



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Alltägliches Handeln in intelligenten Systemen

Philosophische Überlegungen zur
Hybridisierung von Menschen durch
intelligente Technik - Konsequenzen für ein
technisches Design im Ubiquitous Computing

Vom Fachbereich Gesellschafts- und Geschichtswissenschaften der
Technischen Universität Darmstadt zur Erlangung eines Grades des
Doktors der Philosophie
(Dr. phil.)

Dissertation von
Dipl.-Ing. (FH), M.A. Manuel DIETRICH
aus Frankfurt am Main

Erstgutachter: Prof. Dr. Christoph HUBIG
Zweitgutachter: Prof. Dr. Kristof VAN LAERHOVEN
Drittgutachterin: Prof. Dr. Petra GEHRING

Darmstadt 2018
D17

Dietrich, Manuel: „Alltägliches Handeln in intelligenten Systemen: Philosophische Überlegungen zur Hybridisierung von Menschen durch intelligente Technik - Konsequenzen für ein technisches Design im Ubiquitous Computing“

Darmstadt, Technische Universität Darmstadt.

Jahr der Veröffentlichung der Dissertation auf TUPrints: 2018

URN: nbn:de:tuda-tuprints-74779

Tag der mündlichen Prüfung: 25.04.2018

Veröffentlicht unter CC BY-SA 4.0 International

<https://creativecommons.org/licenses/>

Für Svenja und Marlene

Inhaltsverzeichnis

1	Einleitung	1
1.1	Zielstellung, Ausrichtung, Gang der Arbeit und Methode . .	4
2	Konzeptuelle Vorarbeit: Charakterisierungen, Visionen und leitende Beispiele	9
2.1	Charakterisierung intelligenter Systeme	10
2.1.1	Künstliche Intelligenz: Ein Exkurs	10
2.1.2	Charakterisierung aus Sicht maschinellen Lernens . .	16
2.1.3	Systematische Erweiterung	18
2.2	Intelligente Systeme im Ubiquitous Computing	20
2.2.1	Die Vision des Ubiquitous Computing	20
2.3	Zwei Typen von Mensch-Technik-Verhältnissen im Ubiquitous Computing	23
2.3.1	Kontextsensitive Interaktion	24
2.3.2	Self-Tracking-Techniken	26
3	Entwurf einer Philosophie der Mensch-AI-Hybridisierung	29
3.1	Technik als Mittel, Mittler und Vermittler: Auf den Spuren der Medialität der Technik	31
3.1.1	Medialität der Technik als nicht-neutrale Transformation	32
3.1.2	Hybridisierung und das Verschwinden technischer Medien im Vollzug: Medialität von Technik aus Sicht von Sybille Krämer	35
3.1.3	Hybridisierung und die Rolle medialer Erfahrung: Spurenlesen als ein Abduzieren nach Christoph Hubig .	42
3.1.4	Hybridisierung als Blackboxing: Medialität nach Bruno Latour	48
3.1.5	<i>Trivialität</i> nicht-trivialer Maschinen: Heinz von Foersters Maschinenmetapher	55
3.2	Verleiblichung im Handeln mit Technik	60
3.2.1	Einführung	61
3.2.2	Unterscheidung: Leib und Körper	63
3.2.3	Maurice Merleau-Ponty: Erwerb von Gewohnheiten als Verwandlung des Körperschemas	64

3.2.4	Phänomenologie technischen Handelns bei Martin Heidegger	69
3.2.5	Technik-Mensch-Verhältnisse bei Don Ihde: Eine kritische Betrachtung	75
3.3	Philosophie der Mensch-AI-Hybridisierung	79
3.3.1	Zusammenführung: Hybridisierung von Menschen durch Technik	80
3.3.2	Schwierigkeiten einer Mensch-AI-Hybridisierung und Möglichkeiten eines Vertraut-werdens	83
3.3.3	Die Hybridisierung hinsichtlich der zwei Typen von Mensch-AI-Verhältnissen im Ubiquitous Computing	85
4	Verstehbarkeit intelligenter Systeme: Kandidaten zur Vermittlung von Systemmechanismen	91
4.1	Verstehbarkeit intelligenter Systeme im maschinellen Lernen	93
4.1.1	Ein Datensatz und zwei gängige Lernstrategien . . .	93
4.1.2	Entscheidungsbäume	95
4.1.3	Künstliche neuronale Netze	98
4.1.4	Fazit: Verstehbarkeit maschineller Lernalgorithmen .	102
4.2	Verstehbarkeit intelligenter Systeme in der HCI	103
4.2.1	Verstehbarkeit durch Erklärungen	104
4.2.2	Verstehbarkeit im Zuge praktischer Erfahrung	107
4.2.3	<i>In the Wild</i> : Interaktion mit einem intelligenten Heizsystem	109
4.2.4	Fazit: Verstehbarkeit intelligenter Systeme	111
4.3	Kandidaten zur Vermittlung von Systemmechanismen . . .	112
4.3.1	Schritte der maschinellen Klassifikation - Chronologie	113
4.3.2	Quellen	114
4.3.3	Metadaten	115
4.3.4	Semantische Modelle: Domainwissen	117
5	Konsequenzen für die Gestaltung von Self-Tracking-Techniken	119
5.0.5	Methodische Ausrichtung	119
5.1	Technisches Setup	121
5.1.1	Das Sensorsystem: Der <i>Hedgehog</i>	121
5.1.2	Die Technik der Aktivitätserkennung auf Basis von Beschleunigungssensoren	122
5.2	Studie 1: Intelligentes Self-Tracking-System zur möglichen Anwendung in der psychologischen Therapie	124
5.2.1	Szenario: Aktivitätserkennung für die psychologische Therapie	125
5.2.2	Lernansatz: <i>Dense Motif Discovery</i>	126
5.2.3	Visualisierung der Zwischenschritte	129

5.2.4	Übersicht: Studiendesign	130
5.2.5	Ergebnisse der Studie	133
5.2.6	Diskussion der Ergebnisse	136
5.2.7	Fazit: Erste Studie	138
5.3	Studie 2: Self-Tracking-Techniken und die Möglichkeiten und Grenzen einer Selbstreflexion	138
5.3.1	Das Ziel ist die Reflexion: Eine Diskussion der Konzepte	139
5.3.2	Übersicht: Studiendesign	141
5.3.3	Versuchsaufbau	142
5.3.4	Ergebnisse der Studie	146
5.3.5	Diskussion der Ergebnisse	155
5.4	Die Ergebnisse der Studien aus Sicht der Hybridisierungs- theorie	159
6	Schluss und Ausblick	163
6.1	Zusammenfassung	163
6.2	Ausblick: Die Hybridisierung aus einer normativen Perspek- tive	174
6.3	Fazit	177
	Literatur	179

Kapitel 1

Einleitung

Intelligente technische Systeme begegnen uns bereits heute im Alltag. Nicht als menschenähnliche und ebenbürtige Agenten, sondern als Gebrauchstechniken, die in bestimmten Anwendungsbereichen und in einem begrenzten Aktionsrahmen oft sogar als *hochgeistig* geltende Aufgaben souverän erledigen. Zu den von *intelligenten* Techniken geleisteten Aufgaben zählen beispielsweise Vorausplanungen, Situationserfassungen sowie Störungsmanagement; gelegentlich verarbeiten diese Systeme sogar historisches Wissen. Ein aktuelles prominentes Beispiel für die Leistungsfähigkeit *intelligenter* technischer Systeme ist das Computersystem *AlphaGO*, welches das Spiel *Go* mit Bravour meistert und sogar dazu in der Lage ist, einen menschlichen Großmeister zu schlagen.¹ Darüber hinaus gibt es mittlerweile über den ganzen Globus verteilte *intelligente* Computersysteme wie z. B. Suchmaschinen im Internet, personalisierte Online-Werbung, Fahrerassistenzsysteme, Auto-Trading (für Aktien), Heiz- und Klimatechnik im *Smart-Home* sowie eine große Bandbreite an Mensch-Roboter-Interaktionen, die von der industriellen Fertigung bis zu Einsätzen im Alten- und Pflegeheim oder im häuslichen Umfeld reichen.

Die Forschung schreitet schnell voran, so dass in immer mehr Bereichen des alltäglichen Lebens *intelligente* Systeme neu implementiert werden oder vormals etablierte Techniken durch *intelligente* ersetzt werden. Die Technologien der Industrie 4.0, autonom fahrende Automobile sowie algorithmenbasierte Systeme zur Vermittlung von Übernachtungsmöglichkeiten oder Fahrdienstleistungen sind weitere Beispiele für aktuelle Trends. In der Praxis der technischen Gestaltung werden die oben genannten Techniken ganz selbstverständlich in einem undifferenzierten Sprachgebrauch als *intelligente* Systeme bezeichnet.

Es ist Ziel dieser interdisziplinären Arbeit, das alltägliche Handeln in *intelligenten* Systemen philosophisch zu untersuchen und zu diskutieren, welche Konsequenzen sich aus diesen Überlegungen für die technische Gestaltung *intelligenter* Systeme ergeben. In der Arbeit wird argumentiert, dass es

¹Siehe u. a.: <https://www.scientificamerican.com/article/how-the-computer-beat-the-go-master/>.

hinreichende Gründe gibt, Intelligenz als Terminus zur Charakterisierung bestimmter technischer Systeme zu verwenden. Der Geltungsbereich, nach der eine Technik als intelligent bezeichnet werden kann, ist Gegenstand der Erörterung.

In einer ersten Annäherung zeichnen sich intelligente technische Systeme, kurz intelligente Systeme, durch die Möglichkeit aus, selbstständig-adaptiv zu agieren. Von einem selbstständigen Verhalten spreche ich bei technischen Systemen dann, wenn Umgebungsvariablen, wozu Situationen, Nutzer und ihre Präferenzen, aber auch bestimmte ökonomische und soziale Entwicklungen gehören können, von dem technischen System gedeutet werden und sich aus diesen Deutungen ein situationssensitives Agieren folgt, was als Adaption bezeichnet wird. Der Begriff der Deutung wird verwendet, um den sonst naheliegenden Begriff der Interpretation zu vermeiden, der häufig mit dem Anspruch eines Verstehens verbunden wird. Deuten dagegen legt nahe, dass es bloße Anzeichen gibt, dass etwas der Fall ist. Eine Präzisierung des Konzepts maschineller Deutung als Basis des Agierens intelligenter Systeme ist Gegenstand der Arbeit (2.1.2). Beispielsweise deuten Fahrerassistenzsysteme die Verkehrssituation bzw. den Zustand des Fahrers und treffen gegebenenfalls sicherheitsrelevante Entscheidungen. Das Gleiche gilt für Internet-Suchmaschinen, die die Suchergebnisse an das bisherige Suchverhalten von Nutzern und an die zur Verfügung stehenden Daten der geographischen Verortung des Suchenden anpassen.

Was folgt daraus für das alltägliche Verhältnis zwischen Technik und Mensch, wenn die Techniken in der hier skizzierten Weise selbstständig-adaptiv agieren? Es wird argumentiert, dass ein Bedarf besteht, existierende Ansätze, die das alltägliche Mensch-Technik-Verhältnis in philosophischer Hinsicht beschreiben, zu präzisieren und zu erweitern (siehe Kapitel 3).

Diese Auseinandersetzung erfolgt anhand der Analyse einer spezifischen Mensch-Technik-Beziehung, in der sich ein alltägliches Handeln mit Technik ausdrückt, die als Hybridisierung beschrieben wird.² Hybride Formen sind Mischformen aus zwei oder mehreren vormals getrennten Entitäten oder Systemen. Der Begriff Hybridisierung findet in vielen wissenschaftlichen und außerwissenschaftlichen Kontexten Verwendung, die aber

²Eine Reihe von Publikationen in unterschiedlichen Disziplinen beziehen sich explizit auf diesen Begriff: Hård und Jamison, 2013, Latour, 2006, Krämer, 2008, Rammert und Schulz-Schaeffer, 2002.

nicht immer einer einheitlichen und strengen Bedeutung folgen. Unter anderem können Textsorten, kulturelle Identitäten oder Antriebstechniken hybride Verbindungen eingehen.³ Die Hybridisierung von Mensch und Technik ist nicht nur Thema der Technikphilosophie, sondern auch der Techniksoziologie und der Technikgeschichte.

Die Hybridisierung wird in der hier vorliegenden Arbeit nicht als ein Mensch-Technik-Verhältnis als Folge einer bestimmten modernen Technik aufgefasst, sondern als ein die gesamte Menschheitsgeschichte durchziehendes Verhältnis. Von einer Hybridisierung lässt sich immer dann sprechen, wenn Menschen funktionale Verbindungen mit technischen Artefakten und Systemen eingehen, die ihr tagtägliches Agieren und Wahrnehmen leiten. Nicht jedes Verhältnis von Mensch und Technik kann als Hybridisierung beschrieben werden, sondern nur funktionale Verbindungen, das heißt solche, in denen die Zielerreichung in Zusammenhang mit der Technik steht.⁴

Eine Hybridisierung als Prozess hat stattgefunden, wenn der Hausmann gewohnheitsmäßig die Haushaltsgeräte nutzt, um den Alltag mit den Kindern zu organisieren oder der Werkzeuggürtel der Handwerkerin zum selbstverständlichen Accessoire ihres alltäglichen Arbeitens geworden ist.⁵ Geht man von dieser Konzeption aus – deren präzise Rekonstruktion ein zentraler Teil der theoretischen Arbeit darstellt (siehe Kapitel 3) – stellt sich die Frage, unter welchen Umständen das Verhältnis zwischen Menschen und intelligenter Technik als eine Hybridisierung verstanden werden kann.

Im Forschungsfeld des Ubiquitous Computing, auf das ich mich in meiner Arbeit beziehe und das den Rahmen vorgibt, innerhalb dessen die praktischen Experimente dieser Arbeit verortet werden können, gibt es seit seinen Anfängen in den 90er Jahren einen starken Wunsch, Techniken mit *intelligenten Befähigungen* zu entwickeln, die in den menschlichen Alltag integrierbar sind und dabei die Nutzer selbstständig unterstützen.⁶

Bei dem Verhältnis von Mensch und intelligenter Technik handelt es sich um eines, das aktuell aus Sicht der technischen Entwicklung im Ubiquitous Computing als problematisch aufgefasst wird. Das zeigt sich vor

³Beispielsweise die Verbindung von Glosse und wissenschaftlichem Text, die Vermischung kultureller Identitäten im Kontext von Migration oder die funktionale Verbindung von Verbrennungs- und Elektromotor.

⁴Ein Mensch-Technik-Verhältnis, bei dem es sich um keine funktionale Verbindung handelt, ist beispielsweise das Verhältnis von Anwohnern zu einer lauten Umgehungsstraße. Sie können mit ihr nur eine *Beziehung* eingehen, etwa indem sie lernen, die Straße weitestgehend zu ignorieren.

⁵In der Arbeit wird der Begriff der Hybridisierung hauptsächlich im Sinne eines Prozesses verstanden und das *Ergebnis* dieses Prozesses als ein Hybrid.

⁶Siehe die Vision von Weiser, 1991, die in 2.2 ausführlich besprochen wird.

allem in der bestehenden Verunsicherung von Nutzern über die Leistungsfähigkeit und Adaptionsfähigkeit der Technik oder in den Schwierigkeiten bei der Zuordnung der eigenen Handlungsanteile und handlungsähnlichen Anteile des Systems am Gesamtverhalten. So zeigen beispielsweise Untersuchungen, dass Verunsicherungen im alltäglichen Umgang in einer intelligenten Smart-Home-Anwendung auftreten können (4.2). Eine wissenschaftliche Auseinandersetzung mit diesem Verhältnis aus Sicht der Technikgestaltung findet im Forschungsbereich der Mensch-Maschine-Interaktion statt (diese wird in der Arbeit aufgrund der englischen Übersetzung Human-Computer Interaction als HCI abgekürzt).

Im Rahmen des interdisziplinären Interesses der Arbeit dienen die philosophischen Überlegungen zur Hybridisierung als Grundlage dafür, Konsequenzen für die Technikgestaltung zu ziehen.

1.1 Zielstellung, Ausrichtung, Gang der Arbeit und Methode

Im Titel dieser Arbeit wird von einem *Handeln in ..* gesprochen, da dadurch und im Gegensatz zu einem *Handeln mit ..* die sonst naheliegende Charakterisierung der Technik als Werkzeug (Instrument oder Tool) vermieden werden soll. Dabei ist hier das Verhältnis von Menschen und intelligenter Technik aus einer auf das Individuum gerichteten Perspektive von Interesse. Der Einzelne wird als Nutzer bzw. als Akteur im Umgang mit technischen Systemen verstanden. Der Arbeit geht es dabei nicht um die Verhältnisbestimmung von Technik und Mensch aus einer gesellschaftstheoretischen Perspektive.

Die Arbeit bewegt sich in ihrer interdisziplinären Ausrichtung zwischen Philosophie und Informatik, spezifischer zwischen Technikphilosophie und Ubiquitous Computing, wobei Letzteres wiederum ein Teilgebiet des Bereichs der HCI ist. Ziel ist es, dem Anspruch beider Disziplinen nachzukommen. Das heißt, dass die philosophische Auseinandersetzung dahingehend geführt wird, dass sie zu methodischen Problemlösungen beim Gestalten intelligenter Techniksysteme beitragen kann. In diesem Kontext werden die diskutierten philosophischen Konzepte veranschaulicht, geschärft und weiterentwickelt. Mit der interdisziplinären Ausrichtung ist zudem der Anspruch verbunden, dass diese Arbeit in gleicher Weise die Grundlagen und Hintergründe zweier Disziplinen berücksichtigt und transparent zu machen versucht, um sie für Leser beider disziplinärer Hintergründe lesbar zu machen.

