autonomie, Ziel-Orientierung und Kontinuität. Reaktivität in Form einer zeitnahen Reaktion auf asynchrone, externe Reize ist eine Schlüsseleigenschaft eines autonomen Agenten. Deshalb sollte die Informationsverarbeitung in Echtzeit erfolgen [17].

Tabelle 1. Eigenschaften zur Klassifizierung von Agenten (in Anlehnung an [13])

<i>Eigenschaft</i>	<i>Beschreibung</i>
Reaktiv	Reagiert zeitnahe auf Veränderungen in der Umwelt
Autonom	Übt Kontrolle über seine eigenen Aktivitäten aus
Ziel-orientiert	Keine einfache Reaktion auf die Umwelt
Temporär	Kontinuierlich laufender Prozess
Kommunikativ	Kommuniziert mit anderen Agenten, eventuell Menschen eingeschlossen
Lernend	Verändert sein Verhalten basierend auf Erfahrungen
Mobil	Fähigkeit, sich selbstständig auf eine andere Maschine zu transportieren
Flexibel	Aktionen folgen keinem Skript
Charakter	Glaubhafte „Persönlichkeit“ und emotionaler Zustand

[18] heben ergänzend die sozialen Fähigkeiten insbesondere die Kommunikation mit anderen künstlichen oder menschlichen Agenten, als zentrale Eigenschaft hervor. [19] untersuchen soziale Interessen von Agenten in Bezug auf die Entscheidungsfindung. Sie unterscheiden zwischen folgenden drei Formen sozialer Kompetenzen:

- **Eigennutz:** Ein Agent stellt die eigene Zielstellung respektive seine eigenen Vorteile über den globalen Nutzen der Gruppe.
- **Hilfsbereitschaft:** Ein Agent wählt Alternativen, welche aus Sicht der Gruppe vorteilhaft sind. Hierbei sind auch Alternativen möglich, die keinen individuellen Vorteil für den einzelnen Agenten bringen.
- **Kooperation:** Agenten wählen Alternativen, welche einen positiven Nutzen für die Gruppe kooperierender Agenten bringen. Zwar ist eine Wahl von Alternativen möglich, die einen individuellen Nachteil des Agenten nach sich ziehen, allerdings wird dieser Nachteil von einem Vorteil übertroffen, der durch eine in einem Team ausgeführte Aktion erzielt wird.

Vorangehende Ausführungen bezüglich der Konzepte zu Autonomie und zu autonomen Agenten zeigen, dass Informationsverarbeitung auf der einen Seite und Interaktivität auf der anderen Seite ausschlaggebend für die Bestimmung des Autonomielevels von Maschinen sind. Vergleicht man die Menge der innerhalb der Grenzen eines Sys-

tems verarbeiteten Informationen mit der über die Systemgrenzen hinweg ausgetauschten Informationsmenge, lässt sich der Informationsgradient als Indikator des Autonomiegrades eines Systems bestimmen [9].

Bezüglich der Informationsverarbeitung stützen sich [20, 21] auf die vier Funktionen Überwachung, Erzeugung, Auswahl und Implementierung. Diese können wie folgt auf den Prozess der Informationsverarbeitung übertragen werden: Systemstatus wahrnehmen, mögliche Reaktionen ableiten, Aktionen auswählen und Aktionen umsetzen. Während [20] den Grad der Autonomie über die Zuordnung der Grundfunktionen zu Menschen und Maschinen festlegen, beschreiben [21] eine Divergenz des Autonomiegrades innerhalb der einzelnen Funktionen selbst.

[22] liefern eine Taxonomie acht möglicher Kommunikationsmuster zwischen Menschen und Maschinen, im Speziellen Robotern.

- **1 → 1:** Der Mensch nimmt eine steuernde Rolle ein, während die Maschine einerseits Informationen aus der Umwelt wahrnimmt und an den Menschen weiterleitet und andererseits die durch den Menschen vorgegebenen Aktionen ausführt.
- **1 → 1+1+...:** Zwei oder mehr Maschinen werden von einem Menschen gesteuert. Der Mensch ist für die Aufteilung der Aufgabe und die Koordination der Ausführung zuständig. Die Maschinen arbeiten unabhängig voneinander ihre Teilaufgaben ab.
- **1 → n:** Zwei oder mehr Maschinen werden von einem Menschen gesteuert. Die Maschinen nehmen die Anweisung durch den Menschen als Gruppe wahr und koordinieren die Unterteilung der Aufgabe sowie deren Ausführung selbst.
- **n → 1:** Zwei oder mehr Menschen erarbeiten in einem Team Aktionen für eine Maschine und übermitteln diese nach gemeinsamer Abstimmung an eine Maschine zur Ausführung.
- **n → 1+1+...:** Zwei oder mehr Menschen erarbeiten in einem Team Aktionen für zwei oder mehr Maschinen. Die Menschen sind für die Aufteilung der Aufgabe und die Koordination der Ausführung zuständig. Die Maschinen arbeiten dann unabhängig voneinander ihre Teilaufgaben ab.
- **n → n:** Zwei oder mehr Menschen erarbeiten in einem Team Aktionen für zwei oder mehr Maschinen und übermitteln diese nach gemeinsamer Abstimmung an die Maschinen zur Ausführung. Die Maschinen nehmen die Anweisung durch den Menschen als Gruppe wahr und koordinieren die Unterteilung der Aufgabe sowie deren Ausführung selbst.
- **1+1+... → 1:** Zwei oder mehr Menschen übermitteln unabhängig voneinander Aktionen an eine Maschine. Die Maschine übernimmt die Koordination eingehender Aufgaben.
- **1+1+... → n:** Zwei oder mehr Menschen übermitteln unabhängig voneinander Aktionen an eine Gruppe von Maschinen. Die Maschinen entscheiden innerhalb ihrer Gruppe über die Zuweisung der Aufgaben an einzelne Gruppenmitglieder sowie über die Prioritäten bei der Aufgabenbearbeitung.