Im Forschungsfeld der HCI wurden seit den Anfängen dieser Disziplin

bei der Entwicklung technischer Design-Strategien Anleihen bei der Philosophie gemacht.⁷ Die hier vorliegende Arbeit ist methodisch mit diesen Ansätzen verwandt, wobei der Fokus der genannten Arbeiten in einem höheren Maße auf der technischen Umsetzung und Entwicklung einer Design-Strategie liegt, die von den philosophischen Theorien zwar inspiriert sind, selbst aber keine tiefer gehende theoretische Diskussion enthalten. In der Philosophie gibt es andererseits eine Reihe von Arbeiten, die sich mit Computertechnologien und intelligenten Techniksyste men befassen.⁸ Bei diesen Ansätzen spielt zwar die theoretische Auseinandersetzung und begriffliche Schärfung eine zentrale Rolle; die Konzeptionen werden aber weniger mit Blick auf die konkrete Gestaltungspraxis technischer Systeme reflektiert. Es gibt sowohl bei Wieg erling und Hubig, als auch bei Verbeek Vorschläge für eine praktische Gestaltungsstrategie, aber keine Umsetzung und Erprobung der Konzepte. Die hier vorliegende Arbeit ist ein Versuch, die Lücke im Rahmen eines interdisziplinären Ansatzes zu schließen.

Die Arbeit ist in fünf thematische Kapitel unterteilt. Im anschließenden **zweiten Kapitel** werden die für die Arbeit wichtigen begrifflichen Grundlagen erörtert. In diesem Zusammenhang wird ein vertieftes Verständnis von intelligenten technischen Systemen erarbeitet und eine Einführung in die Vision des Ubiquitous Computing gegeben. Ebenso werden zwei grundlegende Mensch-Technik-Verhältnisse im Ubiquitous Computing vorgestellt, die wiederholt der Veranschaulichung der theoretischen Überlegungen dienen.

Das **dritte Kapitel** untersucht sowohl medienphilosophisch als auch leibphänomenologisch das Konzept der Hybridisierung. Das Ergebnis dieses Kapitels ist eine Konzeption, die es zu verstehen ermöglicht, wie Menschen durch Technik in einer alltäglichen, individuellen Konstellation hybridisiert werden und wie sich die Möglichkeiten dieser Verbindung in Bezug auf intelligente Technik verändern. Die theoretischen Überlegungen legen auch frei, warum der Umgang mit intelligenten Systemen problematisch werden kann. Es wird gezeigt, inwiefern von einer weitgehenden *Unmöglichkeit* einer Mensch-AI-Hybridisierung gesprochen werden sollte, die sich dadurch erklären lässt, dass die Spuren der Medialität der Technik im Umgang für die Nutzer verloren gehen.

Die Möglichkeit eines *Spurenlesens* im Erlernen bzw. Erproben von Umgangsweisen ist eine Bedingung für eine Gewöhnung an eine neue Technik, die als Hybridisierung bezeichnet werden kann und die in einem selbstverständlichen und vertrauten Umgang mit der Technik besteht. Im anfänglichen Umgang mit *neuer* Technik können Erwartungen an eine bestimmte

⁷Suchman, 1987, Fällman, 2003, Svanæs, 2013.

⁸Wieg erling, 2011, Hubig, 2015, Verbeek, 2005, Rammert und Schulz-Schaeffer, 2002.

Zielerreichung enttäuscht, erfüllt oder sogar übererfüllt werden. Das Erlebnis einer Enttäuschung, verstanden als Störung des Vollzugs, ist ein Beispiel für eine Erfahrung, die, wenn die Störung identifiziert ist, zu einer Spur des Mediums wird. Das *Spurenlesen* führt zu einer Vorstellung medialer Verfasstheit und macht u. a. eine adäquate Handlungs- und Erwartungskorrektur für die weitere Nutzung möglich und führt zu einem *selbstverständlicheren* Umgang.

Auch im Erproben bzw. Erlernen eines Umgangs mit intelligenter Technik kann es zu den genannten Erlebnissen kommen. Der Unterschied besteht darin, dass die Erlebnisse aber meist nicht hinsichtlich ihrer Ursächlichkeit hin identifizierbar bzw. erfahrbar sind. Die Wirkmechanismen der intelligenten Technik entziehen sich einer *Zugänglichkeit*. Die Überlegungen zur Problematik der Mensch-AI-Hybridisierung wird im letzten Teil des Kapitels anhand der beiden in Unterkapitel 2.3 beschriebenen Mensch-Technik-Verhältnisse im Ubiquitous Computing veranschaulicht und vertieft.

Im **vierten Kapitel** der Arbeit wird eine Methode (d. h. ein Framework) für das Design von intelligenten technischen Systemen entwickelt. Die Ermöglichung eines hybriden Verhältnisses von Menschen und intelligenter Technik, die durch einen nachhaltigen und selbstverständlichen Umgang gekennzeichnet ist, entspricht dem, was im Forschungsfeld des Ubiquitous Computing als ein *gutes* technisches Design gilt. Es wird argumentiert, dass dem zuvor problematisierten Verlust von Spuren in der Gestaltung intelligenter Systeme begegnet werden kann.

Die selbstständig-adaptiven Wirkmöglichkeiten intelligenter Techniken basieren auf der Verwendung lernender Algorithmen. Es ist Ziel dieses Kapitels, die Systematiken lernender Algorithmen dahingehend zu diskutieren, welche *Aspekte der Mechanismen*, die vor allem die Prozessualität aktueller Nutzung betreffen, dem Nutzer vermittelt und visuell dargestellt werden sollen. Solche Aspekte können bestimmte Zwischenschritte in der algorithmischen Aktivität betreffen oder Metadaten über die kontextabhängige Leistungsfähigkeit des Systems sein. Das Ziel ist es, dass der Zugang zu Aspekten der Mechanismen dem Nutzer dazu dient, Erlebnisse der Nutzung zumindest teilweise auf ihrer Ursächlichkeit hin identifizierbar bzw. erfahrbar zu machen.

Im **fünften Kapitel** werden die im Kapitel zuvor entwickelten Kandidaten als Teil eines technischen Systemdesigns im Bereich intelligenter Self-Tracking-Systeme umgesetzt. Bei Self-tracking-Techniken handelt es sich

um technische Systeme, die Sensordaten als personenbezogene Daten maschinell deuten, um es den Nutzern zu ermöglichen, ihre Daten retrospektiv zu betrachten und daraus eventuell weitere Schlüsse zu ziehen. In diesem Kapitel werden zwei Nutzerstudien beschrieben, die von mir mit jeweils 15 Teilnehmenden durchgeführt wurden. In beiden Studien wurde ein Self-Tracking-System verwendet, welches eine maschinelle Deutung von alltäglichen Aktivitäten (Freizeitaktivitäten) auf Basis von Sensoren, die die Probanden am Körper trugen, ermöglicht. Der Fokus des praktischen Teils liegt auf der technischen Umsetzung des Kandidaten-Frameworks, das im Kapitel zuvor entwickelt wurde, wobei disziplinäre Forschungsfragen zur Möglichkeit der Anwendung der Technik in psychologischer Therapie und die Möglichkeit zur technik-vermittelten Selbstreflexion untersucht werden.

Die empirischen Ergebnisse der Nutzerstudien dienen zudem als Grundlage dafür, die theoretischen Überlegungen zur Hybridisierung zu verhandeln. Das Untersuchungsdesign der Studien war so ausgerichtet, dass die Teilnehmer entweder nur die Ergebnisse der Deutung der Aktivitäten erhielten oder zusätzlich auch Aspekte der Mechanismen in visualisierter Form angeboten bekamen. Die Frage war, ob die dargestellten Aspekte der Systemmechanismen es Personen ermöglichte, Spuren zu lesen, das heißt, beispielsweise produktiv mit eventuellen enttäuschten Anfangserwartungen umzugehen, um ein erweitertes Verständnis über die Möglichkeiten und Grenzen des Umgangs mit dem System zu entwickeln.

Das **sechste Kapitel** schließt die Arbeit mit einer Zusammenfassung ab und gibt einen Ausblick, in dem die normative Dimension der thematisierten Hybridisierung angesprochen wird.

Kapitel 2

Konzeptuelle Vorarbeit: Charakterisierungen, Visionen und leitende Beispiele

Im Folgenden werden die begrifflichen Vorarbeiten geleistet. Mit intelligenten technischen Systemen, kurz intelligente Systeme, verbinden sich, wie in der Einleitung angesprochen, eine Fülle von Bedeutungen bzw. eine Vielzahl an realen Umsetzungen. Intelligente Systeme konzeptuell zu charakterisieren, ist Inhalt des ersten Teils dieses Kapitels (2.1). Ich spreche von einer Charakterisierung und nicht von einer Definition, weil das Ergebnis der Diskussion nicht dem Anspruch einer grundlegenden Definition von künstlicher Intelligenz gerecht werden wird.

Dazu wird als erstes eine kurze Diskussion der Historie der Forschung im Feld der künstlichen Intelligenz präsentiert und eine prominente Linie philosophischer Reflexion, die sich vor allem an der weitreichenden und anhaltenden Auseinandersetzung von Hubert Dreyfus mit diesem Thema orientiert, vorgestellt. Im Anschluss wird die Charakterisierung aus Sicht des maschinellen Lernens konkretisiert, da maschinelle Lernalgorithmen ein zentrales Element intelligent geltender technischer System darstellen.

In zweiten Teil des Kapitels wird die Vision des Ubiquitous Computing vorgestellt, in dem die Arbeit in ihrem praktischen Teil verortet ist (2.2). Dabei werden einige der Begrifflichkeiten, die innerhalb der Vision definiert wurden, diskutiert, wobei die Terminologie der Vision bis heute starken Einfluss auf die Forschung in diesem Bereich hat.

Im dritten Teil des Kapitels werden zwei Typen von intelligenten Systemen im Ubiquitous Computing eingeführt, die in der Arbeit als leitende Beispiele dienen werden (2.3). Dabei handelt es sich zum einen um kontext-sensitive Systeme und zum anderen um Self-Tracking-Techniken.

2.1 Charakterisierung intelligenter Systeme

2.1.1 Künstliche Intelligenz: Ein Exkurs

Ein eigenständiges Forschungsfeld, das sich mit dem Thema künstliche Intelligenz (AI) befasst, formte sich in den 1950er Jahren.¹ Die ursprüngliche Motivation für die *Erschaffung* von künstlichen Intelligenzen — spezifischer sprach man von elektronischen Gehirnen — lag darin, das menschliche Denken besser zu verstehen, indem man es versucht, nachzubilden.² Neben ersten technischen Erfolgen waren es vor allem hohe Erwartungen, die diese Disziplin zu Beginn schnell entwickelte, was sich unter anderem in der Terminologie des „Electronic Super-Brains“ widerspiegelt.

Viele der Erwartungen in eine umfassende künstliche Intelligenz erfüllten sich jedoch nicht und damit wandte sich die Disziplin von einem generellen Erkenntnisinteresse, das heißt Intelligenz zu erschaffen, um Intelligenz zu verstehen, ab und wurde zu einer spezialisierten technischen Disziplin.³ Demnach wurden statt des Ziels eines umfassenden Verständnisses von menschlichem Denken und menschlicher Intelligenz leistungsfähige Werkzeuge entwickelt.⁴ Ein Zitat aus dem umfassenden Werk zur künstlichen Intelligenz von Russell und Norvig fasst das treffend zusammen, wobei sich die Liste der spezifischen Aufgaben seit Veröffentlichung stark erweitert hat:

„AI currently encompasses a huge variety of subfields, from general-purpose areas such as perception and logical reasoning, to specific tasks such as playing chess, proving mathematical theorems, writing poetry, and diagnosing diseases“.⁵

Die Vorstellung einer Selektivität intelligenter Befähigung – eine Art von Inselbegabung – wird zumindest von technisch-orientierten Informatikern als unproblematisch empfunden, widerstrebte jedoch Forschern, die den Traum hatten, mit dem Nachbilden von Intelligenz auch menschliche Intelligenz zu verstehen. Wiederum wird bis heute ein selektives Verständnis von Intelligenz vor allem von Philosophen, die Intelligenz als eine umfassende Befähigung (Vermögen) betrachten, kritisiert (dazu mehr in der folgenden *phänomenologischen Kritik*).

¹Künstliche Intelligenz wird in der hier vorliegenden Arbeit als *AI* abgekürzt, aus dem engl. *artificial intelligence*.

²Eine Einführung vgl. Russell und Norvig, 2010, S. 3.

³Man verabschiedete sich größtenteils von der Idee einer „strong AI“ und wandte sich hin zu einem schwächeren Verständnis mit geringeren Ansprüchen

⁴Weizenbaum spricht vielmehr von „Abfallprodukten“ der Forschung im Bereich künstlicher Intelligenz, siehe Weizenbaum, 1990, S. 302.

⁵Russell und Norvig, 2010, S. 4.

Strömungen

In den folgenden Abschnitten werden einige grundlegende Strömungen aus der Historie der Forschung zur künstlichen Intelligenz vorgestellt. Der Abschnitt orientiert sich an dem zuvor schon zitierten Werk von Russell und Norvig.⁶ Die Autoren unterscheiden in ihrer historischen Einführung zwischen einer Forschung, die das hauptsächliche Ziel hat, menschlich-intelligente Leistungsfähigkeit nachzubilden und Ansätzen, die auf ein idealisiertes Konzept von Intelligenz abzielen, das allgemein als Rationalität bezeichnet wird.

Bezüglich des ersten Forschungsinteresses, welches als der *thinking-and-acting-humanly*-Ansatz bezeichnet wird, ist die Arbeit von Alan Turing als prominent zu nennen. Der Turing-Test als eine experimentelle Anordnung, die den Anspruch hat, maschinelle Intelligenz bzw. maschinelles Denken zu testen, ist über wissenschaftliche Grenzen hinweg bekannt. Turing nannte seinen Test „imitation game“, den er in seinem 1950 veröffentlichten Artikel „Computing machinery and intelligence“ vorstellte.⁷ Turing bezieht sich bei seinen Überlegungen zur maschinellen Intelligenz bzw. zur Beantwortung der Frage, ob Maschinen denken können, auf eine Versuchsanordnung, in der eine Person in der Interaktion mit zwei weiteren Personen, die sie nicht sieht, mit denen sie einzeln und nur schriftlich kommuniziert, entscheidet, welcher der Interaktionspartner ein Mann und welcher eine Frau ist. Daraus entwickelt er die Frage, was passiert, wenn der Computer die Rolle eines der Interaktionspartner übernimmt und am Spiel teilnimmt.⁸ In einer so modifizierten Versuchsanordnung des *imitation game* müsste eine Person, die bei Turing *interrogator* heißt, entscheiden, bei welchem von zwei Interaktionspartnern es sich um einen Menschen oder eine Maschine handelt.

Würde eine Person in einem solchen Versuch vom Computer davon überzeugt werden, ein Mensch zu sein, gälte der Computer aus Sicht von Turing als denkend bzw. intelligent. Wenn nicht, wäre das jedoch nicht ein ausschließendes Kriterium, wenn man Dennetts Deutung von Turings Test folgt: „failing this test is not supposed to be a sign of lack of intelligence“.⁹ Die Leistungsfähigkeit künstlicher Intelligenz wird bei Turing mittels der Kommunikationsfähigkeit der Maschine mit einem Menschen gemessen. Teil seiner Argumentation ist es, dass die Fähigkeit, in einer Konversation zu überzeugen, ein klares Zeichen, sogar ein hinreichender Indikator,

⁶Russell und Norvig, 2010.

⁷Turing, 1950.

⁸vgl. Turing, 1950, S. 434.

⁹Dennett, 2004, S. 3 über Turing, 1950.

für Intelligenz ist.¹⁰ Gleichzeitig schloss er aber auch nicht aus, dass es andere Indikatoren von Intelligenz gibt. Folgt man dieser Argumentation, so könnte man sich einen *Intelligenztest* für Maschinen vorstellen, in dem eine künstliche Intelligenz ein Kunstwerk erschafft, das beispielsweise auch aktuelle politische und soziale Tendenzen aufnimmt und ein kunstinteressiertes Publikum davon überzeugt, dass das Kunstwerk von einem Menschen erschaffen wurde. Demnach könnte man argumentieren, dass die Maschine Intelligenz im Sinne einer kreativen Intelligenz besitzt. Ob Turing kreatives Vermögen als hinreichendes Kriterium für generelle Intelligenz annehmen würde, wie er es beim kommunikativen intelligenten Verhalten tut, lässt sich nicht sagen.

Die Forschung im Bereich künstlicher Intelligenz, die auf einem idealisierten, rationalen Verständnis von Intelligenz fußt, wird als symbolische AI bezeichnet. In diesem Bereich ist es das erklärte Ziel, Maschinen als *rational agents* zu konstruieren, meist unter der Annahme, dass das menschliche Denken im Kern rational (vernünftig) ist, wenn es nicht durch irrationale Bedürfnisse und Triebe gestört wird, und dass es dieser Anteil ist, der menschliches Verhalten und Denken intelligent macht.

Folgt man dieser Idee, liegt es nahe zu sagen, dass Maschinen im höheren Maße als Menschen das Vermögen haben, intelligent, weil rational, zu handeln, eben da ihnen nicht der Makel eines teilweise irrationalen Denkens *anhaftet*. In der symbolischen AI wird die Welt unter anderem als eine Fülle von Gegenständen, Ereignissen, Personen oder Phänomenen in einer maschinellen Struktur zeichenhaft repräsentiert. Ist das *Weltwissen* in ein von Maschinen lesbares Format transformiert worden, kann dieses die Grundlage für Operationen sein kann, welche rein logischen Regeln folgen. Ob die Ergebnisse dieser Operationen sich dann aber wiederum mit grundlegenden Erfahrungen in der Welt decken, hängt von den Annahmen bzw. Kriterien ab, nach denen das Weltwissen in die Systeme eingegeben wurde. Diese Kriterien sind jenseits der Sphäre logischer Regeln.

„Symbolic AI assumes, as its principal point of departure, that intelligence is a matter of manipulating symbols following fixed and formal rules. A series of assumptions is made to arrive at this idea. A first, necessary assumption is that all intelligent processes, including perceiving, reasoning, calculating, and language use, are forms of information processing, that is, of uptaking information from the environment, processing or manipulating of this information, and providing a response. Thus,

¹⁰Neben Daniel Dennett hat sich auch John R. Searle prominent mit Turings *imitation game* auseinandergesetzt, was in dem Gedankenexperiment des „Chinese Room“ Bekanntheit erlangte, siehe dazu Searle, 1984.

when one adds numbers, one first determines the addend, performs a certain operation on this information, and then exhibits the solution".¹¹

In diesem Zusammenhang sprach man in den 70er Jahren von sogenannten „microworlds“, in denen technische Systeme fähig waren, bestimmte Probleme zu lösen, für deren Lösung man davon ausging, dass Intelligenz eine Voraussetzung ist („chose limited problems that appeared to require intelligence to solve“.¹²). Die Grenzen symbolischer Repräsentation und der eingeschränkte Rahmen, in dem logische Schlüsse zu sinnvollen Ergebnissen in der Praxis führten, zeigten, dass man sich nicht auf der richtigen Spur hin zu einer grundlegenden künstlichen Intelligenz befunden hatte. Die anfängliche Annahme, dass die Grenzen des symbolischen Ansatzes durch die mangelnde Leistungsfähigkeit der Computer und der nicht ausreichenden Kapazität der Speicher lag, wurde durch die Vermutung eines substanziellen Problems abgelöst.¹³ Unter anderem die Kritik von Hubert Dreyfus an der symbolischen Intelligenzforschung, die im Folgenden behandelt wird, beschleunigte die Verbreitung dieser Annahme in der Forschung.

Phänomenologische Kritik

Ein Aufsatz von Hubert Dreyfus mit dem Titel „What Computers Can't Do“ wurde im Jahre 1972 veröffentlicht und schlug hohe Wellen in der AI Community.¹⁴ Die Kritik, die Dreyfus an dem bisher dominanten symbolischen Ansatz anbrachte, war von seiner Arbeit im Bereich phänomenologischer Philosophie geprägt. Zu Beginn war es vor allem Martin Heideggers „Sein und Zeit“, das ihm als Inspirationsquelle diente, ein Buch, bei dessen Übersetzung ins Englische er beteiligt war und zu dessen Verbreitung im englischsprachigen Raum er maßgeblich beitrug.

Dreyfus zentrale Argumentation ist es, dass die symbolische AI eine *letzte Inkarnation* einer als veraltet geltenden Wissenschaftsauffassung ist, die als Rationalismus bezeichnet wird. Unter anderem bezieht er sich auf René Descartes, der Dreyfus zufolge einer der zentralen Vertreter dieser veralteten Auffassung ist, wobei Descartes nicht nur bei ihm immer wieder als Strohmann dient. Ich spreche von einem Strohmann, weil er sich inhaltlich nicht weitergehend mit Descartes Schriften beschäftigt, sondern das Kontrastbild, das schon bei andern Phänomenologen vor ihm verwendet wurde, einfach übernimmt. Dreyfus versteht die Grundannahme einer

¹¹Brey, 2001, S. 39.

¹²Russell und Norvig, 2010, S. 19.

¹³Vgl. Russell und Norvig, 2010, S. 21.

¹⁴Wie er später selbst sagt: Dreyfus, 1972.

rationalistischen Erkenntnistheorie im Sinne Descartes, so, dass sie auf einem Dualismus von Geist (Gehirn) und körperlicher Umwelt aufbaut. Die räumliche Umwelt wird dabei durch die Sinne (bei Dreyfus: „input from the universe“), vermittelt, wobei die Stimuli vom Gehirn hinsichtlich von Merkmalen abstrahiert werden und somit als Grundlage dafür dienen, eine Repräsentation der Welt zu konstruieren.¹⁵ Diese Annahme, nach der die Welt rational repräsentiert werden kann, leitet auch die symbolische AI-Forschung.

Dreyfus Verständnis orientiert sich an der neuen Auffassung der Phänomenologie, deren Erkenntnistheorie er in der AI-Forschung dieser Zeit nicht wiederfindet. Er bezieht sich dabei unter anderem auf Heidegger und dessen philosophischen Überlegungen zum Werkzeuggebrauch. In Heideggers Überlegungen in *Sein und Zeit* ist der Hammer eben nur ein Hammer, weil man mit ihm hämmern kann und in Relation zu anderen *Dingen*, wie Nägel, sowie der Fähigkeit eines solchen Werkzeuggebrauchs.¹⁶ Die Fähigkeit, zu Hämmern, oder andere Fähigkeiten seien, wie Dreyfus sagt, „*stored*, not as representations in the agent’s mind, but as the solicitations of situations in the world“.¹⁷ Demnach ist es der Horizont von Dingen (als *Zeugganzes*) sowie die Tätigkeit des damit Umgehens, die den Hammer zuallererst charakterisiert, die nicht in Symbolen repräsentierbar sind. In Abschnitt 3.2.4 wird die Technikphilosophie Heideggers hinsichtlich der Hybridisierung ausführlich vorgestellt.

Dieser phänomenologischen Auffassung folgend argumentiert Dreyfus, dass Menschen in den seltensten Fällen von internen Repräsentationen in ihrem Handeln geleitet werden. Vielmehr sei menschliche Intelligenz situativ und *embodied*, das heißt, durch die Anforderungen von Situationen in der Welt geleitet. Der situative Ansatz von Dreyfus wurde in der Forschung zur künstlichen Intelligenz nach anfänglichem Widerstand immer mehr zum konkreten Forschungsthema und etablierte sich sogar unter dem Begriff der *Heideggerian AI*, die im Grunde eine *Dreyfus AI* ist und seinen Einfluss widerspiegelt. Unter anderem wurde der Fokus bei der Gestaltung von intelligenten Systemen mehr als zuvor auf eine physische Präsenz und real-weltliche Aktivität von Computern gelegt, die dann meist als Roboter mit der Umwelt interagierten (in Form von Kommunikation und Handhabung) und deren Weltauffassung sich hauptsächlich an den Erfahrungen des eigenen Agierens und nicht an menschlichen Wissenssystemen orientierten.

¹⁵Vgl. Dreyfus, 2007, S. 1150.

¹⁶Heideggers phänomenologische Überlegungen unter anderem zum Werkzeuggebrauch werden in Abschnitt 3.2.4 eingehend behandelt.

¹⁷Dreyfus, 2007, S. 1144.

Demnach ist der Kernaspekt des phänomenologischen Ansatzes bei Dreyfus die Idee eines „Learning without Representation“. Dieses findet er in neueren Entwicklungen im Bereich der AI, den „Künstlichen Neuronalen Netzen“. ¹⁸ Diese orientieren sich an Konzepten, die die Funktionsweise des menschlichen Gehirns beschreiben sollen. Mehr zu der Systematik von neuronalen Netzen in Abschnitt 4.1.3.

An dieser Stelle ist es vor allem zentral, dass die neuronalen Netze (ich spreche von neuronalen Netzen, wenn ich künstliche neuronale Netze meine) eine mehrschichtige Struktur von Knoten und Verbindungen haben, die in ihrer Gesamtheit Wissen *speichern*, das in ihrer Aktualisierung (Anwendung) wirkt, aber sonst als Blackbox erscheint, über die keine exakten Vorhersagen über ihr Wirken getroffen werden kann. Das Netz verfestigt sich zuvor mit Hilfe von Trainingsdaten in einem iterativen Prozess. Hier sieht Dreyfus die Anforderungen an ein menschenähnliches leibliches Lernen erfüllt, weil man bei dem stabilisierten Netz eben nicht mehr von einer symbolischen Repräsentation sprechen kann. Es handelt sich nach Dreyfus dabei aber auch nicht um eine bloße Struktur von Assoziationen. ¹⁹ Gleichzeitig weist er aber auch auf die Grenzen der neuronalen Netze hin und widerspricht der Annahme, wieder auf der Spur hin zu *strong AIs*, das heißt zu einem generellen künstlich intelligenten Agenten, zu sein. Dreyfus sagt dazu: „Still there are many important ways in which neural nets differ from embodied brains. Some of them seem to be limitations that can be overcome by further research“. ²⁰

Es ist zu bezweifeln, ob die Forschung intelligenten Agenten im dreyfusschen Sinne in den letzten Jahren nähergekommen ist, obwohl viele praktische Erfolge der AI als ein Tool zu verzeichnen sind. Gründe für den Erfolg liegen anteilig in der schnellen Entwicklung von Rechenleistung (vor allem durch schnelle Grafikkarten), der ausgereiften Methodik zur Initialisierung der Netze sowie dem Zugang zu großen Mengen von Trainingsdaten (Phänomen des Big Data). ²¹ Es ist jedoch nicht Teil dieser Arbeit, die Erforschung der künstlichen Intelligenz philosophisch aufzuarbeiten.

Der Exkurs zur Forschung und Historie der künstlichen Intelligenz hat gezeigt, dass es richtig sein kann, von intelligenten Systemen zu sprechen, wenn diese in der Erfüllung *spezifischer* Aufgaben intelligent *erscheinen*. Das heißt, dass die Systeme ein bestimmtes Verhalten aufweisen, welches uns

¹⁸Dreyfus, 2002.

¹⁹Dreyfus, 2007, S. 1151.

²⁰Dreyfus, 2002, S. 374.

²¹Aktuelle Erfolge der Deep-Learning-Algorithmen, medienwirksam beispielsweise die AI, der es möglich war, einen menschlichen Großmeister im Spiel Go zu schlagen, das als viel schwerer als beispielsweise Schach gilt, lässt AI-Puristen wieder von einer generellen Kompetenz im Sinne einer *strong AI* („general purpose reasoning“ und „creativity“) träumen. In einem Artikel von Lake et al. wird, Stand 2016, das diskutiert Lake et al., 2016, S.1.

(Menschen) normalerweise dazu veranlasst, auf Intelligenz zu schließen. Menschen können in ein *interaktives* Verhältnis mit einer solchen intelligenten Technik eintreten. Obwohl es sich im engeren Sinne um eine *simulierte* Intelligenz handelt, macht es Sinn, sich mit dieser Konstellation im Verständnis einer Mensch-AI-Hybridisierung auseinanderzusetzen.

In den folgenden Abschnitten werden intelligente Systeme, wie sie in dieser Arbeit verstanden werden, aus der Sicht maschinellen Lernens weiter charakterisiert.

2.1.2 Charakterisierung aus Sicht maschinellen Lernens

In diesem Abschnitt wird diskutiert, wie intelligente Systeme aus technischer Sicht konkretisiert werden können. Bei den meisten technischen Systemen, die als intelligent charakterisiert werden, ist eine starke konzeptuelle Nähe zum maschinellen Lernen zu finden. Das heißt, dass maschinelle Lernalgorithmen ein basaler Bestandteil des Systems sind. Das ist insofern kein Zufall, als dass die Methoden, die sich in dem Feld des maschinellen Lernens versammeln, häufig ihren historischen Ursprung in der AI-Forschung haben.

Um die Fülle von Methoden zu unterscheiden, wird in der Forschung zwischen überwachtem, unüberwachtem und bestärkendem Lernen unterschieden. Beim überwachten Lernen werden Daten verwendet, die von Personen hinsichtlich einer bestimmten Bedeutung gekennzeichnet wurden. In der internationalen Forschung spricht man bei der Kennzeichnung von einem Labeling, was ich im Folgenden auch machen werde. Beim unüberwachtem Lernen gibt es keine Labels, sondern die Vorschläge für eine Deutung werden vielmehr vom System selbst gemacht. Beim bestärkenden Lernen ist der Lernprozess durch ein aktives Agieren und durch die Bewertung der Ergebnisse dieses Verhaltens (Reaktion) in der realen Umgebung geprägt. Auf die Unterscheidung wird im Abschnitt ?? ausführlicher Bezug genommen. Es ist üblich, die Technik des maschinellen Lernens auf ein „statistical pattern recognition“ (statistische Mustererkennung) zu reduzieren. Das Erkennen von Mustern in Daten (oft Sensordaten) auf Basis einer repräsentativen Menge von Beobachtungen (statistisch) ist ohne Frage ein zentrales Element maschinellen Lernens. Es gibt einige Forschungsbereiche, wie eingangs erwähnt unter anderem die AI-Forschung oder das Feld des Data Mining, die starke inhaltliche Überschneidungen haben.²² Auf den Aspekt der Mustererkennung wird im Folgenden weiter Bezug genommen, um die intelligenten Systeme weiter zu charakterisieren.

²²Siehe Einführung bei Alpaydm, 2008.

Die Systematik der Mustererkennung als *pattern recognition* lässt sich wie folgt skizzieren: Die grundlegende Idee ist, dem System beizubringen, wie bestimmte Muster in Sensordaten oder sonstigen maschinell lesbaren Quellen (Wissensdatenbanken) gedeutet werden sollen. Muster sind örtliche Regelmäßigkeiten, die sich durch ihren Horizont, das heißt hinsichtlich der Perspektive der Betrachtung (Granularität, *nah oder fern*), von anderen (relevanten) Mustern und von unregelmäßigen, chaotischen bzw. heterogenen Erscheinungen unterscheiden.²³ Ich denke, es ist schlüssig anzunehmen, dass man nicht allgemein entscheiden kann, was ein Muster ist und was nicht, sondern dass es von der Perspektive abhängt und dass jeder noch so heterogenen Beobachtung hinsichtlich bestimmter Kriterien mit genügend großem Abstand oder zeitlichem Horizont ein Muster *abzugewinnen* ist.

Bestimmte Eigenschaften, (*features*), in den Sensordaten können als Indikatoren für einen bestimmten Deutungsinhalt dienen. Die Grundlage für die Verknüpfung von Features und Deutungsinhalt ergibt sich aus den Trainingsdaten. Der Deutungsrahmen wird dabei meist von einem gewählten Kategoriensystem vorgegeben. Beispielsweise können auf Basis eines binären Kategoriensystems Porträtfotos dahingehend unterschieden werden, ob Männern oder Frauen auf den Fotos abgebildet sind. Beim überwachten Lernen ist das Kategoriensystem meist von den Entwicklern eines solchen Systems vorgegeben, welches bestimmten funktionalen Vorgaben folgt. Die Kategorien dienen als Labels für die Trainingsdaten, mit Hilfe derer das System lernt, wie Eigenschaften in den Daten mit den gewünschten Deutungsinhalten zusammenhängen. Das Lernen ist eine Optimierung auf *Diskriminierbarkeit*, das heißt, die Suche nach den besten Sets an Eigenschaften, die es dem System erlauben, zwischen den Kategorien zu unterscheiden bzw. zu entscheiden, dass keine der Kategorien zutrifft. Dass auf eine Unterscheidbarkeit optimiert wird, heißt auch immer, dass die Eigenschaften, nach denen das System klassifiziert wird, keine generellen Charakteristiken hinsichtlich der klassifizierten Entitäten sind, sondern nur solche, die relativ zu den anderen Kategorien bestehen.

Im Umgang mit einem intelligenten System können Nutzer auf unterschiedliche Weise mit den Klassifikationen von Mustern als *etwas* (Deutungsinhalt) konfrontiert sein. Je nach Anwendung kann die Klassifikation die Grundlage für eine Visualisierung für den Nutzer sein oder sich ein Systemverhalten im Sinne von *control tasks* anschließen. Beide Fälle werden es in den Typen von Anwendungen geben, die später als leitende Beispiele besprochen werden (2.3).

Wenn diese Bedingungen einer erlernten Mustererkennung gegeben

²³Eine Diskussion zur Realität von Mustern, siehe „Real Patterns“ von Dennett, 1991.

sind, kann man aus meiner Sicht von einem (minimalen) intelligenten System sprechen. Für diese Charakterisierung ist es wichtig, dass die Entscheidungen der Systeme auf Ansätzen des maschinellen Lernens beruhen, das heißt, dass sie nicht Ursache eines sogenannten *hard-codings* sind. Das wäre der Fall, wenn einem Algorithmus von Entwicklerseite explizit bestimmte Schwellenwerte (im engl. *thresholds*) vorgegeben werden. Dann kann man im engeren Sinne nicht mehr von einem intelligenten System sprechen. Eine erweiterte maschinelle Intelligenz, das heißt Systeme, die mehr als eine minimale Intelligenz besitzen, kann man Systemen zuschreiben, die auf Ansätzen des bestärkenden Lernens (auch Online-Lernen) beruhen. Diese können sich auch während der Laufzeit des Systems, das heißt nach abgeschlossener Lernphase, hinsichtlich ihrer Deutungs- und Handlungsfähigkeiten, aber immer nur im Rahmen von bestimmten Möglichkeiten (die s. g. *cost functions* vorgeben), weiterentwickeln. Der Unterschied zwischen einem Verhalten, das Ergebnis eines Lernens ist und einem explizit programmierten Verhalten ist nicht eindeutig, da auch beim maschinellen Lernen viel explizites Wissen einfließen kann.