Die Fähigkeit der Gruppenbildung weist Parallelen zu komplexen adaptiven Systemen auf [23], in denen Agenten miteinander interagieren, aber auch lernen und sich anpassen. Eine weitere sich daraus ergebende Eigenschaft ist die Selbstähnlichkeit.

Zur Übertragung von Informationen ist ein Kommunikationsmedium erforderlich. [24] beschreibt Varianten der Kommunikation eines Menschen mit einer Maschine und vice versa. Diese orientieren sich an den klassischen fünf Sinnen Sehen, Hören, Riechen, Fühlen und Schmecken.

[25] bezeichnen Vertrauen als Basis der Verlässlichkeit bei der Kommunikation. Vertrauen ist immer dann ein notwendiger Faktor, wenn eine Bewertung von Handlungsalternativen aufgrund von Unsicherheiten und hoher Komplexität nicht möglich ist. Der Grad des Vertrauens dient dann als ein Indikator dafür, ob eine Entscheidung akzeptiert oder selbst getroffen wird. Attribute für die Vertrauenswürdigkeit von Kommunikation sind bspw. Aufmerksamkeit, Verlässlichkeit, Konformität und Sicherheit.

3.2 Anforderungen an autonome Agenten

Agenten agieren in einer Umwelt. Sie müssen Reize mittels Sensoren wahrnehmen, Aktionen wählen und diese Aktionen durch den Einsatz eigener Effektoren an ihre Umwelt zurückgeben. Die Umwelt kann dabei als holonisches Multi-Agenten-System betrachtet werden. Das bedeutet, dass ein Agent aus der Vogelperspektive beobachtet wie ein einzelnes Teil mit sichtbaren Sensoren und Effektoren wirkt. Im Querschnitt dieses Agenten wird deutlich, dass sich der Aufbau des Hauptagenten aus einer Vielzahl von Subagenten zusammensetzt. Es besteht eine rekursive Beziehung zwischen Agenten, die eine beliebige Granularität zulässt.

Anforderung 1: Die Spezifikation muss hierarchische Anordnungen von autonomen Agenten abbilden.

Ein Agent besitzt beliebig viele Sensoren und beliebig viele Effektoren. Mit Effektoren gibt ein Agent Informationen eines bestimmten Typs ab. Der Agent selbst oder andere Agenten in seiner Reichweite nehmen diese Informationen mit den jeweils entsprechenden Sensoren wahr.

Anforderung 2: Die Spezifikation muss abbilden, welche Arten von Sensoren und Effektoren von Agenten unterstützt werden und welche Sensoren die Datenobjekte welcher Effektoren wahrnehmen können.

Die Richtung einer Interaktion bestimmt den Datenfluss zwischen den Teilnehmern an einer Interaktion. Hierdurch wird auch die Abhängigkeit zwischen Agenten sowie die Gesamtstruktur eines Netzwerkes aus Agenten ausgedrückt.

Anforderung 3: Die Spezifikation muss die Richtungsdefinition einer Interaktion abbilden, so dass deren kommunikatives Verhalten sichtbar gemacht wird.

Ein Agent ist genau dann autonom, wenn dieser Aufgaben selbstständig ausführen kann. Hierzu zählt das Treffen relevanter Entscheidungen bzgl. Ressourcen und Ausführungszeitpunkten.

Anforderung 4: Die Spezifikation muss einfache und komplexe Regeln, welche das Handeln eines Agenten beschreiben, abbilden.

Ergänzend müssen die Entscheidungen auf Basis des internen und externen Zustandes getroffen werden.

Anforderung 5: Die Spezifikation muss verschiedene Zielstellungen, Nutzenfunktionen und Zustände abbilden.

Die Anforderung, in Echtzeit auf Stimuli zu reagieren, zeigt insbesondere bei Menschen, dass die Auswertung von Sensordaten nur in einem beschränkten Maße möglich

ist. Die als Aufmerksamkeitskapazität bezeichnete Fähigkeit beschreibt sowohl eine Beschränkung des Kurzzeitgedächtnisses als auch der Phasen Wahrnehmung, Entscheidungs- und Aktionsauswahl sowie Aktionsdurchführung. Aus ökonomischen oder technologischen Gründen kann eine vergleichbare Beschränkung auch auf Maschinen zutreffen.

Anforderung 6: Die Spezifikation muss eine kapazitive Beschränkung von Funktionalitäten durch ein Konstrukt ähnlich der Aufmerksamkeit abbilden.

Unter dem Aspekt der Gruppierung respektive der Teambildung lassen sich mehrere Agenten zu einer Gruppe zusammenfassen. Beispielsweise übergibt ein Mensch eine Aufgabe an eine Gruppe von Maschinen, welche die Arbeitsaufteilung und Aufgabenbewältigung selbst koordiniert. Wie beschrieben können Agenten auch aus beliebig vielen Subagenten aufgebaut und durch Zielstellungen, Nutzenfunktionen, Regeln und Zustände spezifiziert werden. Diese müssen nicht konfliktfrei sein.

Anforderung 7: Die Spezifikation muss den Aspekt der sog. sozialen Rationalität abbilden und Formen sozialer Kompetenzen darstellen sowie eine Möglichkeit abbilden, Agenten unterschiedlichen Typs zu Gruppen zusammenzufassen. Das gilt auch für Agenten desselben Typs, wenn deren Kommunikationsverhalten eindeutig bestimmt ist. Dies erfordert die Bildung von Agentenklassen, bei denen die Anzahl möglicher, konkreter Instanzen individuell festgelegt werden kann.

Sicherheit ist einerseits ein Mechanismus, der Korruption bei der Nachrichtenübertragung verhindert und somit bewusste Manipulationen und Angriffe von Dritten auf Datenströme erschwert. Andererseits wird dem Sicherheitsaspekt die Eigenschaft der Zuverlässigkeit bei der Nachrichtenübermittlung zugeordnet.

Anforderung 8: Die Spezifikation muss Vertrauens-Parameter einer Beziehung respektive eines Datenübertragungskanal, wie Aufmerksamkeit, Verlässlichkeit, Konformität und Sicherheit, abbilden.

Die Beispiele zu den allgemeinen Koordinationsprotokollen sowie die Ansätze aus dem Bereich des Spezialfalls der Supervisory Control [26] zeigen, dass strukturelle Vorgaben für die Koordination von Agenten über deren Zustände und Regeln abgebildet werden können. Folglich entstehen aus dem Bereich der Koordination keine gesonderten Anforderungen.

3.3 Anforderungen an die Interaktion autonomer Agenten

Es gibt mehrere Konzepte die sich mit der Einteilung von Autonomie in unterschiedliche Stufen bei der Interaktion zwischen Mensch und Maschine beschäftigen [2, 20, 21, 26-30]. Tabelle 2 fasst die verschiedenen Konzepte zusammen und bildet die entsprechenden Stufen ab. Eine detaillierte Besprechung der verwandten Arbeiten findet sich in Kapitel 5. Eine eindeutige stufenartige Staffelung, welche sich zwischen den beiden Extremen der vollständigen Autonomie einer Maschine S (Stufe 1) oder eines Menschen M (Stufe 19) bewegt, ist nicht möglich.

Die beiden Extremfälle auf Stufe 1 und 19 beschreiben jeweils Situationen, in denen entweder ein Mensch oder eine Maschine die vollständige Kontrolle über die Entscheidungsfindung und die Aktionsdurchführung besitzen. Die bereits für autonome Agenten identifizierten Anforderungen genügen, um beide Sachverhalte adäquat abzubilden. Folglich entstehen aus den Randbereichen keine gesonderten Anforderungen.