2.1.3 Systematische Erweiterung

Die Charakterisierung intelligenter Systeme aus technischer Sicht wurde mit Bezug auf das maschinelle Lernen vorgenommen. Eine minimale Intelligenz kann man Systemen unterstellen, denen auf Basis von Trainingsdaten eine algorithmische Deutungsperformanz eingeschrieben wurde, das heißt, die Möglichkeit, dass Muster wiedererkannt werden, und basierend auf dieser Deutung sich bestimmte Aktionen anschließen. Im Folgenden wird ein Verständnis von intelligenten Systemen hinsichtlich ihres Potentials zur Mustererkennung systematisch erweitert.

Die Mustererkennung hat das Ziel, einem computertechnischen System eine Struktur einzuverleiben, die es ermöglicht, Observablen auf eine bestimmte Art und Weise zu deuten. Observablen können unter anderem Sensordaten sein, die so aufbereitet werden, dass eine automatische Etikettierung (*Labeling*) möglich ist. Unter vorherigen, vergleichbaren Bedingungen wurde dem System *beigebracht*, welches Label zu einem bestimmten sensorischen Phänomen gehört.

Die Labels, die die Daten als Trainingsdaten für ein überwachtes Lernen verwendbar machen, werden in der Entwicklung festgelegt. Sie basieren auf bestimmten Kriterien, die auf einem etablierten Kategoriensystem basieren können, sowie einem bestimmten Instanzwissen bei den Personen, die das *Labeling* durchführen. Unter Instanzwissen werden die impliziten Regeln, denen Personen folgen, wenn sie eine beobachtete Instanz einer bestimmten Kategorie zuordnen, verstanden. Ein Szenario ist vorstellbar, in

dem Videos (beispielsweise von YouTube) dahingehend gelabelt werden, welche Sportarten Personen in den Videos ausüben. Die Labels könnten dann dazu dienen, einem System beizubringen, selbstständig Sportarten in Videos zu erkennen.²⁴ In einem solchen Szenario könnte den Personen, die das Labeling durchführen, zur Vorbereitung ein Kategoriensystem zur Hand gegeben werden. In dem Kategoriensystem kann festgelegt sein, was (relevante) Sportarten sind; gegebenenfalls kann auf verifiziertes Wissen Bezug genommen werden (etwa, was in einem Lexikon als Sportart aufgeführt oder vom olympischen Sportbund vorgegeben wird). Trotz dieser gemeinsamen Basis kann ein Labeling, das von unterschiedlichen Personen durchgeführt wurde, sehr unterschiedlich aussehen. Nicht, weil sie nicht gewillt sind, einem gemeinsamen Kategoriensystem zu folgen, sondern beispielsweise, weil sie ein anderes Instanzwissen haben, das heißt, sie eine andere beobachtete Aktivität als Sportart X bezeichnen würden bzw. unterschiedliche Vorstellungen über den Anfang und das Ende der Aktivität haben.

Allgemein kann man sagen, dass in den Trainingsdaten eine Menge an normativen Gehalt stecken kann, sowohl aufgrund der Auswahl des Kategoriensystems als auch der Praxis der Zuordnung.²⁵ Ein bestimmtes Muster in den Beobachtungsdaten (Observablen) kann als Indikator für eine bestimmte maschinelle Deutung stehen, die immer bestimmten expliziten und impliziten bzw. subjektiven und objektiven Kriterien folgt. Bestimmte Deutungen als Observablen zweiter Ordnung können dann wiederum Grundlage sein, um sie eventuell in Kombination mit anderen Observablen zu deuten. Auf diese Weise können mehrstufige Systeme entstehen, die kontinuierlich Deutungen vornehmen, die wiederum Grundlage für weitere maschinelle Deutungen sind. Die Kriterien, die die Grundlage für die Deutung sind, *verschwinden* im Algorithmus des Systems, wenn sie nicht anderweitig dokumentiert sind, das heißt, nur der Deutungsmechanismus bleibt bestehen.

Maschinelles Lernen auf Basis von Mustererkennung kommt in vielen Bereichen zur Anwendung. Klassisch sind es Anwendungen im Feld der *Computer Vision* oder allgemeiner der *Machine Perception*, wobei die Grundlage der Deutung von Mustern die Sensordaten sind, die einem technischen System als Videos oder Fotos vorliegen.

²⁴Die automatische Deutung von Videos ist eine mögliche Praxis, siehe beispielsweise hier: Miyamori und Iisaku, 2000.

²⁵Mit zunehmender Verbreitung maschineller Lernansätze in unterschiedlichen *Produkten* stellt sich die Frage, ob es einen *Bias* in den Trainingsdaten gibt, der bestimmte Personengruppen diskriminiert, beispielsweise Systeme zur Gesichtserkennung für bestimmte ethnische Gruppen schlechter funktionieren, da sie in den Trainingsdaten unterrepräsentiert waren.

Ein Anwendungsbereich unter vielen, in dem das maschinelle Lernen als Mustererkennung eine Rolle spielt, sind beispielsweise die *Digital Humanities*. Beim sogenannten „Distant Reading“ werden Ansätze aus dem maschinellen Lernen verwendet, um inhaltliche, stilistische oder grammatikalische Muster in Texten zu deuten.²⁶ Dieser Ansatz ist unumstritten, hatte aber ohne Frage einen starken Einfluss in der Forschung: „It aims to generate an abstract view by shifting from observing textual content to visualizing global features of a single or of multiple text(s)“.²⁷ Mit diesem Ansatz können große Mengen von Texten hinsichtlich bestimmter *Muster* gedeutet werden. Inwieweit diese Quantifizierung einer wissenschaftlichen Erkenntnis zuträglich ist, ist die zentrale Frage in der Diskussion.

In diesem Teil des Kapitels wurde ein Verständnis von intelligenten Systemen entwickelt, das für die weitere Arbeit ausreichend ist.

Im Folgenden wird die begriffliche bzw. konzeptuelle Vorarbeit fortgesetzt, indem das Ubiquitous Computing und seine prägende Vision vorgestellt und diskutiert wird. Zudem wird auf die Rolle von intelligenten Systemen in der Vision und der aktuellen Forschungspraxis eingegangen.

2.2 Intelligente Systeme im Ubiquitous Computing

Das Ubiquitous Computing ist ein interdisziplinärer Forschungsbereich, der in der Informatik verortet ist, und dessen Entstehung maßgeblich einem wissenschaftlichen Artikel von Mark Weiser zu verdanken ist, der den Terminus und die Idee einführte. Zentral ist die Idee einer Computerisierung der alltäglichen Lebenswelt, die durch den technischen Fortschritt, der immer kleinere Computersysteme hervorbringt, und durch die Vorstellung einer unauffälligen Assistenz von Menschen in ihren alltäglichen Aufgaben geleitet ist. Die Wahl der Mittel zur Realisierung der Assistenz sind häufig Lernalgorithmen, wie sie zuvor charakterisiert worden sind.

2.2.1 Die Vision des Ubiquitous Computing

Die Vision des Ubiquitous Computing, aus der sich in der Folge ein ganzer Forschungsbereich entwickelte, basiert auf einem Artikel, den Mark Weiser, den er 1991 in der wissenschaftlichen Zeitschrift „Scientific American“ veröffentlichte.²⁸ Mark Weiser gehörte in dieser Zeit zu einer Riege von Wissenschaftlern des renommierten Xerox PARC Forschungsinstituts, wo

²⁶From Close to Distant and Back: How to Read with the Help of Machines, Bonfiglioli und Nanni, 2016.

²⁷Jänicke et al., 2015.

²⁸Weiser, 1991.

er in einem interdisziplinären Team arbeitete. In einem Text über den Einfluss auf Mark Weisers Arbeit schreibt Paul Dourish: „Weiser’s ubicompo was informed by a number of non-technical impulses and sources“.²⁹ Vor allem die Arbeit mit der Anthropologin Lucy Suchman, die sein Team bei PARC leitet, wird hervorgehoben.³⁰

Das Ubiquitous Computing etablierte sich als die „third wave‘ of computing technologies“, welches nach Auffassung der Forscher in diesem Bereich den beiden vorherigen Meilensteinen, dem Großrechner und dem Personal Computer, folgt.³¹ Aus technischer Sicht zeichnet sich diese dritte Welle durch kleine und leistungsstarke Computer aus, „worn, carried or embedded in the world arround us“.³² Folgt man dieser Wellen-Metapher kann man allgemein sagen, dass die Computersysteme mit jeder weiteren Welle physisch kleiner und gleichzeitig leistungsstärker werden. Allein die Tatsache, dass durch eine geringere physische Größe der Hardware diese auch handlicher werden, eröffnet einen Horizont neuer Möglichkeiten des Einsatzes. Bei der zweiten Welle ist es der private Bereich des Zuhauses und das eigene Büro und bei der dritten Welle erfolgt der Einsatz überall.

Weiser startet seinen Artikel im Scientific American mit den Sätzen „The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it“.³³ Er stellt sich eine Computertechnik vor, die sich mit dem *Gewebe* des alltäglichen Lebens *verflechtet*, so dass sie aus der Wahrnehmung der Nutzer verschwindet. Das Verschwinden wird dabei nicht hauptsächlich als ein physisches Verschwinden gesehen, sondern als ein Verschwinden hinsichtlich der bewussten Wahrnehmung der Nutzer: „Such a disappearance is a fundamental consequence not of technology but of human psychology“.³⁴ Hier kommt der anfänglich angesprochene interdisziplinäre Einfluss zum Tragen, wobei er in seinem Text unter anderem auf die Phänomenologie Heideggers hinweist, die auch Lucy Suchman, mit der er bei PARC zusammengearbeitet hatte, in ihrem Werk als Inspirationsquelle diente.³⁵ Nur wenn die Systeme so gestaltet sind, dass sie keine oder kaum aktive Aufmerksamkeit benötigen, handelt es sich um eine Ubiquitous Computing-Technik. Die Idee der Reduktion aktiver Aufmerksamkeit wird durch eine,

²⁹Dourish und Bell, 2011, S. 10.

³⁰Lucy Suchman hatte 1987 mit ihren philosophisch-ethnologischen Arbeiten sowohl auf konzeptueller und auch empirischer Ebene großen Einfluss (Suchman, 1987).

³¹Dourish und Bell, 2011, S. 1.

³²Dourish und Bell, 2011, S. 2.

³³Weiser, 1991, S. 94.

³⁴Weiser, 1991, S. 94.

³⁵Vgl. zur Historie Dourish und Bell, 2011, Kapitel 2: „Contextualizing Ubiquitous Computing“, S. 9ff, Suchman, 1987.

wie er sagt, *nahtlose* Integration von technischen Geräten in den Alltag realisiert. Diese "seamless intergration" ist eine der Leitideen, die in vielen Publikationen in der direkten Folge von Weisers Vision als Motivation Verwendung fand und bis heute sehr wichtig ist.³⁶

Im Hauptteil von Weisers Artikel wird die konzeptuelle Einbettung größtenteils nicht weitergeführt, sondern vor allem, wie es sich für einen Artikel der Informatik gehört, die technischen Lösungen und deren Nutzen thematisiert. Hier stehen die zu dieser Zeit sprunghaft verbesserte Hardware kleiner Sensoren und deren günstige Verfügbarkeit sowie deren Vernetzung im Mittelpunkt: "cheap, low-power computers [...] and a network that ties them together".³⁷ Als Beispiele nennt er unter anderem technische Systeme, die kontextsensitiv sind („context-aware“) und bei deren späterer Realisierung Ansätze aus dem maschinellen Lernen zur Anwendung kamen. Kontextsensitive Techniken sind solche, die je nach Situation, in der sich der Nutzer mit dem technischen Gerät befindet, sein Verhalten anpasst, ohne dass der Nutzer diese Anpassung explizit veranlassen muss. Ein bekanntes Beispiel ist das eines kontextsensitiven Kommunikationsgerätes (z. B. Smartphone), welches erkennt, dass der Nutzer sich gerade in einer Situation befindet, in der er nicht gestört werden möchte, beispielsweise in einem Meeting. Mit diesem Wissen über die Situation kann das System selbstständig den Modus der Benachrichtigung des Nutzers, beispielsweise über den Erhalt einer E-Mail oder einer Kurznachricht, von einem Benachrichtigungston zu einem Vibrationssignal umstellen. Eine solche Anwendung ist, folgt man der Vision, nahtlos, weil sie keine aktive Intervention von Seiten des Nutzers benötigt, so dass dieser nicht jedes Mal, wenn er nicht gestört werden will, den Modus der Benachrichtigung *händisch* einstellen muss. Bei diesem Beispiel und bei anderen dieser Art stellt sich die Frage, ob die Annahme Weisers richtig war, da eine automatische Anpassung, wenn der Nutzer die Bedingungen der Anpassung nicht genau kennt, oft eher irritiert. Eine kritische Auseinandersetzung mit der Vision des Ubiquitous Computing wird an verschiedener Stelle in meiner Arbeit vorgenommen (u. a. siehe Abschnitt 3.3).

Mark Weiser spricht in seinem Artikel nicht explizit von intelligenten Systemen. In vielen seiner Szenarien ist es aber ein zentraler Bestandteil, dass sich die Systeme an die Intentionen und Bedürfnisse von Menschen anpassen. Das heißt, sie lernen deren Gewohnheiten und unterstützen diese, ohne dass der Nutzer explizit dazu auffordern muss. Eines seiner Beispiele ist eine Kaffeemaschine, die *weiß*, wann eine Person aufstehen will

³⁶Weitere Diskussion dieser Begrifflichkeit: Chalmers und Maccoll, 2003.

³⁷Weiser, 1991, S. 100.

und den Kaffee dann schon zubereitet hat.³⁸ Zudem bespricht er einen sogenannten „foreviewer“, der eine Person dabei unterstützt, den besten Weg zur Arbeit und anschließend einen Parkplatz zu finden.³⁹ Eine Anpassung an die Nutzer, wobei deren Verhalten und die Umgebung im Rahmen des maschinellen Lernens gedeutet werden, ist ein zentraler Aspekt vieler Anwendungen im Ubiquitous Computing.

2.3 Zwei Typen von Mensch-Technik-Verhältnissen im Ubiquitous Computing

Im Folgenden werden zwei Typen von Mensch-Technik-Verhältnissen vorgestellt, die in der Forschung und Gestaltung des Ubiquitous Computings eine wichtige Rolle spielen. Die beiden Typen werden in dieser Arbeit als leitende Beispiele dienen, auf die regelmäßig Bezug genommen wird. Der Beispielcharakter ist insofern wichtig, als dass es sich um zwei wichtige, aber nicht die einzigen möglichen Mensch-Technik-Verhältnisse im Ubiquitous Computing handelt.

- Kontextsensitive Interaktion
- Self-Tracking-Techniken

Die beiden Typen wurden unter anderem deshalb ausgewählt, weil sie in Mark Weisers Artikel schon besprochen worden sind und ihnen bis heute eine wichtige Bedeutung in Forschung und Gestaltung zukommt. In einem Artikel von Abowd und Mynatt wurden die Anwendungsbeispiele in Weisers Arbeit systematisch diskutiert. In einer Klassifizierung, die sich aus Weisers Veranschaulichungen ergibt, sprechen sie zum einen davon, dass „ubicom applications need to be context-aware, adapting their behavior based on information sensed from the physical and computational environment“ und zum anderen, dass „ubicom applications strive to automate the capture of live experiences“.⁴⁰ Sie behandeln in ihrer Diskussion von Mark Weisers Artikel neben den beiden genannten Typen auch die Rolle von *naturlicher* Kommunikation („natural interfaces that facilitate a richer variety of communications capabilities“) wie Rede, Gestik und Mimik für die Gestaltung von technischen Systemen im Ubiquitous Computing.⁴¹ Letzteres wird in dieser Arbeit aber nicht eigens behandelt.

³⁸Weiser, 1991, S. 102.

³⁹Weiser, 1991, S. 102.

⁴⁰Abowd und Mynatt, 2000, S. 30.

⁴¹Abowd und Mynatt, 2000, S. 30.

Bei den beiden ausgewählten Typen handelt es sich um zwei grundlegende, selbstverständlich nicht ausschließliche, Mensch-Technik-Verhältnisse. Zum einen Technik als Partner, der in einem engeren Sinne als Interaktionspartner gesehen werden kann und zum anderen als Mittler in der Selbstwahrnehmung.⁴²

Andere Mensch-Technik-Verhältnisse, die im Ubiquitous Computing eine Rolle spielen, sind beispielsweise die technisch-vermittelte Kooperation, wie sie im Bereich des *Computer-supported Cooperative Work* (CSCW) Thema ist. In meiner Arbeit werden sie aber nicht explizit behandelt.

2.3.1 Kontextsensitive Interaktion

Kontextsensitive Systeme sind eine der Technologien, die in Mark Weisers einflussreichem Artikel behandelt wurden und seitdem als eigener Teilbereich große Wichtigkeit in Forschung und Gestaltung erlangten. Die Idee ist, dass sich Systeme an bestimmte Situationen der Interaktion anpassen, indem sie die sensorisch-erfasste Umgebung maschinell deuten und darauf basierend Handlungsoptionen auswählen. Im Rahmen dieser Forschung wurden verschiedene Versuche unternommen, eine allgemeingültige Definition von Kontext zu entwickeln, welche die Forschung leiten soll. Die bekannteste und allgemeinste Definition ist die von Abowd et al.:

„Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves“.⁴³

In diesem Sinne ist der Kontext eine Ansammlung von relevanten Informationen, die unterschiedlichen Typen zugeordnet werden können. Abowd et al. nennen explizit den Ort, die Identität, die Zeit und die Aktivität als, wie sie sagen, primäre Kontextinformationen. Es wird nicht begründet, warum es sich bei den genannten Informationen um primäre Kontextinformationen handelt. Vielmehr scheinen die *State-Of-The-Art*-Möglichkeiten der sensorischen Erfassung den Rahmen vorzugeben, was als Kontext gilt.