Tabelle 2. Übersicht der Autonomiestufen

<i>Nr</i>	<i>Beschreibung</i>	<i>Quelle</i>
1	S trifft Entscheidung komplett und führt Aktionen autonom aus, keine Funktion durch M	[2, 20, 21, 26, 27, 30]
2	S trifft Entscheidung komplett und führt Aktionen autonom aus, M überwacht S und greift ggf. steuernd ein	[20, 27, 28]
3	S trifft Entscheidung komplett und führt Aktionen autonom aus, M überwacht bestimmte Bereiche von S und greift ggf. steuernd ein	[29]
4	M trifft Entscheidung und führt Aktion autonom aus bis zur Übergabe an S, dann führt S Aktionen autonom aus	[20, 26]
5	M trifft Entscheidung und übergibt Aktion zur Ausführung an S	[20]
6	S trifft Entscheidung und führt Aktion solange aus bis M ein Veto einlegt	[2]
7	S wählt Aktion und gibt M beschränkte Zeit für Veto	[21, 26, 30]
8	S wählt Aktion und führt nur auf Zustimmung von M aus	[20, 21, 26, 27, 30]
9	S bietet M alle Alternativen zur Auswahl an	[21]
10	S bietet M vorausgewählte Alternativen zur Auswahl an	[2, 21, 26, 27, 30]
11	S bietet M genau eine Alternative an, M kann dennoch ablehnen	[2, 21, 26, 27, 30]
12	S und M schlagen Alternativen vor, M wählt beste Alternative aus	[20]
13	M schlägt S Alternativen vor, S wählt die beste Alternative aus und führt diese aus	[20]
14	S führt Aktion aus und benachrichtigt M nach eigenem Ermessen	[21, 26]
15	S führt Aktion aus und benachrichtigt M auf Nachfrage von M	[21, 26, 30]
16	S führt Aktion aus und benachrichtigt M immer	[21, 30]
17	M trifft Entscheidungen und steuert S	[28]
18	M trifft Entscheidungen und steuert, aber S übernimmt Teile der Steuerung und führt diese autonom aus	[28]
19	M trifft Entscheidung komplett und führt Aktionen autonom aus, keine Funktion durch S	[20, 21, 27, 30]

Stufe 2 und 3 beschreiben eine Situation, in der einer der beiden Interaktionspartner Aktionen auswählt und ausführt, während der andere Partner die Rolle einer überwachenden Instanz einnimmt und in kritischen Situationen steuernden Einfluss auf die Ausführung nimmt. Die beiden Stufen unterscheiden sich im Grad der Überwachung, welcher sich über die Menge der überwachten Funktionen festlegen lässt. Beide Stufen entsprechen dem Konzept der Supervisory Control und den Anforderungen, die mit diesem Konzept verknüpft sind (siehe oben).

Eine weitere Gruppe bilden die 4. und 5. Stufe. Bei beiden liegt die Autorität einer Entscheidung bei einem Menschen. Beide Stufen unterscheiden sich durch den Grad der Aufteilung einer Aufgabe bei deren Ausführung. Durch das Konzept der autonomen Agenten wurde bereits eine Anforderung definiert, welche Datenobjekte als Teil des

hier auszuarbeitenden Meta-Modells beschreibt, die zwischen zwei oder mehreren Agenten ausgetauscht werden. Der Fokus dieser beiden Gruppen liegt auf dem Austausch von Aufgaben und Teilaufgaben.

Anforderung 9: Die Spezifikation muss Aufgabenlisten als speziellen Typ eines Datenobjektes abbilden.

Ein ähnlicher Fall ergibt sich aus den Stufen 9 bis 11, welche unterschiedliche Mengen an Alternativen, die einem Agenten von einem anderen Agenten vorgeschlagen werden, beschreiben: (1) alle Alternativen, (2) eine Auswahl an Alternativen sowie (3) genau eine Alternative. Stufe 12 und 13 beschreiben das gleiche Verhalten aus gegensätzlicher Perspektive.

Anforderung 10: Die Spezifikation muss Datenobjekte des Typs Aufgabenlisten entsprechend der beschriebenen Abstufung abbilden.

Stufe 6 bis 8 beschreiben die Situation, in der eine Maschine selbstständig Aktionen auswählen und ausführen kann. Die Stufen unterscheiden sich darin, dass ein Mensch unterschiedliche Möglichkeiten besitzt, durch sein Veto die Ausführung einer Aktion zu unterbrechen oder gar zu verhindern.

Anforderung 11: Die Spezifikation muss Datenobjekte zur Zustimmung und Ablehnung von Aufgaben abbilden.

Eine weitere Kategorie kann aus den Stufen 14 bis 16 abgeleitet werden. Diese beschreiben die Situation, dass ein Agent eine Aktion ausführt und dabei einen weiteren Agenten auf unterschiedliche Art über seinen aktuellen Zustand benachrichtigt: (1) nach eigenem Ermessen benachrichtigen, (2) immer benachrichtigen und (3) auf Anfrage benachrichtigen.

Anforderung 12: Die Spezifikation muss den Typ der Benachrichtigung abbilden.

Die verbleibenden Stufen 17 und 18 zeigen starke Ähnlichkeit zu den Stufen 2 und 3. Der Unterschied zwischen beiden Gruppen liegt in der aktiven beziehungsweise in der passiven Haltung des menschlichen Agenten. Auf der einen Seite nimmt dieser eine passive, beobachtende Rolle ein und wird nur in kritischen Situationen steuernd aktiv. Auf der anderen Seite ist der menschliche Agent durchgehend aktiv, trifft zu einem gewissen Grad Entscheidungen und steuert den zweiten Agenten. Auf abstrakter Ebene betrachtet, macht die Frequenz der Eingriffe den Unterschied zwischen den Stufen aus, welche sich über die Komplexität der zu steuernden Aufgaben ergibt. Folglich kann dieser Sachverhalt über das Formulieren von Regeln für die beteiligten Agenten abgebildet werden und stellt keine weiteren Anforderungen.

4 Meta-Modell zur Spezifikation autonomer Agenten

Der folgende Abschnitt definiert die abstrakten Bestandteile einer Sprache, ein Meta-Modell, das die Spezifikation von Autonomie erlauben soll. Abbildung 1 stellt den Aufbau eines Agenten in drei Kategorien dar: Bestandteile, kennzeichnende Elemente und beschreibende Elemente. Diese modellierten Informationen könnten dann für die Charakterisierung der Autonomie in CPS, bspw. von (Sub-)Prozessen im Bereich Business Process Management oder von Event Processing Agents bzw. Networks im Bereich

Complex Event Processing, verwendet werden, um zu signalisieren, wie autonom sich diese Prozesse bzw. Analyse-Netzwerke in Bezug auf ihre Umwelt verhalten sollen.

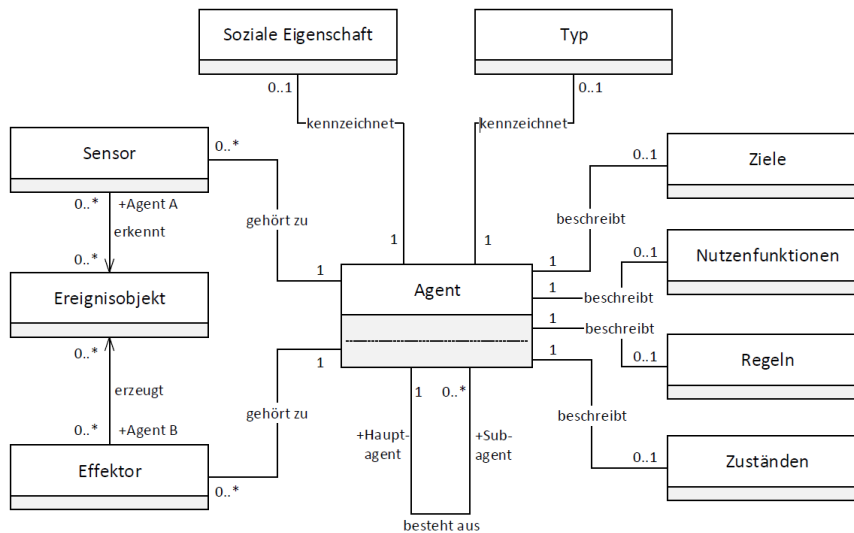


Abbildung 1. Übersicht Meta-Modell

Die Kategorie Bestandteile umfasst die Sensoren und Effektoren eines Agenten. Im traditionellen Sinne ist ein Sensor immer genau einem Agenten zugeordnet. Maschinen sind jedoch grundsätzlich in der Lage gemeinsam auf Sensoren und Effektoren zuzugreifen. Weiterhin kann ein Agent beliebig viele Sensoren zugeordnet bekommen. In der gleichen Beziehung steht ein Agent zu seinen Effektoren. In beiden Fällen sind auch die Extreme möglich, dass einem Agenten keine Sensoren oder keine Effektoren zugeordnet werden. Im Meta-Modell ist dies implizit über die Hierarchisierung von Agenten abgebildet.

Es gibt grundsätzlich sechs Kanäle über die Effektoren und Sensoren kommunizieren: visuell, auditiv, olfaktorisch, taktil, gustatorisch und generisch. Die Einteilung der ersten fünf Kanäle orientiert sich hierbei an den von Schmauk als klassische Modalitäten eingeordneten fünf Sinnen. Der generische Sensor dient als offener Typ, der über zusätzliche Parameter näher bestimmt werden kann und somit alle diejenigen Kanäle abbilden kann, welche nicht durch eine der anderen fünf Klassen spezifiziert werden können. Hierzu zählen bspw. elektronische Kommunikation von Maschinen oder Hirnströme bei Menschen.

Zwei oder mehrere Agenten können in einer Beziehung stehen, wenn sie Sensoren und Effektoren besitzen, die kompatibel zueinander sind, d. h. wenn Ereignisobjekte erzeugt werden, die von einem der Sensoren des anderen Agenten erkannt werden können. Eine Beziehung zwischen zwei Agenten kann durch optionale Parameter beschrieben werden. Diese beschreibenden Parameter können von Agenten bei deren Entscheidungsfindung herangezogen werden und können somit Einfluss auf die Entscheidung selbst haben. Ein Beziehungsparameter offenbart grundlegende Eigenschaften, die das

Verhältnis zwischen zwei Agenten betreffen. Beziehungsparameter spezialisieren sich zu Aufmerksamkeit, Verlässlichkeit, Konformität und Sicherheit.