Beispielsweise gilt der Ort als eine Kontextinformation, die zentrale Bedeutung in den Anfängen der kontextsensitiven Systemen erlangt hat (*location-sensitive systems*). Ort als *location* wird dabei meist als eine GPS-Koordinate verstanden, die abhängig vom zugehörigen Kartenmaterial verschiedenen Bedeutungen des Ortes zulässt. Durch die zunehmende Anzahl

⁴²Eine konzeptuelle Unterscheidung zwischen diesen beiden Typen als Basis für ein HCI-Design wurde hier publiziert: Dietrich und Laerhoven, 2015.

⁴³Abowd et al., 1999, S. 3f.

von Sensoren in Computergeräten (Smartphone als *multi-sensor device*) und vor allem die Vernetzung von Geräten zum Austausch von Daten nimmt die Menge der Quellen zur Bestimmung des *Kontexts* immer weiter zu.⁴⁴ Bei der Deutung von Sensordaten als einem bestimmten Kontext spielt das maschinelle Lernen eine Rolle, beispielsweise in der Deutung einer Situation als *Meeting*, wie es im vorherigen Abschnitt besprochen wurde.

Ein kontextsensitives System reagiert immer aktiv auf die Deutung der Umgebung als Kontext, das heißt, es folgt meist eine darauf basierende Verhaltensanpassung. Eine Information ist laut Definition nur dann eine Kontextinformation, wenn sie relevant ist und relevant ist sie, wenn es Teil der Funktion des Systems ist. Ein Beispiel dafür ist ein System, das sich an die aktuelle Leistungsfähigkeit eines Nutzers anpasst, beispielsweise indem es den Schwierigkeitsgrad eines Computerspiels graduell verändern kann. Dabei wäre es nützlich, die aktuelle körperliche und geistige Verfassung des Nutzers zu kennen. Wissen über die Tag- und Nachtzeit könnte dabei eine relevante Information zur Deutung der aktuellen Verfassung sein. Stellt sich aber bei der Personalisierung des Systems an den Nutzer heraus, dass dieser keinen tageszeitabhängigen Leistungsunterschied zeigt, ist die Tag- und Nachtzeit zwar eine Information, auf die die Anwendung Zugriff hat, aber eben keine Kontextinformation, weil sie für die (spezifische) Nutzung nicht relevant ist. Allgemein kann man sagen, dass jede Information, die relevant für die Nutzung ist, als Kontext und jede automatische Anpassung daran als kontextsensitiv gilt.

Kontextsensitive Computersysteme werden bei Mark Weiser und später bei vielen anderen vor allem mit dem Anspruch einer nahtlosen (*seamless*) Interaktion verknüpft. Der Nutzen der Nahtlosigkeit liegt in der Einbettung der Technik in die Praktiken des schon Bekannten, als Routinen und Gewohnheiten, indem beispielsweise das sonst übliche aufwendige Einüben von neuen Weisen der Nutzung möglichst gering gehalten werden: „Reducing the distractions that such interactions currently introduce“, sagen Chalmers und MacColl in ihrer Diskussion von *seamless interaction*.⁴⁵ Nahtlos ist die Interaktion, gerade weil das technische System weniger explizite Inputs des Nutzers benötigt und sich an bestimmte Änderungen relevanter Umgebungsvariablen selbstständig anpasst.

Ein Spezialfall der kontextsensitiven Systeme ist der, wenn sich die Kontextfassung und Anpassung des Agierens Bestandteil einer Interaktion mit dem Nutzer ist. Dabei sind die unterstellten Erwartungen des Nutzers Teil der Modellierung des Handlungshorizonts und legen fest, welche systemischen Anpassungen vorgenommen werden. Dies geschieht oft in Form

⁴⁴Perera et al., 2014.

⁴⁵Chalmers und Maccoll, 2003.

einer Anpassung des Interfaces, was dazu führt, dass sich die Optionen für den Nutzer verändern mit der Technik weiter umzugehen.

2.3.2 Self-Tracking-Techniken

Bei Self-Tracking-Techniken handelt es sich um technische Systeme, die, wie es in der Literatur zum Ubiquitous Computing heißt, *alltägliche Erlebnisse* erfassen können.⁴⁶ Es handelt sich hier selbstverständlich nicht im engeren Sinne um die Erlebnisse selbst, sondern vielmehr um Informationen oder erst einmal nur um Daten, die alltägliches Erleben begleiten. Dabei handelt es sich im weitesten Sinne um die Aktivitäten des Tages, deren Spuren sich in Quantifizierbarem wie motorischen Bewegungen des Körpers, in der Herzfrequenz oder in der Kreditkartennutzung zeigen können.

In der Forschung versammeln sich die Self-Tracking-Techniken vor allem unter den Überbegrifflichkeiten des *Quantified Self* und der *Personal Informatics*.⁴⁷ Ein gängiges technisches Setup in diesen Bereichen bilden tragbare Sensorsysteme, sogenannte *Wearables*, wie Smartphones, Smartwatches oder speziell dafür ausgelegte Geräte (Fitnessarmbänder). Daneben ist auch die alltägliche Verwendung von Computertechnik beispielsweise zur Kommunikation eine naheliegende Quelle für personenbezogene Daten. Dazu können E-Mails, Webseitenaufrufe, Filmstreaming oder Online-Bankaktivitäten gehören.⁴⁸ Die aus unterschiedlicher Quelle stammenden quantifizierten personenbezogenen Informationen werden dann in der Regel fusioniert, maschinell gedeutet und für die Nutzer visualisiert. Die so visualisierten Daten können wiederum dazu dienen, dass Nutzer ihr eigenes Verhalten verstehen und reflektieren können: „We define personal informatics systems as those that help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge“.⁴⁹

Eine häufige verwendete Technik des Self-Tracking ist die Technik der Aktivitätserkennung (*activity recognition*), wobei es sich um eine sensorische Erfassung und Deutung physischer Aktivität handelt. Auf Seiten der Hardware werden zum einen *on-body*-Sensoren, beispielsweise ein Armband mit einem Bewegungssensor⁵⁰, zum anderen sogenannte *ambient sensors*, die so heißen, weil sie in einer technisierten Umgebung wie in einem *smart-home*.

⁴⁶Siehe oben: „capture of live experience“.

⁴⁷Li, Dey und Forlizzi, 2011, Rooksby et al., 2014, Choe et al., 2014.

⁴⁸Li et al. stellen eine lange Liste von Techniken vor, die sie auf Basis einer Studie, bei der sie Interviews mit sogenannten *Quantified Selfers* geführt haben, erstellt haben Li, Dey und Forlizzi, 2011.

⁴⁹Li, Dey und Forlizzi, 2011, S. 558.

⁵⁰Meist eine *IMU = Incremental Measurement Unit*, die unter anderem aus Beschleunigungssensoren und einem Gyroskop besteht

zu finden sind, verwendet. Klassische *ambient sensors* sind visuelle Sensoren, wie Kameras, die die Aktivitäten der Bewohner aufzeichnen.

Im Bereich der Aktivitätserkennung spielt das maschinelle Lernen eine entscheidende Rolle. Die Sensordaten, die in unterschiedlicher Weise aufgezeichnet werden, müssen in bestimmter Hinsicht gedeutet werden, damit sie kombinierbar und vor allem, damit sie für den Menschen verstehbar und somit nutzbar sind.⁵¹ In der aktuellen Anwendung sind es sogenannte IMUs (die maßgeblich auf Beschleunigungssensoren beruhen), die Bewegung in hoher Frequenz aufzeichnen. Werden diese am Körper getragen, sind Schlüsse auf das physische Verhalten der Personen möglich.

Es gibt eine große Zahl von Projekten in der aktuellen Forschung, in denen zahlreiche unterschiedliche Aktivitäten maschinell gedeutet werden. Neben Aktivitäten alltäglicher Fortbewegung (Laufen, Radfahren, Treppensteigen) gibt es Projekte, in denen spezielle Freizeit- und Sportaktivitäten erkannt werden.⁵² Im Feld des *Quantified Self* geht man erst einmal mit Blick auf mögliche Anwendungen von einem allgemeinen Interesse potentieller Nutzer hinsichtlich ihrer alltäglichen Aktivitäten aus. Daneben gibt es Projekte, die die Aktivitätserkennung in psychologischer und physiologischer Therapie einsetzen möchten.⁵³ Dazu gehören auch Anwendungen zur Erkennung von Rauchaktivitäten, Früherkennung von Panikattacken oder die automatische Dokumentation der Einnahme von Insulin für ein Diabetes-Krankheitsmanagement.⁵⁴

Ein wichtiger Bestandteil von Self-Tracking-Techniken ist die Aufbereitung und Vermittlung der personenbezogenen Daten. Teil der Aufbereitung ist es, neben der Kombination von verschiedenen Datenquellen, beispielsweise die Kombination eines Schrittzählers mit einem GPS-Modul, und der auf maschinellen Lernalgorithmen und Trainingsdaten basierenden Deutung, die Daten an den Handelnden selbst mit Hilfe von Visualisierungen zu vermitteln. In einer Reihe von Forschungsprojekten wird sich intensiv mit dieser Schnittstelle, das heißt der Visualisierung von personenbezogenen Daten beschäftigt.⁵⁵ Der Aspekt der Selbstreflexion mit Self-Tracking-Techniken wird im praktischen Teil dieser Arbeit konzeptuell und empirisch behandelt (5.3).

⁵¹Ein Beispiel für einen schon lange verwendeten Bewegungssensor zum Self-Tracking ist der Pedometer (Schrittzähler). Hierbei handelt es sich klassisch um ein mechanisches Gerät, das die Erschütterungen misst, Le Masurier und Tudor-Locke, 2003. Hier wird kein maschinelles Lernen angewandt und die Deutung folgt einem simplen Schema: jede Erschütterung (über einem bestimmten Grenzwert, der mechanisch eingestellt ist) ist ein Schritt, egal wie er erzeugt wurde.

⁵²Berlin, 2014.

⁵³Singh et al., 2014, Dietrich, Berlin und Laerhoven, 2015, Helal, Cook und Schmalz, 2009.

⁵⁴Scholl und Van Laerhoven, 2012, Cruz et al., 2015, Chiauzzi, Rodarte und Dasmahapatra, 2015.

⁵⁵Choe, Lee und Schraefel, 2015, Cuttone, Petersen und Larsen, 2014.

Gerade mit Bezug auf aktuell weit verbreitete Standardsensoren und immer leistungsstärker werdenden maschinellen Lernalgorithmen lässt sich menschliches Verhalten auf unterschiedlichste Weise aufzeichnen, kombinieren, deuten und präsentieren. Da alltägliches Verhalten maschinell gedeutet wird, ist die Charakterisierung von Self-Tracking-Techniken als alltägliche intelligente Systeme zutreffend.

Kapitel 3

Entwurf einer Philosophie der Mensch-AI-Hybridisierung

Ziel des folgenden Kapitels ist es, das Verhältnis von Mensch und intelligenter Technik als eine Form der Hybridisierung zu konkretisieren. Das Ergebnis ist ein Schritt in Richtung einer Philosophie der Mensch-AI-Hybridisierung.

Das Kapitel ist in drei Abschnitte gegliedert. Im ersten Teil (3.1) werden Überlegungen aus dem Bereich der Medienphilosophie hinsichtlich des Hybridisierungsprozesses diskutiert. Im zweiten Teil (3.2) wird die Hybridisierung des Menschen durch Technik aus Sicht der Phänomenologie des Leibes (im engl. „embodiment“) betrachtet. In einem abschließenden Abschnitt (3.3) werden die Überlegungen anhand der beiden Typen von Mensch-Technik-Verhältnissen im Ubiquitous Computing (2.3), die Self-tracking-Techniken und die kontextsensitive Interaktion, vertieft und reflektiert.

Es gibt eine Fülle an theoretischen Ansätzen sowohl in der Medientheorie bzw. Medienphilosophie als auch in der Phänomenologie des Leibes, die das Verhältnis von Mensch und Technik behandeln. Für die Auswahl der hier behandelten Ansätze ist es zentral, dass sie dazu beitragen können, den Hybridisierungsprozess von Menschen durch Technik aus Sicht eines *bilateralen* Zusammenwirkens von Mensch und Technik zu rekonstruieren, wobei der Mensch in erster Linie die Rolle des Nutzers einnimmt und die Technik die Rolle des Genutzten, ohne dabei einem instrumentellen Verständnis zu folgen, das heißt, dass das Handeln vollständig auf die Initiative des Menschen zurückzuführen ist.

Im medienphilosophischen Kapitel werden verschiedene Medienkonzepte hinsichtlich der Wirkkraft von Medien mit Blick auf die Hybridisierung betrachtet. Zunächst wird das Medienverständnis der angloamerikanischen Technikphilosophie vorgestellt (3.1.1). Das Ziel dieses Abschnitts ist es, ein einleitendes Verständnis der Medialität von Technik im Sinne einer *technical mediation* zu erlangen. Im Anschluss daran wird Sybille Krämers medienphilosophischer Ansatz diskutiert (3.1.2). Krämer versteht die

Konstellation von Menschen und Medien – wobei sie Techniken weitestgehend als Medien auffasst – u. a. als eine *operative Einheit*, wobei die Medien als Vermittelnde oder Mittler in der Aktivität der Einheit hinsichtlich ihrer Rolle aus Sicht der Menschen verschwinden. Diese Einheit wird von mir als eine hybride Konstellation verstanden. Im Anschluss daran wird Christoph Hubigs Medienphilosophie diskutiert, welche er im Rahmen seiner Technikphilosophie entwickelt. Sein Fokus liegt auf der Entstehung hybrider Einheiten im Sinne einer Hybridisierung von Menschen durch Technik, wobei die Medialität der Technik etwas ist, das im Umgang erfahren wird und in diesem Zuge die Einheit konstituiert. Des Weiteren werden Bruno Latours Arbeiten zu Hybriden von menschlichen und nicht-menschlichen Akteuren vorgestellt, wobei es sich bei letzterem maßgeblich um Techniken handelt. Die hybriden Akteure sind ein zentraler Teil seiner Gesellschaftstheorie. Die allgemeinen Überlegungen zur Medialität von Technik werden mit Blick auf die Maschinenmetapher von Heinz von Foerster in Richtung der intelligenten Technik konkretisiert (3.1.5).

Im zweiten Abschnitt des Kapitels wird die Hybridisierung von Menschen durch Technik aus Sicht der Phänomenologie des Leibes betrachtet. Es wird gezeigt, dass die Hybridisierung eine Weise der Verleiblichung darstellt. Diese Sichtweise wird vor allem durch Rückgriffe zum einen auf Maurice Merleau-Ponty und zum anderen auf Martin Heidegger entwickelt. Beide beziehen sich in ihrer Arbeit explizit auf technische Systeme und stellen die grundlegende Bedeutung von Technik für die menschlichen Weltverhältnisse heraus. Ein zentraler Begriff in Merleau-Pontys Arbeit ist das Körperschema, nach dem der Körper als Leib in seiner Gesamtheit bekannt ist und wirksam werden kann, das heißt als ein Gemeinsames agiert. Merleau-Pontys Ansicht nach werden Techniken in der Praxis alltäglichen Handelns und Wahrnehmens in das Körperschema einverleibt (3.2.3). Eine weitere Betrachtung der Hybridisierung von Mensch und Technik findet anhand Heideggers phänomenologischer Theorie statt. Auch wenn Heidegger seine Arbeit nicht explizit als eine Philosophie des Leibes bezeichnet, können seine Überlegungen in „Sein und Zeit“ in diese Richtung aufgefasst werden (3.2.4). Zum Abschluss des zweiten Teils des dritten Kapitels wird ein Ansatz aus der Tradition der angloamerikanischen Phänomenologie diskutiert. Don Ihde setzt sich mit den Arbeiten der klassischen Phänomenologie, vor allem mit Heidegger und Merleau-Ponty, auseinander und beschäftigt sich dabei explizit auch mit intelligenten Techniken (3.2.5).

Im dritten Teil des Kapitels werden die Ansätze zur Hybridisierung aus medienphilosophischer und phänomenologischer Sicht zusammengeführt und hinsichtlich des *neuen* Verhältnisses von Menschen und intelligenter Technik präzisiert.

3.1 Technik als Mittel, Mittler und Vermittler: Auf den Spuren der Medialität der Technik

In Debatten über neue technische Entwicklungen (Technologien) oder in der historischen Betrachtung von *Technologiesprüngen* wird diskutiert, in welcher Weise die Nutzung bestimmter Techniken über ihr instrumentelles Bewirken hinaus einen grundlegenden Einfluss auf das soziale Miteinander, die basalen Weltverhältnisse oder erstrebenswerten Lebensweisen haben.

In den letzten Jahrzehnten gab es *große* technologische Veränderungen, beispielsweise zwischenmenschliche oder beruflichen Kommunikation via E-Mail, im Wissensmanagement durch digitale kollektive Online-Datenbanken (z. B. Wikipedia) oder Geräte zur Navigation (GPS). Mit bestimmten technischen Entwicklungen ändern sich bestimmte Handlungs- und Kommunikationsweisen. Dass die Einführung neuer Technologien nicht nur neue Zwecke ermöglicht und dabei hilft, einzelne Arbeitsabläufe, Prozeduren oder generell den Alltag effizienter zu gestalten, sondern auch negative Effekte mit sich bringt, zeigt sich anhand folgender *Alltagsmeinungen*. Im Arbeitsalltag kann man Personen sagen hören: „Ich rufe lieber einmal an, als fünf E-Mails hin und her zu schicken.“ oder auf dem Rastplatz belauschen: „Ich schaue lieber auf die Karte, als mich auf das Navi zu *verlassen*“. Im ersten Beispiel steht die Nützlichkeit der E-Mail für bestimmte Anliegen in Frage und beim zweiten wird der digitalen Karte und ihren systembedingten Abhängigkeiten (der Verbindung mit Satellitenortung) kein Vertrauen geschenkt.