Die Aufmerksamkeit bezeichnet wie oben eingeführt eine kapazitive Obergrenze an Informationen, die in den verschiedenen Stufen der Verarbeitung wahrgenommen werden kann. Wir nutzen dieses Konzept, um den von einem Effektor erzeugten Anteil der Ereignisobjekte einer Beziehung zu beschreiben, der von einem Sensor erkannt wird, oder um technische Limitationen bei der Auswertung zum Ausdruck zu bringen. Verlässlichkeit und Konformität leiten sich aus den Konzepten bezüglich des Vertrauens bei der Interaktion zwischen Mensch und Maschine ab. Der Unterschied zwischen den beiden Parametern liegt in ihrer Wirkungsrichtung. Verlässlichkeit beschreibt das Vertrauen eines wahrnehmenden Agenten gegenüber dem sendenden Agenten. Es ist also ein Indikator dafür, zu welchem Grad die Signale zur Entscheidungsfindung von einem wahrnehmenden Agenten herangezogen werden. Die Konformität wirkt entgegengesetzt zu der Verlässlichkeit und beschreibt somit das Vertrauen eines sendenden Agenten gegenüber dem empfangenden Agenten. Sicherheit beschreibt als letzter Aspekt die Vertrauenswürdigkeit des Kommunikationskanals.

Die zweite Kategorie enthält die kennzeichnenden Elemente eines Agenten. Hierzu zählen einerseits die sozialen Eigenschaften und andererseits die unterschiedlichen Typenausprägungen eines Agenten. Eine gesonderte Betrachtung der sozialen Eigenschaften zu den unterschiedlichen Agententypen ist notwendig, um soziale Eigenschaften als Indikator für das Auflösen von Konflikten zwischen den Zielen, Nutzenfunktionen, Regeln und Zuständen eines Hauptagenten zu seinen Subagenten bzw. Konflikten zwischen Agenten in einem Netzwerk zu nutzen und somit das spezifische, autonome Verhalten eines Agenten zu zeigen. Die Kennzeichnung durch den Typ eines Agenten gibt Aufschluss über das allgemeine Verhalten bei der Ereignisverarbeitung. Einem Agenten kann maximal eine soziale Eigenschaft zugeordnet werden. Die Assoziation zwischen einem Agenten und seinem Typ folgt der gleichen Kardinalität wie bei den sozialen Eigenschaften.

Eine dritte Kategorie bilden die vier beschreibenden Elemente Ziele, Nutzenfunktionen, Regeln und Zustände. Die Summe der Elemente respektive deren Zusammenspiel beschreiben das autonome Verhalten eines Agenten. Die Kardinalitäten der Assoziation eines Agenten mit den jeweiligen Elementen dieser Kategorie entspricht den bereits beschriebenen Kardinalitäten der kennzeichnenden Merkmale. Folglich sind auch diese Elemente für die Spezifikation optional und zeigen primär an, ob ein Agent durch eine Menge von Zielen, Nutzenfunktionen, Regeln und Zuständen beschrieben wird oder nicht.

Eine weitere Eigenschaft eines Agenten, die durch Abbildung 1 dargestellt ist, ist eine Assoziation zu sich selbst. Diese reflexive Assoziation zeigt an, dass sich ein Hauptagent in beliebig viele Subagenten zerlegen lässt.

5 Verwandte Arbeiten

[26] stellen eine zehnstufige Skala für die Unterscheidung verschiedener Grade der Automatisierung bei der Entscheidungsfindung vor. Zwischen den beiden Extremen der

vollständigen Entscheidungskompetenz liegen Stufen, welche durch das Bilden von Interaktionsmustern zwischen Mensch und Maschine differenziert werden.

Auch [21] beschreiben eine graduelle Abstufung von Autonomie in zehn Stufen. Jede dieser Stufen ist hierbei auf die einzelnen Phasen der menschlichen Informationsverarbeitung anwendbar. Neben der Stufe der vollständigen Autonomie werden die Stufen nach den drei Kategorien Art der Information, Vetorechte und Anzahl der angebotenen Handlungsalternativen differenziert. [30] reduzieren diese Skala auf acht Stufen. Die Autoren gehen hierbei speziell auf den Fall einer durch den Computer ausgewählten Handlungsalternative ein.

[31] beschreibt ein Modell, welches die grundlegenden Eigenschaften aus dem Konzept der Supervisory Control abbildet. Er beschreibt eine fünfstufige Skala für die Bestimmung des Automatisierungsgrades in Bezug auf den Prozess der Informationsverarbeitung. [29] identifiziert sieben Arten der Beeinflussung gegenseitiger Autonomie durch Veränderung von Systemparametern. Bisherige Ansätze werden dahingehend erweitert, dass autonomes Verhalten nicht nur von der eigenen Autorität über das Treffen von Entscheidungen abhängt, sondern auch von weiteren Entitäten im System beeinflusst und geprägt wird.

[32] schlägt 13 Eigenschaften eines Agenten vor, anhand derer die autonome Abhängigkeit zu seiner Umwelt festgestellt werden kann. [14] bauen auf diesen Eigenschaften der Dimensionierung von Autonomie auf und zeigen in ihren Untersuchungen Autonomie in Bezug auf Multi-Agenten-Organisationen.

[27] zeigt vier unterschiedliche Grade von Autonomie auf. [33] greifen diese Skala auf und vergeben für die autonome Abstufung die folgenden Bezeichnungen: keine Systemunterstützung, Entscheidungsunterstützung, konsensuelle und überwachte, künstliche Intelligenz sowie vollständige Autonomie. [20] erweitern die Skala zu einer Taxonomie mit zehn Stufen. Diese reichen von manueller Kontrolle bis zu vollständiger Autonomie.

[28] klassifiziert unterschiedliche Grade von Autonomie: Überwachung und logisches Denken, Planung des Fahrwegs, Kollisionsverhinderung und einfache Steuerung. Auch diese Abstufung reicht von der vollständigen Kontrolle des Menschen bis zur Überwachung von Zielvorgaben durch den Menschen.

[19] beschreiben in ihrem Ansatz das Verhalten von Agenten in Gruppen. Sie knüpfen Autonomie an die soziale Bereitschaft eines Agenten, welche über die Zielstellung eines Agenten an die Maximierung eines individuellen oder gesellschaftlichen Nutzens gebunden ist. [8] wählen ebenfalls einen Ansatz, welcher Autonomie unterschiedliche Dimensionen zuordnet. Sie beschreiben die Anpassung von Autonomie entlang gewählter Dimensionen durch das Einschränken oder das Erweitern von Berechtigungen, Fähigkeiten, Möglichkeiten und Verpflichtungen eines Partners, welcher sich in einer Beziehung zur Erfüllung gemeinsamer Aufgaben befindet.

[2] untersuchen wie sich die automatisierungsbedingte Leistungsfähigkeit von Menschen bei unterschiedlichen Automatisierungsstufen verändert. Die Autoren weisen auf einen Kosten-Nutzen-Kompromiss hin, der sich daraus ergibt, dass zwar Routine-Aufgaben bei korrekter Arbeitsweise auch bei veränderlicher Arbeitsbelastung schneller erledigt werden können, aber die Performance und das Situationsbewusstsein deutlich sinkt, wenn das Entscheidungsunterstützungssystem fehlerhaft arbeitet.

6 Zusammenfassung und Ausblick

Aufbauend auf einer umfangreichen Literaturstudie zur IT-Autonomie wurden 12 Anforderungen für die Spezifikation autonomer Agenten in CPS-Systemen bestimmt. Diese wurden übersichtlich in einem abstrakten Meta-Modell zusammengefasst, welches die Möglichkeit eröffnet, die Autonomie von Agenten eindeutig zu beschreiben. Es ist zunächst bewusst abstrakt gehalten und kann für verschiedene Komponenten eines Mensch-Maschine-Systems verwendet werden, wie bspw. ein CPS-System, das Menschen sowie Maschinen mit oder ohne physischer Präsenz (bspw. virtualisierte Prozess- oder Analyse-Systeme) enthält.

Die nächsten Schritte werden sich mit der Umsetzung des abstrakten Modells in eine konkrete Notation und die Verknüpfung derer mit existierenden Modellierungssprachen beschäftigen. Außerdem werden wir Entwurfsmuster für die Mensch-Maschinen-Kommunikation basierend auf den vorgestellten Autonomie-Levels entwickeln, um die Implikationen für den Kommunikationsablauf besser darstellen zu können.