Techniken in der Mannigfaltigkeit ihres Wirkens in alltäglichen sozialen Zusammenhängen in ihrer Systematik zu untersuchen, ist Ziel einer Konzeption von Technik als Medium. Die Charakterisierung eines Mediums in seinem wörtlichen Ursprung als Vermittelndes legt nahe, dass die Konzeption vor allem für Kommunikationstechniken oder sogenannte Massenmedien (Fernseher, Internet) Geltung besitzt. Eine Philosophie des Mediums bzw. der Medialität geht aber darüber hinaus, so dass u. a. technische Artefakte, Methoden und Verfahren, Theorien und Begriffe hinsichtlich ihres Wirkens in medialer Hinsicht betrachtet werden. Medien bilden allgemein gesprochen Handlungsräume, die festlegen, wie überhaupt gehandelt, das heißt, kommuniziert, wahrgenommen und gelebt werden kann und wie nicht.

Philosophisches Technikverständnis

Bevor die Medialität von Technik als nicht-neutrale Transformation diskutiert wird, geht es mir im Folgenden zunächst darum zu skizzieren, was

unter Technik bzw. technischen Handeln aus philosophischer Sicht zu verstehen ist.

Technisches Handeln lässt sich als das planvolle Einsetzen von Mitteln zur Realisierung von Zwecken verstehen. Beispielsweise kann ein Werkzeug planvoll eingesetzt werden, um eine Gerätschaft oder Maschine zu bauen. In einer solchen Konstellation bezeichnet das Substantiv Technik unter anderem das Werkzeug selbst, genauso wie das materielle Ergebnis dieser zweckrationalen Handlung und die Fähigkeiten bzw. Fertigkeiten zum Werkzeuggebrauch.¹ Der Begriff der Technik gilt, wenn er sich auf Gegenstände bezieht, nicht als Eigenschaft, die es ermöglicht, zwischen Gegenständen zu unterscheiden, beispielsweise zwischen technischen Gegenständen und Naturgegenständen. Vielmehr gilt Technik als ein Reflexionsbegriff, das heißt als ein Begriff, der Verhältnisse beschreibt, beispielsweise das von Menschen und Dingen, wobei das Ding als Technik verstanden wird, wenn es zum Beispiel von Menschen regelmäßig als Mittel zur Sicherung bestimmter Bedürfnisse eingesetzt wird.²

In den Fallbeispielen, die in der hier vorliegenden Arbeit behandelt werden, wird von Techniken in erster Hinsicht als Artefakte oder Systemen gesprochen, die die alltägliche Lebenspraxis unterstützen. Weniger ist die Rede von Technik als Strategie oder Fähigkeit. Die Techniken als alltägliche Gebrauchstechniken sind dabei Produkte, die erfunden, entworfen, gestaltet und gebaut wurden und in einem beruflichen oder privaten Bereich von Personen oder Personengruppen genutzt werden.

3.1.1 Medialität der Technik als nicht-neutrale Transformation

In der anfänglichen Charakterisierung können Techniken in Form von Werkzeugen, Geräten oder Systemen als Mittel verstanden werden, die dazu dienen, bestimmte Ziele zu erreichen. In der Beschreibung des technischen Handelns als ein planvolles Agieren geht man davon aus, dass die Handelnden definierte Zwecke verfolgen und eine konkrete Vorstellung davon haben, wie diese mit Hilfe der Technik zu erreichen sind. Ein mögliches Verständnis technischen Handelns ist es, dass der Handelnde ein Ziel vor Augen hat, eine klare Intention, die er genau mit der Nutzung dieser Technik erreichen kann. Die Technik dient als das Vehikel, das die Wünsche

¹Christoph Hubig benennt hier noch weitere Ebenen: Techniken sind ihm zufolge „(1) einschlägige Fähigkeiten und Fertigkeiten, (2) die in bestimmten Verfahrensschemata [...] (3) das Wissen um diese Schemata [...], (4) das konkrete Agieren und Prozessieren des Bewirkens, (5) die bei diesem Bewirken eingesetzten Artefakte als raumzeitliche Entitäten und schließlich (6) die Ergebnisse eines derartigen Bewirkens als realisierte Zwecke“, siehe Hubig, 2011, S. 1.

²Dazu: Reflexionsbegriffe „sind Begriffe für die Konzeptualisierung von Operationen an Gegenständen, nicht Begriffe der Unterscheidung *zwischen* Gegenständen“ Hubig, 2011, S. 7.) oder siehe auch Grunwald und Julliard, 2005, Nordmann, 2008.

des Handelnden erfüllt. Diese Position gilt als das instrumentelle Verständnis von technischem Handeln. Nach dieser Auffassung werden Techniken als neutral charakterisiert, da sie zuallererst nur durch die Handhabung eines Menschen ihr Wirkpotential entfalten und das Ergebnis des Technikgebrauchs mehr oder weniger dem entspricht, was vom Menschen intendiert wurde.

Die Gegenposition besagt, dass Technik (meist) nicht in der beschriebenen Weise neutral ist, weil in und durch das menschliche Handeln mit Technik auch andere Effekte *mitrealisiert* werden, die nicht intendiert waren. In der englischsprachigen Technikphilosophie wird in diesem Zusammenhang von einer „non-neutral transformation“ gesprochen, die u. a. Autoren wie Longdon Winner, Don Ihde und Peter-Paul Verbeek vertreten Verbeek.³ Dieser Ansatz wird im Folgenden diskutiert.

Don Ihde charakterisiert die neutrale Position als eine Sichtweise auf Technik, die bloß auf menschliches Handeln und Entscheiden zurückzuführen ist, das heißt, unter souveräner Kontrolle menschlichen Agierens stattfindet. Danach sind alle technischen Artefakte und Systeme „merely results of human decisions and manipulations“.⁴ Ihde folgt Longdon Winner, wenn er herausstellt, dass die Nicht-Neutralität eine Annahme ist, die den Bedarf für eine philosophische Auseinandersetzung mit Technik motiviert. Die Annahme einer Nicht-Neutralität gilt nach Ihde für jede Technik:

„For every enhancement of some features, perhaps never before seen, there is also a reduction of other features. To magnify some observed objects, optically, is to bring it forth from a background into a foreground and make it present to the observer, but it is also to reduce the former field in which it fits, and – due to foreshortening – to reduce visual depth and background. Such non-neutral transformation belongs to all technologies“.⁵

Ihde betrachtet große, tiefgreifende technologische Fortschritte und zeigt, wie Entwicklungen im Bereich der Medizin helfen konnten, Probleme zu lösen, diese aber gleichzeitig unbeabsichtigte oder in Kauf genommene Nebenfolgen („Side effects“) entfalteten.⁶ Es gibt demnach technologische Erfindungen und Entwicklungen sowie einen Gebrauch dieser Techniken, die sich nicht ausschließlich auf die ihnen zugrundeliegenden Intentionen zurückführen lassen. Dies sind die sogenannten Neben- und Folgeeffekte. Zum Beispiel beschreibt Landgon Winner in seiner historischen Abhandlung zur medialen Rolle von Technik, wie bestimmte technische Produkte,

³Verbeek, 2005, Ihde, 1990 und Winner, 1980.

⁴Ihde, 1993 mit Bezug auf Winner, 1980, S. 100.

⁵Ihde, 1993, S. 111.

⁶Ihde, 1990, S. 163.

beispielsweise die Einführung des „Snowmobils is reshaping the entire set of social and human/animal relations of the Lapps“.⁷ Ein anderes, von Winner beschriebenes, sehr populäres Beispiel beschäftigt sich mit einem Bauprojekt in Long-Island bei New York. Er argumentiert, dass die Brücken einer Schnellstraße so gebaut worden seien, dass der Zugang zu den Gebieten, in denen die Strände von Long-Island lagen, mit Bussen nicht mehr möglich war. Dies hätte dazu geführt, dass die ärmere schwarze Bevölkerung, die normalerweise den Bus nahm, um zum Strand zu kommen, diesen nunmehr schwerer erreichen konnten. Dies ist ein Beispiel dafür, dass die Beschaffenheit eines bestimmten technischen Artefakts (Schnellstraße) einen starken Einfluss auf die lokale soziale Ordnung hat. Trotz der von verschiedenen Autoren formulierten Kritik, bleibt dieses Beispiel anschaulich und kann hier zur Verdeutlichung des Ansatzes angeführt werden.⁸

Je nach Erzählung kann dieses umstrittene Beispiel unterschiedlich bewertet werden. Wenn die Planer, Gestalter und Politiker nicht nur ein Straßenbauprojekt intendierten, das als Mittel zur Entlastung des Innenstadtverkehrs diente, sondern auch noch rassistische Gründe eine Rolle gespielt haben sollten, dann ist das kein Argument gegen die instrumentelle Perspektive. Wenn dagegen diese Wirkung nur nicht bedacht wurde, dann war es eine nicht-intendierte Nebenfolge, die eben nicht auf Menschen zurückzuführen ist (höchstens in ihrer Nichtbeachtung) und stellt insofern ein Beispiel für den transformativen Charakter von Technik dar.

Allgemein kann man sagen, dass in der retrospektiven Betrachtung eine technische *Entdeckung* als ein Auslöser für eine unendliche Reihe von Wirkungen aufgefasst werden kann, zu denen beispielsweise Folgetechniken oder Anwendungen in nicht intendierten Kontexten gehören. Ein Beispiel ist die Auswirkung der Schreibmaschine auf die Arbeitswelt und damit auf die sozialen Verhältnisse. Die Sekretärin als Berufsbild hat sich erst im Rahmen der Verbreitung der Schreibmaschine etabliert, da Männer, die zuvor Sekretäre waren, diese Techniknutzung als Herabwürdigung ihrer Arbeit empfanden und somit mit der Zeit größtenteils aus dem Berufsfeld verschwanden.⁹

In dieser historischen Perspektive können bestimmte Techniken als Ursachen bzw. als konstituierend für spätere Entwicklungen gesehen werden

⁷ Ihde, 1993, S. 101.

⁸ Zur Kritik siehe: Woolgar und Cooper, 1999.

⁹ Lupton, 1993.

oder, wie im Beispiel der Schneemobile, eine soziale Ordnung und das Verhältnis zur Natur neu strukturieren. Die historische Sichtweise ist erst einmal unabhängig davon, ob bestimmte Folgen und Folgefolgen einer Technik vorhersagbar sind und ob man daran *glaubt*, dass technische Entwicklungen *aufgehalten* bzw. signifikant in ihrem Kurs verändert werden können.¹⁰

Die von den Autoren in diesem Abschnitt beschriebenen Effekte, die von der Technik ausgehen und die als Transformationen aufgefasst werden, beziehen sich hauptsächlich auf ihre politischen und sozialen Neben- und Folgefolgen. Eine Präzisierung der nicht-neutralen Wirksamkeit von Techniken als Medien mit dem Ziel, die Hybridisierung von Menschen durch Technik auf einer lokalen bzw. bilateralen Ebene zu erklären, erfolgt im nächsten Abschnitt anhand eines Ansatzes von Sybille Krämer.

3.1.2 Hybridisierung und das Verschwinden technischer Medien im Vollzug: Medialität von Technik aus Sicht von Sybille Krämer

In diesem Abschnitt wird eine philosophische Theorie der Medialität von Sybille Krämer präsentiert.¹¹ Der Fokus der Darstellung von Krämers Theorie liegt auf ihren Überlegungen zum operativen Verhältnis von Menschen und Medien, das als hybride Konstellation verstanden werden kann.

Krämers Auseinandersetzung mit Medialität bzw. Medien hat das Ziel, eine allgemeine Systematik auszuarbeiten, wobei sie ihren Ansatz anhand von verschiedenen Gegenstandsbereichen diskutiert, beispielsweise anhand der Rede sowie der anonymen Kommunikation in einem globalen Netz (Internet), der Orientierung und Navigation mithilfe von Karten oder auch der psychologischen Therapie. Das Letztere betrifft Überlegungen zur Psychoanalyse, in der aus Sicht Krämers der Arzt als ein Medium in der Therapie eines Patienten fungiert.¹² Techniken in Form von technischen Artefakten oder technischen Systemen spielen als unterschiedliche Arten von Medien eine Rolle in ihrer Arbeit, wobei sie nicht alle Techniken als Medien begreift. Sie diskutiert Kriterien, die erfüllt sein müssen, damit bestimmte Techniken als Medien aufgefasst werden können. Diese Kriterien betreffen nicht die Gegenständlichkeit der Technik, sondern bestimmte Gebrauchsweisen (mehr dazu in Abschnitt 3.1.2).

Ein zentraler Aspekt ihrer Medienphilosophie ist die Eigenart der Medien, die im gelungenen bzw. „störungsfreien Vollzug“ in bestimmter

¹⁰Zu Fragen des Technikdeterminismus siehe Winner, 1977, Hubig, 2015.

¹¹Unter anderem ihre Monographie: „Sybille Krämer (2008): Medium, Bote, Übertragung: kleine Metaphysik der Medialität“, Krämer, 2008.

¹²Siehe dazu u. a. hier Krämer, 2008, S 272f.

Hinsicht zurücktreten bzw. verschwinden.¹³ Das Verschwinden wird nicht als ein buchstäbliches verstanden, sondern als ein sich Entziehen aus der Wahrnehmung der Akteure, die mit oder in dem Medium agieren. Beispielsweise tritt das Kommunikationsmedium hinter dem Gehalt des Übertragenen zurück¹⁴: In der Rede als Art der Kommunikation werden Informationen zwischen einem Sprecher und einem Empfänger ausgetauscht, wobei die Rede im Medium der Stimme erfolgt. Die Stimme als Medium der Rede ist dem Sprecher und dem Zuhörer als ein Mittel ihrer Kommunikation nicht präsent, wenn sie sich erfolgreich verständigen (Nicht nur die Stimme ist ein Medium der Sprache, sondern beispielsweise auch die Luft, welche die Schallwellen überträgt).¹⁵

Dieses Konzept des Verschwindens ist, wie Krämer selbst sagt, an bekannte Positionen der Medienphilosophie angelehnt, wobei sie die Autoren Heider, Mersch oder Luhmann nennt.¹⁶ Eine Erweiterung bzw. Neufassung dieser Überlegungen zur Medialität erfolgt in ihrer Arbeit anhand des Konzepts bzw. der Leitmetapher des sogenannten „postalische[n] Prinzip[s]“.¹⁷ In dieser Konzeption steht der Bote, der zwischen dem Sender und Empfänger *liegt* bzw. *agiert*. Der Bote als Medium überbringt die Botschaft in der Kommunikation, liegt räumlich in der Mitte zwischen den Entitäten und ist dabei Vermittler bzw. Mittler des Austauschs.¹⁸ Diese Metaphorik ist mit Bezug auf Kommunikationsmedien naheliegend, wird aber auch auf andere Arten von Medien erweitert. Das Verschwinden des Mediums im Vollzug wird in ihrer Konzeption terminologisch als eine *Selbstneutralisierung* des Boten in seiner Funktion als Mittler gefasst (Sie spricht auch von dem „sterbende[n] Bote[n]“¹⁹).

In ihrer Abhandlung spricht Krämer von einer Einheit aus Mensch und Medium, wenn sie den funktionalen Aspekt eines störungsfreien Umgangs mit Medien beschreibt. Ein solches Verhältnis erkennt sie beispielsweise in der Nutzung einer Karte als Medium der Orientierung in einem unbekanntem Terrain.²⁰ In der Gebrauchssituation sind Mensch und Karte als eine „operative Einheit“ bzw. „fungierende Einheit“ zu verstehen.²¹ Bei einer solchen Verbindung handelt es sich, bezogen auf die Terminologie meiner Arbeit, um eine hybride Konstellation.

In den folgenden Unterabschnitten werde ich den Ansatz von Krämer

¹³Krämer, 2008, S. 27.

¹⁴Krämer, 2008, S. 38.

¹⁵Krämer, 1998a, S. 79.

¹⁶Krämer, 2008, S. 273f verweist auf Heider, 1927, Luhmann, 2001, Mersch, 2003.

¹⁷Einführung: Krämer, 2008, S. 15.

¹⁸Krämer, 2008, S. 37.

¹⁹Krämer, 2008, S. 36.

²⁰Krämer, 2008, S. 334.

²¹Krämer, 2008, S. 334.

etwas detaillierter diskutieren und untersuchen, wie die Überlegungen zu einem gehaltvollen Konzept von Hybridisierung beitragen kann. Den Anfang macht eine Diskussion zur Wirkkraft des Mediums, die in Anschluss an die Unterscheidung von neutraler und nicht-neutraler Technik im vorherigen Abschnitt betrachtet wird (3.1.1).

Wirkkraft des Mediums

Krämer verortet sich mit ihrer Medienphilosophie im Problemfeld zwischen einer „avantgardistische[n] Medientheorie“, die für ein autonomes Verständnis von Medien werben²² und einem Verständnis von Medien als transparente Vermittler, die neutrale, bloß sinntransportierende Kräfte besitzen. Letztere Annahme geht davon aus, dass Medien selbst keinen Anteil an dem Inhalt haben, der vom Sender zum Empfänger vermittelt wird. Seit den Anfängen der Medientheorie wird die Annahme eines transparenten Mediums in Frage gestellt. Die „Transparenz des Mediums zu unterminieren, bildet die Präsupposition [...], von der nahezu alle einem Medienapriori verpflichteten Positionen ausgehen“.²³ Vielmehr wird davon ausgegangen, dass das Medium am Inhalt beteiligt ist, das heißt eine Wirkung entfaltet, die vom Sender nicht intendiert ist.