Literaturverzeichnis

1. Kopetz, H.: Real-Time Systems: Design Principles for Distributed Embedded Applications. 2nd edn. Springer, New York, NY (2011)
2. Onnasch, L., Wickens, C., Li, H., Manzey, D.: Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis. *Journal of the Human Factors and Ergonomics Society* 56, 476-488 (2014)
3. Ashton, K.: That "Internet of Things" Thing. *RFID Journal* (2009)
4. Weiser, M.: The Computer for the 21st Century. *Scientific America* 265, 94-104 (1991)
5. Gill, H.: From Vision to Reality: Cyber-Physical Systems. In: *Proceedings of the HCSS National Workshop on New Research Directions for High Confidence Transportation CPS: Automotive, Aviation, and Rail*, Tyson's Corner, VA (2008)
6. Lee, E.A.: *Cyber Physical Systems: Design Challenges*. Technical Report No. UCB/EECS-2008-8 University of California, Berkeley, CA (2008)
7. Geisberger, E., Broy, M.: *agendaCPS: Integrierte Forschungsagenda Cyber-Physical Systems*. Springer, Heidelberg (2012)
8. Bradshaw, J., Feltovich, P., Jung, H., Kulkarni, S., Uszok, W.T.A.: Dimensions of Adjustable Autonomy and Mixed-initiative Interaction. *Agents and Computational Autonomy*. LNCS, vol. 2969, pp. 17-39. (2004)
9. Richling, J., Werner, M., Jaeger, M., Mühl, G., Heiß, H.-U.: *Autonomie in verteilten IT-Architekturen*. De Gruyter, München (2011)
10. Chestnut, H.: Automation: What it is and What are the Problems it Poses. *Automatica* 1, 241-252 (1963)
11. Parasuraman, R., Riley, V.: Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39, 230-253 (1997)
12. Russell, S., Norvig, P.: *Artificial Intelligence: A Modern Approach*. Pearson, Harlow (2014)
13. Franklin, S., Graesser, A.: Is it an Agent, or Just a Program?: A Taxonomy for Autonomous Agents. *Intelligent Agents III: Agent Theories, Architectures and Languages*. LNCS vol. 1193, pp. 21-35. Springer, Berlin (1997)

14. Schillo, M., Fischer, K.: A Taxonomy of Autonomy in Multiagent Organisation. Agents and Computational Autonomy. LNCS vol. 2969, pp. 68-82. Springer, New York, NY (2004)
15. Munroe, S., Luck, M.: Agent Autonomy Through the 3M Motivational Taxonomy. Agents and Computational Autonomy. LNCS vol. 2969, pp. 55-67. Springer, New York, NY (2004)
16. Jennings, N.: Agent-based Computing: Promise and Perils. In: Proceedings of the 16th International Joint Conference on Artificial Intelligence, pp. 1429-1436. Stockholm (1999)
17. Brustoloni, J.: Autonomous Agents: Characterization and Requirements. Technical Report, Pittsburgh, PA, Carnegie Mellon University (1991)
18. Wooldridge, M., Jennings, N.: Intelligent Agents: Theory and Practice. The Knowledge Engineering Review 10, 115-152 (1995)
19. Kalenka, S., Jennings, N.: Socially Responsible Decision Making by Autonomous Agents. In: Proceedings of the 5th International Colloquium on Cognitive Science, pp. 135-149. Springer, Donostia-San Sebastián (1999)
20. Endsley, M., Kaber, D.: Level of Automation Effects on Performance, Situation Awareness and Workload in a Dynamic Control Task. Ergonomics 42, 162-192 (1999)
21. Parasuraman, R., Sheridan, T., Wickens, C.: A Model for Types and Levels of Human Interaction with Automation. IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 30, 286-297 (2000)
22. Yanco, H., Drury, J.: Classifying Human-robot Interaction: An Updated Taxonomy. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, pp. 2841-2846. The Hague (2004)
23. Holland, J.H.: Studying Complex Adaptive Systems. Journal of Systems Science and Complexity 19, 1-8 (2006)
24. Agah, A.: Human Interactions with Intelligent Systems: Research Taxonomy. Computers & Electrical Engineering 27, 71-107 (2000)
25. Lee, J., See, K.: Trust in Automation: Designing for Appropriate Reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society 46, 50-80 (2004)
26. Sheridan, T., Verplank, W.: Human and Computer Control of Undersea Teleoperators. Technical Report Massachusetts Institute of Technology, Cambridge, MA (1978)
27. Endsley, M.: The Application of Human Factors to the Development of Expert Systems for Advanced Cockpits. In: Proceedings of the Human Factors and Ergonomics Society 31st Annual Meeting, pp. 1388-1392. New York, NY (1987)
28. Jipp, M.: Levels of Automation: Effects of Individual Differences on Wheelchair Control Performance and User Acceptance. Theoretical Issues in Ergonomics Science 15, 479-504 (2014)
29. Sheridan, T.: Adaptive Automation, Level of Automation, Allocation Authority, Supervisory Control, and Adaptive Control: Distinctions and Modes of Adaptation. IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 41, 662-667 (2011)
30. Sheridan, T., Parasuraman, R.: Human-automation Interaction. Reviews of Human Factors and Ergonomics 1, 89-129 (2005)
31. Sheridan, T.: Supervisory Control. In: Salvendy, G. (ed.) Handbook of Human Factors, pp. 1025-1052. Wiley, New York, NY (1997)
32. Castelfranchi, C.: Founding Agent's 'Autonomy' On Dependence Theory. In: Proceedings of the 14th European Conference on Artificial Intelligence, pp. 353-357. Amsterdam (2000)
33. Endsley, M., Kiris, E.: The Out-of-the-loop Performance Problem and Level of Control in Automation. Human Factors: The Journal of the Human Factors and Ergonomics Society 37, 381-394 (1995)