In der nach Krämer *avantgardistischen* Medientheorie findet sich eine Zuspitzung dieser Position in dem Sinne, dass nicht mehr von einer Beteiligung an der Botschaft, sondern von einer signifikanten Formung der Botschaft gesprochen wird.²⁴ Diese konstruktivistische Position lebt von der Charakterisierung, nach der Medien nicht nur nicht auf personelle Intentionen reduziert werden können, sondern autonom als „sinnmitemzeugende“ Kraft wirken.²⁵

Gegen diese *avantgardistischen* Ansätze der Medientheorie wendet Krämer ein, dass „allzu leicht [...] das Medium als generativer, mithin bedingender Mechanismus aufgefasst und darin autonom gesetzt“ werde.²⁶ Sie argumentiert, dass Medien fremdbestimmt (heteronom) seien.²⁷ Doch sollte dies ihrer Auffassung nach nicht in einem instrumentellen, neutralen Sinne verstanden werden, vielmehr würden Medien zwar als Instrumente eingesetzt, doch seien die Effekte dieses instrumentalen Agierens nicht vollständig auf den menschlichen Akteur zurückzuführen. Es handelt sich nach

²²Krämer, 1998b, S. 82.

²³Krämer, 1998b, S. 68.

²⁴Das ist die Position von McLuhan, Fiore und Agel, 1999, eine der prägenden Figuren der Medientheorie, siehe dazu Krämer, 1998a, S. 75.

²⁵Krämer, 1998a, S. 73.

²⁶Krämer, 2008, S. 35.

²⁷Krämer, 2008, S. 335.

Krämer um ein „verteiltés Potential [...], bei dessen Produktivität menschliche und nichtmenschliche Komponenten stets zusammenwirken“.²⁸ Es gibt demnach dem Medium inhärente Potentiale, die sich aber nur im Zusammenwirken mit menschlichen Agenten aktivieren.

Technik als Medium

Die medientheoretische Konzeption von Krämer entwickelt sich aus Überlegungen zu Kommunikationsmedien, die auf Grund der Terminologie des Boten, der eine Botschaft überbringt, naheliegend ist. Krämer betont allerdings eine weiterreichende Gültigkeit ihres theoretischen Ansatzes. Um das beispielhaft aufzuzeigen, diskutiert Krämer die Karte als Medium der Orientierung bzw. Navigation mit Hilfe ihres Ansatzes.

Eine Funktion von Karten ist es, sie als Technik zur Wegfindung in einem unbekanntem Territorium einzusetzen. Unter einer Karte versteht Krämer „jene Elementarformen [...], die den Anspruch haben, ein mehr oder weniger umfangreiches räumliches Territorium in einer handlichen Darstellung so zu zeigen, dass wir uns innerhalb dieses Territoriums dann *handelnd orientieren* können“.²⁹ Die Ziele des Nutzers werden erst durch Nutzung der Technik erfüllbar. Das Medium Karte als vermittelnde Entität verschwindet dabei im *alltäglichen Gebrauch*.³⁰ Das Finden des Weges mit Hilfe einer Karte wird durch die spezifische Weise der Darstellung wie auch durch die Möglichkeiten und Fähigkeiten des Lesen-Könnens mitgeprägt.

Bei Krämer entfaltet sich die Wirksamkeit von Medien nur in der Konstellation zusammen mit dem menschlichen Akteur als Aktualisierung von „verteilt[e]n Potential[en]“.³¹ Die Karte als Medium verschwindet in der „operative[n] Einheit“ aus Kartennutzer und Karte im Kontext des Zurechtfindens im Territorium.³² Krämer sagt, eine „Karte zur alltäglichen Orientierung und Selbstlokalisierung zu gebrauchen gelingt nur, wenn wir ‚blind bleiben dürfen‘ für die Verzerrungen, die der kartographischen Projektionsmethode eigen sind“.³³ Mit Verzerrung der Karte meint sie, dass jede Karte notwendig nur eine Weise der Darstellung unter vielen möglichen Darstellungsweisen eines Territoriums ist.

Krämer sieht in dem „Sich-Zurücknehmen“ von Medien jeglicher Art die Eigenart der Medien, die zwingend mit der Ermöglichungsleistung von

²⁸Krämer, 2008, S. 335.

²⁹Krämer, 2008, S. 299f.

³⁰Krämer, 2008, S. 306.

³¹Krämer, 2008, S. 335.

³²Krämer, 2008, S. 334.

³³Krämer, 2008, S. 336.

Medien zusammenhängt.³⁴ Diese Eigenart, so Krämer, unterscheidet Medien von bloßen Techniken. Ein Beispiel für bloße Techniken, denen sie eine mediale Verfasstheit abspricht, sind Werkzeugtechniken. Diese nehmen sich, so ihre Auffassung, nicht in der genannten Weise zurück und sind somit auch im störungsfreien Vollzug präsent. Doch die Annahme einer stetigen Präsenz des Werkzeugs im Gebrauch ist fraglich. Der Hammer liegt in der Hand des erfahrenen Zimmermanns und verschwindet subjektiv hinter dem Ziel, den Dachstuhl fertigzustellen. Zudem ist auch der Hammer in gleicher Weise wie die Karte oder bestimmte Kommunikationsmedien eine den Sinn mitbestimmende Entität, wie es zuvor beschrieben wurde. Der Hammer in der Hand lässt nach dem gleichen Schema, nach dem die Karte das Territorium hinsichtlich einer möglichen Begehbarkeit deutet, die Umgebung hinsichtlich einer *Bearbeitbarkeit* erscheinen. Ein Sprichwort, das in verschiedenen künstlerischen Werken Einzug gehalten hat, bringt die Medialität von Werkzeugtechnik auf den Punkt. So *rapped* beispielsweise Kate Tempest, in einem Refrain eines Tracks: „When all you’ve got is a hammer, everything looks like nails“.³⁵

Erfahrung des Medialen: Spuren und Spurenlesen

Krämer argumentiert, wie ich zuvor skizziert haben, dass das Medium, das sie nach dem Botenmodell rekonstruiert, sich dem Zugriff einzieht, indem der Bote sich beim Überbringen selbst neutralisiert. Diese Eigenart von Medien, nur wirksam zu werden, indem sie sich entziehen, macht es unmöglich, dass ein Medium in seiner Verfasstheit als Potential Gegenstand einer direkten Vorstellung werden kann. Krämer arbeitet mit dem Begriff der Spur, um ausgehend von dieser Annahme Medien und ihre Wirkmacht in wissenschaftlicher Hinsicht begreifbar zu machen.³⁶ Spuren sind Hinterlassenschaften, resultierend aus dem Umgang in oder mit dem Medium, die zwar keinen direkten aber zumindest einen indirekten Zugriff auf die potentielle Verfasstheit zulassen.

Das Beispiel der Stimme als Medium der Rede veranschaulicht Krämers Überlegungen. Der Inhalt der Rede ist das absichtsvoll gebrauchte Zeichen und die Stimme ist das Medium, welches die Botschaft mit formt. „Die Stimme macht Aussagen, aber sie kommentiert auch das Gesagte“.³⁷ Das Mediale an der Stimme sind vor- und außersprachliche Momente, die in der

³⁴Krämer, 2008, S. 274

³⁵Kate Tempest, eine Rapperin aus London, das Zitat stammt aus dem Track „A Hammer“ aus dem Album „Everybody Down“ von 2014; ein Zitat, das in etwas veränderter Form Mark Twain zugeschrieben wird („Wenn Dein einziges Werkzeug ein Hammer ist, wirst Du jedes Problem als Nagel betrachten“), zitiert nach Hubig, 2006, S. 158.

³⁶Krämer, 2008, S. 283.

³⁷Krämer, 1998a, S. 79.

Stimme mittransportiert werden, so z. B. der Klang oder die Stimmung. So sagt beispielsweise der Stimmenklang etwas über den Wahrheitsgehalt der Aussage aus. Die Stimme als Medium der Sprache bleibt uns, wie Krämer sagt, meist im Sprechen verborgen, hinterlässt aber ständig Spuren während des Sprechens.

Die Spur ist wie die inhaltlichen Aussagen der Kommunikation eine Wirkung des Agierens im medialen Raum, eine, die aber nicht vom Sprecher intendiert war. Krämer argumentiert: „Das Hinterlassen von Spuren ist zwar ein Effekt, nie aber Absicht, Ziel oder Zweck eines Handelns“.³⁸ In Krämers Terminologie ist die Spur die Rückseite der Botschaft, die auf deren Medialität verweist. „Das Medium ist nicht einfach die Botschaft; vielmehr bewahrt sich an der Botschaft die Spur des Mediums“.³⁹ Allgemeiner sind die Spuren die „nicht intendierten Hinterlassenschaften eines Tuns“, was jedes Agieren in einem Medium betrifft.⁴⁰

Anderes als die Aussage sagt uns die Spur nichts, so Krämer, sondern sie zeigt uns nur etwas.⁴¹ Das, was sie zeigt, ist das Medium und in diesem Sinne sind die Spuren Indikatoren des Mediums. Um Spuren als Hinterlassenschaften des Mediums zu begreifen, so die Annahme von Krämer, müssen diese interpretiert werden.

Das Interpretieren von Spuren als Spurenlesen kann beispielsweise Aufgabe einer sprachwissenschaftlichen Untersuchung sein, wie es Krämer in ihrer Arbeit darstellt.⁴² Das Lesen von Spuren ist eine Aktivität, die im Anschluss an das bewusste Agieren im Medium oder die Beobachtung dessen erfolgen kann, das heißt a posteriori.⁴³ Diese Interpretation wird als Spurenlesen bezeichnet und folge dabei einer „Logik der Narration“, die laut Krämer stets „unterschiedliche kohärente Erzählungen“ ermögliche.⁴⁴ Die Logik der Narration besagt hier, dass die Spuren als Verweisende (Indikatoren) einen Deutungsrahmen bilden, der verschiedene Schlussfolgerungen zulässt.⁴⁵

So wie ich ihre Ausführungen verstehe, führt Krämer das Konzept des Spurenlesens als Teil einer wissenschaftlichen Praxis ein, als eine Methode der Medientheorie, die sich mit der Wirksamkeit von Medien beschäftigt. Sie spricht von einer Suchbewegung des Spurenlesers, das ein „intensives praktisches und theoretisches Beschäftigen mit dem Material“ voraussetzt,

³⁸Krämer, 2008, S. 280.

³⁹Krämer, 1998a, S. 81.

⁴⁰Krämer, 1998b, S. 86.

⁴¹Krämer, 1998a, S. 79.

⁴²Krämer, 1998a, S. 79.

⁴³Krämer, 2008, S. 295.

⁴⁴Krämer, 1998b, S. 87.

⁴⁵Krämer, 1998b, S. 87.

um die Effekte einer Mediennutzung als Spuren zu interpretieren.⁴⁶

Eine solche Perspektive nimmt sie in etwa ein, wenn sie sich in ihren Ausführungen mit der Karte als Medium oder mit dem Computer als Medium (anonymer) Kommunikation, das heißt, sich kritisch mit Kommunikationskultur und Urheberschaft in ein einem globalen Netz (Internet) befasst.⁴⁷ In ihrer Arbeit zur Karte als Medium spricht sie von einer „medienkritischen Epistemologie“, in deren Zuge sie Überlegungen zur Verzerrung von Karten, die eine bestimmte Weltanschauung nahelegen, aus einer politischen Perspektive diskutiert.⁴⁸ Einen Aspekt von Medium und Spur, den sie in ihrer Arbeit nicht bespricht, ist die Rolle von Spuren und des Spurenlesens aus der Perspektive von Personen, die mit Medien umgehen. Aus dieser Perspektive können die Überlegungen zum Spurenlesen auch verwendet werden, um zu erklären, wie Personen ein Agieren in einem für sie neuem Medium erlernen. Das ist eine Sichtweise, die ich im folgenden Abschnitt besprechen werde, da dieser für ein Verständnis von Hybridisierung relevant ist (3.1.3).

Medium und Hybridisierung

Krämer spricht davon, dass sowohl natürliche als auch technische Medien im störungsfreien Vollzug verschwinden, das heißt, sie entziehen sich der Wahrnehmung. Das ist der Fall, wenn Mensch und Medium in einer hybriden Konstellation eine operative Einheit gebildet haben.

Ein Aspekt, den sie nicht betrachtet, der aber für ein Verständnis der Hybridisierung von Interesse ist, besteht darin, dass diese (weitgehende) Störungsfreiheit das Resultat einer Gewöhnung ist. Der störungsfreie Vollzug in einem Medium ist nicht ursprünglich gegeben, beispielsweise die Rede im Medium der Stimme musste in der Kindheit allererst erlernt werden. So werden Laute ausprobiert und nachgeahmt und ihre Wirkungen im Raum und beim menschlichen Gegenüber erfahren, lange bevor sie in der Rede entfaltet werden.

Krämer befasst sich demnach nicht mit der Entstehung, aus der die zuvor getrennten Entitäten zu einer hybriden Einheit hervorgehen. Sie interessiert nicht der Prozess des Verbindens, d. h. der Vorgang der Hybridisierung. Wie verbindet sich beispielsweise der Kartennutzer mit der Karte, damit die *operative Einheit* entsteht? Es ist nicht anzunehmen, dass es ein ursprüngliches Verschwinden der Karte im Vollzug gibt, sondern dieses vielmehr Ergebnis eines (Lern-)Prozesses ist. Auch wenn es nicht ihr vorrangiges Interesse ist, ist der Prozess der Entstehung von Interesse, um die

⁴⁶Krämer, 1998b, S. 87.

⁴⁷Krämer, 1998a, S. 86ff.

⁴⁸Krämer, 2008, S. 335.

Verteilung von Handlungsmacht in Mensch-Medien-Konstellationen als eine zentrale Frage der Medienphilosophie besser zu verstehen.

Im nächsten Abschnitt werde ich mich mit der Medien- und Technikphilosophie von Christoph Hubig befassen, der sich genau mit dem Prozess des Verbindens befasst (3.1.3). In diesem Zusammenhang wird die Spur diesmal nicht als direkte Erfahrungsgrundlage für eine wissenschaftliche Betrachtung der Wirksamkeit von Medien verstanden, sondern als Grundlage einer indirekten Erfahrung der Medien aus Sicht derjenigen, die im Medium handeln, aufgefasst.

3.1.3 Hybridisierung und die Rolle medialer Erfahrung: Spurenlesen als ein Abduzieren nach Christoph Hubig

Die Medienphilosophie, die Christoph Hubig entwickelt, ist Teil seiner Technikphilosophie. Sein Interesse an der Medienphilosophie ergibt sich daraus, dass er Technik als Medium versteht. Zu seinem Technikverständnis gehören nicht nur technische Apparate oder Geräte, sondern u. a. auch Methoden und Fähigkeiten (siehe 3.1.1).

Christoph Hubig charakterisiert das Medium in erster Hinsicht als einen „strukturierte[n] Möglichkeitsraum“, dessen Potentialität von Menschen als „veranlassende Instanzen“ entfaltet werden müssen.⁴⁹ Analog zu Krämer und anderen bereits genannten Autoren der Medientheorie ist für Hubig das Medium etwas, das sich einem direkten Zugang entzieht. Nach Hubig kann das Medium aber indirekt anhand von Spuren Gegenstand einer partiellen Vorstellung werden.

Die Idee einer Spur, die im Umgang mit dem Medium hinterlassen wird, spielt in seinem Ansatz vor allem eine Rolle, wenn es darum geht, wie Personen einen Umgang mit Medien bzw. ein Agieren im Medium praktisch erlernen. Dieser Aspekt, der sich von Krämers Ansatz unterscheidet und für eine Konzeption einer Hybridisierung von entscheidendem Interesse ist, wird im Folgenden diskutiert.

Rekonstruktion des technischen Mediums im Vollzug

Christoph Hubig argumentiert, dass sich das Mediale der Technik den Personen, die mit dem Medium umgehen, auf indirekte Weise zeigt. Der Modus, der das Medium partiell verfügbar macht, ist die „Medialitätserschließung“, die es genauer zu rekonstruieren gilt.⁵⁰ Wie Hubig sagt, handelt es

⁴⁹Hubig, 2006, S. 180 und S. 182.

⁵⁰Hubig, 2006, S. 160.

sich um ein „negatives Sich-Zeigen“ in Form von enttäuschten Erwartungen oder auch Überraschungen im Umgang mit einem Medium.⁵¹ Im Unterschied zu Krämer geht es Hubig nicht darum, ein primär wissenschaftliches Interesse am Verstehen von Medien zu folgen, sondern um die philosophische Entfaltung einer Medienpraxis, das sich im Umgang mit dem Medium aktualisieren kann.

Das kann in Anknüpfung an die präferierten Beispiele Krämers anhand eines technischen Kommunikationsmediums veranschaulicht werden. Personen tauschen mit Hilfe eines Telefons Informationen über Distanzen aus, wobei sie im Medium der Rede kommunizieren. Dabei kann es sein, dass eine Kommunikation mit Hilfe eines Telefons aus einen oder mehreren Gründen nicht gelingt. Dieses Nichtgelingen als Störung kann die buchstäbliche Übertragung betreffen, beispielsweise wenn die Verbindung des Mobiltelefons in bestimmten räumlichen Kontexten (Tiefgarage, Altbau) nicht möglich ist. Die Störung zeigt die Grenzen der Technik als Medium der Kommunikation auf und damit auch die Möglichkeiten eines Ungestörtseins – wenn man sich in den genannten Orten aufhält.

Im Erlebnis der Störung ist das Mediale der Technik, der oben genannte strukturierte Möglichkeitsraum – der sich hier in dessen Begrenzung zeigt – für die Beteiligten partiell erfahrbar. Hubig bezeichnet eine Störung, die wie im Beispiel identifiziert wurde, als eine (bloße) Irritation und gleichzeitig als Spur des Mediums, das heißt, sie trägt dazu bei, das Mediale dieser Technik zu erfahren.

„So hebt beispielsweise eine identifizierte Störung - eine Irritation - die mediale Intransparenz des Systems auf, und zeigt als »Spur« [...] die mediale Verfasstheit des Systems“.⁵²

Damit eine Technik als Medium im Sinne Krämers größtenteils störungsfrei genutzt werden kann und im Vollzug verschwindet, muss sich die Beziehung stabilisieren. Die identifizierte Störung ist ein Beispiel für eine Spur des Mediums, das heißt, sie ist Grundlage für eine Erfahrung des Mediums, die im Umgang mit einer *neuen* Technik oder einer bekannten Technik in einem neuen Handlungs- bzw. Nutzungskontext auftreten. Erst wenn die mediale Verfasstheit eines Systems in bestimmter Hinsicht erfahren wurde, kann sich das Verhältnis stabilisieren: Das Medium verschwindet in der Selbstverständlichkeit der Nutzung, als transparentes Mittel oder als Vermittler zwischen Personen. Ist dies der Fall, bilden die Personen und

⁵¹Hubig, 2006, S. 160.

⁵²Hubig, 2006, S. 176f.

die Kommunikationstechnik eine hybride Einheit. Hubig spricht hinsichtlich der Entstehung einer stabilen Mensch-Technik-Beziehung von einer „hybride[n] Konstruktion“.⁵³

Die Medialität eines technischen Systems zu verstehen, heißt, es hinsichtlich seines allgemeinen möglichen Wirksamwerdens in relevanten Kontexten näher zu bestimmen. Die Möglichkeit des Wirksamwerdens schließt u. a. die konkrete Nutzung für einen singulären Zweck, wiederholter Befriedigung von bestimmten Bedürfnissen bis hin zum Möglichmachen von bestimmten Lebensentwürfen mit ein.

Die Rolle der Spur in der medialen Erfahrung wird im Folgenden mit Blick auf Hubigs Konzept der Abduktion genau ausgearbeitet.

Abduzieren

Das abduktive Schließen ist eine Form praktischen Schließens, die vom induktiven und deduktiven (logischen) Schließen unterschieden wird. Im Gegensatz zum deduktiven Schließen wird nicht wie im klassischen Syllogismus aus den beiden Prämissen, der Fall (1) und die Regel(2), auf eine Conclusio (Resultat) (3) geschlossen.

Zum Beispiel:

(1): Es hat geregnet. (Fall)

(2) Wenn es regnet, wird die Straße nass. (Regel)

(3) Die Straße ist nass. (Resultat)

Dagegen wird beim abduktiven Schluss, wie er in der pragmatistischen Tradition bei Peirce eingeführt wurde, wie folgt argumentiert.⁵⁴ Es gibt Schlüsse, die in zentraler Weise unser alltägliches Handeln bestimmen, in denen von dem Resultat in umgekehrter Richtung auf den Fall oder die Regel geschlossen wird. Hinsichtlich der Beobachtung, dass die Straße nass ist und der angenommenen Regel, dass wenn es regnet, die Straße nass wird, wird geschlossen, dass es geregnet haben muss. Warum aber ist dieser Schluss nicht deduktiv? Weil er kein *sicherer* bzw. logischer Schluss ist, da es ja auch andere Gründe geben könnte, weshalb die Straße nass ist. Beispielsweise könnte ein angrenzender Bewohner etwas *unkontrolliert* gegossen haben oder ein Hydrant geöffnet worden sein, um für die Kinder der Nachbarschaft Abkühlung zu schaffen.⁵⁵ Oft gibt es gerade beim abduktiven Schluss eine wahrscheinliche Erklärung, die aus der Erfahrung rührt,

⁵³Hubig, 2006, S. 184.

⁵⁴Peirce, 1867.

⁵⁵Ikonisch: Fotos in New York der 70er.

die aber eben nicht logisch zwingend ist, aber unser Handeln, Wahrnehmen und Erkennen leitet.

Durch abduktives Schließen kann aber auch bei Kenntnis des Falls und des Resultats auf die zweite Prämisse, die Regel, geschlossen werden. Diese Arten der Abduktion werden als „hypostatische Abduktionen“ bezeichnet.⁵⁶ Ein kleines Kind kann, nachdem es vom Fenster aus den Garten beobachtet und gesehen hat, dass es im Garten regnet, später, wenn es zum Spielen auf die Straße geht, merken, dass die Straße nass ist, die Regel bilden: „Wenn es (im Garten) regnet, wird die Straße nass“. Der abduktive Schluss auf die Regel ist dem induktiven Schluss insofern gleich, als hier auf Regeln/Gesetze geschlossen wird. In der auf induktivem Schließen beruhenden empirischen Wissenschaft werden klassisch zwei Phänomene beobachtet (Observable (1): es regnet; Observable (2): die Straße ist nass) und dann überprüft, ob und wie diese zusammenhängen. Es müssen dabei die Bedingungen konstant gehalten werden und die Phänomene müssen wiederholt auftreten, bis etwas induktiv als allgemeine Regel gelten kann. Beispielsweise kann nach 1.000 Beobachtungen dieser Phänomene eine signifikante Korrelation angenommen werden. Dagegen ist die Abduktion in der wissenschaftlichen Praxis in der Hypothesenbildung zu finden, die bestimmte Regeln darüber, wie mehrere Phänomene zusammenhängen, vorschlägt, um sie dann für weitere (induktive) Untersuchungen in Frage zu stellen. Peirce spricht bei der Abduktion auch von einer *logic of discovery* oder einer *weit vorstoßenden Induktion*.⁵⁷

In der alltäglichen Lebenspraxis ist das abduktive Schließen die dominante Schlussform, die zur Anwendung kommt, wenn explizit oder, was weitaus häufiger vorkommt, implizit Hypothesen über Zusammenhänge aufgestellt werden und aufgrund dessen gehandelt wird. Ein Beispiel im Bereich der Wahrnehmungsabduktion soll das zeigen. Im alltäglichen Wahrnehmen sind Personen mit den *Regeln* des Wahrnehmens durch unsere Erfahrungen vertraut. Beispielsweise werden Abstände eingeschätzt (nicht in Zahlenwerten, sondern im Modus „Komme ich da durch?“ oder „Schaffe ich diesen Sprung?“), Wahrnehmung bei schlechten Lichtverhältnissen oder perspektivischer Verzerrungen kontextualisiert („das ist näher und deshalb größer“). Hubig sagt, dass eine Abduktion auf die Regel „im Wesentlichen ex negativo“ stattfindet.⁵⁸ Die *Regeln* bzw. das Schema, nach dem man allgemein wahrnimmt, wird allererst explizit, wenn man sich täuscht. Schritt für Schritt wird die Medialität des visuellen Wahrnehmungsapparats erschlossen. Beispielsweise dachte man, dieses Gebäude sei

⁵⁶Hubig, 2006, S. 210.

⁵⁷Peirce, 1867.

⁵⁸Hubig, 2006, S. 210.

viel größer als ein anderes, doch wenn man näher kommt, wird man sich der Täuschung bewusst. Die praktischen Fähigkeiten der Wahrnehmung sind somit abhängig von der Erfahrung. So können beispielsweise manche Berufsgruppen in bestimmten Wahrnehmungsaufgaben besser als andere geschult bzw. geübt sein. Das gilt für den eigenen Wahrnehmungsapparat, der von Geburt an trainiert wird, in gleicher Weise wie für die Nutzung von Techniken jeglicher Art.

Spurenlesen als Abduzieren

Es ist der abduktive Schluss, der den Modus beschreibt, mit dem Personen mediale Erfahrungen machen, das heißt auf die Verfasstheit des Mediums schließen können.

Das Medium kann, wie schon weiter oben angesprochen, in den „Aktualisierungen seiner Ermöglichungsleistung [...] (indirekt) zum Gegenstand einer Vorstellung werden“, wobei die Aktualisierung als die Nutzung der Technik zur konkreten Zweckerfüllung verstanden wird.⁵⁹ Das kann im „Rahmen von Überraschungs-, Enttäuschungs-, Widerfahrnis- und Gelingenserlebnissen“ geschehen, wobei diese Erlebnisse als Spuren des Mediums wirksam werden können, das heißt, dass sie die Basis dafür sein können, die Medialität partiell abduktiv zu erschließen.⁶⁰ Im Rahmen solcher Erlebnisse tritt das Mediale der Technik beispielsweise als *Widerständiges* in Erscheinung und wird so zum indirekten Gegenstand einer Vorstellung. Im Abschnitt zuvor wurde argumentiert, dass die Schemata, nach denen Personen wahrnehmen, das Mediale des natürlichen Wahrnehmungsapparats, sich vor allem in der Täuschung zeigt. Was für das Wahrnehmen und Hantieren ohne Technik gilt, lässt sich auch für die Technik in ihrer Rolle als Vermittler oder Mittel übertragen.

Die Möglichkeit eines Spurenlesens im Modus der Abduktion im Umgang mit Medien, im Folgenden beim Handeln mit Technik, wird durch eine sequenzielle Darstellung veranschaulicht. Techniknutzer haben bestimmte implizite und explizite Erwartungen, wie sie mithilfe einer Technik ein Ziel erreichen können. Die Erwartungen betreffen das Ziel und die Vorstellungen darüber, welches Agieren mit der Technik zu diesem Ziel führt. Der durch die Zielvorstellungen induzierte Handlungsvollzug führt zu einem Ergebnis, welches vom Nutzer hinsichtlich des vorgestellten Ziels erfahren wird. Je nachdem, ob das Agieren als geglückt aufgefasst wird oder nicht, kommt es zu den oben genannten Erlebnissen. Diese Erlebnisse können die Grundlage für Spuren bilden und verweisen dann auf das

⁵⁹Hubig, 2006, S. 221.

⁶⁰Hubig, 2006, S. 221.

Mediale der Technik, das heißt u. a. auf die Möglichkeiten und Grenzen eines Umgangs. Die Bedingungen für Spuren und das Lesen von Spuren wird anhand von Enttäuschungserlebnissen weiter betrachtet. Im Zuge eines Enttäuschungserlebnisses wird die Differenz zwischen der Erwartung und dem beobachteten Resultat als Störung empfunden. Die Differenz Erfahrung, die als Störung empfunden wird, ist eine Spur des Mediums, wenn sie im Kontext ihrer Entstehung – z.B. hinsichtlich des Wissens über bzw. der Beobachtung von zugrundeliegenden (u.a. physischen oder elektronischen) Wirkzusammenhängen – erfahren wird. Hubig spricht hierbei von einer identifizierten Störung, die eine mediale Erfahrung ermöglicht.

Versteht man das Spurenlesen als eine Abduktion, ergibt sich die folgende Schlussfolgerungspraxis: Im Fall eines Enttäuschungserlebnisses wird darauf geschlossen, wie der strukturierte Möglichkeitsraum (das Medium), den die Technik aufspannt, beschaffen sein muss, so dass das vorangegangene eigene Agieren im Umgang mit der Technik zu dem beobachteten Ergebnis und nicht dem gewünschten Ziel geführt hat. Die beobachtete Differenz wird von Hubig als *Spur von* bezeichnet, die zu einer (erweiterten) Vorstellung der Technik als Möglichkeitsraum führt und damit auf die Bedingungen möglicher (zukünftiger) Nutzung als eine *Spur für* verweist.⁶¹

Jeder anfängliche Umgang mit einer Technik, ich sprach von einem Erproben bzw. Erlernen einer Nutzung, wird u. a. durch Erlebnisse begleitet, in denen der Umgang gestört ist, das heißt, ursprüngliche Erwartungen nicht eintreten. Werden die Störungen im beschriebene Sinne identifiziert, tritt die Medialität in Erscheinung, die die Grundlage für eine weitere Nutzung bildet, in der das Agieren verfeinert wird bzw. Erwartungen an die Erfahrungen angepasst werden.

Zusammenfassend ist hervorzuheben, dass das Konzept der Spuren und des Spurenlesens geeignet ist, zu erklären, wie die Hybridisierung von Menschen durch Technik vonstattengeht. Die partielle Rekonstruktion des Mediums ist zentral, damit der Umgang mit einem technischen System nicht nur zu einem gewohnten, sondern zu einem selbstverständlichen Agieren wird, da so Möglichkeiten und Grenzen erfahren und deshalb implizit und explizit im Umgang gewusst werden. Mit der Selbstverständlichkeit des Umgangs als Ergebnis einer Hybridisierung verschwindet die Technik als Vermittelndes bzw. als Mittel im störungsfreien Vollzug.

⁶¹Hubig, 2011, S. 4.

3.1.4 Hybridisierung als Blackboxing: Medialität nach Bruno Latour

Bruno Latour befasst sich in seiner Soziologie, die auch technikphilosophische Standpunkte diskutiert, mit Hybriden und mit deren (sozialer) Vernetzung in Kollektiven. Hybride sind bei Latour Verbindungen von menschlichen und nicht-menschlichen Akteuren.⁶² Prominente Vertreter von Nicht-Menschen sind Techniken, die nach seiner Auffassung im engeren Sinne keine (bloßen) Objekte (im Sinne von technischen Artefakten oder Geräten) sind, sondern es handelt sich um Dinge, die Teil einer sozialen Praxis sind und in dieser Rolle (alltägliche) Beziehungen stabilisieren.⁶³

Das Kollektiv bezeichnet eine soziale Ordnung, in der Menschen und nicht-menschliche Akteure *zusammenleben*. Latour verwendet die Terminologie des Kollektivs, um seinen Ansatz von der klassischen Soziologie zu unterscheiden, in der Gesellschaften in dominanter Weise von Menschen bevölkert werden.⁶⁴ Nach seinem Verständnis der sozio-technischen Ordnung sind Menschen und Nicht-Menschen symmetrisch zu betrachten. Das steht der klassischen Trennung von dem Menschen als Subjekt entgegen, das heißt als die handelnde und auslösende Instanz gesellschaftlicher Aktivität und den Nicht-Menschen als passive Objekte. Latour beschäftigt sich mit der Frage, wie sich Hybride in Kollektiven formen und diese verändern. Diese Überlegungen zur Genese und Transformation von hybriden Konstellationen ist für die Fragestellung in diesem Kapitel zur Konzeption der Hybridisierung von Menschen durch Technik von Interesse.

Seine *Erzählung* zur Veränderung des hybriden Kollektivs beschreibt den „Eintrittspunkt der Nicht-Menschen in das Kollektiv“.⁶⁵ Entitäten, die nicht Teil des Kollektivs sind, können in bestimmter Hinsicht ausgehend von bestimmten Interessenlagen von Akteuren für das Kollektiv rekrutiert bzw. mobilisiert werden.⁶⁶ Werden sie so Teil des Kollektivs, treten sie als *eigenständige* Akteure in eine hybride Beziehung zu den anderen menschlichen und nicht-menschlichen Entitäten ein.

Latour diskutiert seinen Ansatz u. a. anhand eines Projekts, das das Ziel hatte, eine autonome Untergrundbahn für den Süden von Paris zu bauen.⁶⁷ Das Bahnprojekt fasst er als nicht-menschliche Entität auf, die in Form der Idee und als Entwurf in das Kollektiv eingetreten ist. Das geschah, indem

⁶²Wobei er bei Nicht-Menschen auch von Aktanten statt von Akteuren spricht, Latour, 2006, S. 488.

⁶³Latour, 2006, S. 501.

⁶⁴Latour, 2006, S. 504.

⁶⁵Latour, 2006, S. 504.

⁶⁶Latour, 2006, S. 504.

⁶⁷Latour, 2006, S. 506f.

sie von verschiedenen Interessenvertretern (Politikern, Institutionen, Stadtplanern, Industrievertretern) jeweils als etwas anderes rekrutiert bzw. mobilisiert wurde. Die U-Bahn trat dabei bereits in ihrer Form als Entwurf in eine soziale Interaktion mit Individuen und Institutionen ein.

Latour nennt weitere Beispiele, in denen er beschreibt, wie technische Objekte eine alltägliche soziale Rolle einnehmen. Eines davon ist der Hotelschlüssel, der durch ein Gewicht beschwert wurde. Der beschwerte Schlüssel veranlasst Gäste dazu, diesen beim Portier anzugeben, wenn sie das Hotel verlassen. Es handelt sich nach Latour bei diesem Schlüssel um einen hybriden Akteur, weil ihm eine soziale Norm eingeschrieben wurde, das heißt, ihm wurden „Eigenschaften unserer Sozialordnung inskribier[t]“. ⁶⁸ Die soziale Regel, den Schlüssel abzugeben, wenn man das Hotel verlässt, wird nicht durch das Pflichtbewusstsein des Gastes zu einer *gelebten* Norm, sondern über den Umweg der Bequemlichkeit, den schweren Schlüssel nicht herumtragen zu müssen, realisiert.

Es ist hervorzuheben, dass Latours Beispiele keine singulären Phänomene beschreiben, sondern dass Latour davon ausgeht, dass letztlich jedes Objekt Teil eines Kollektivs ist. Er sagt: „Objekte, die einfach als Objekte existieren, abgeschlossen, nicht Teil eines kollektiven Lebens, sind unbekannt“. ⁶⁹

Die Überlegungen Latours zu hybriden Akteuren in sozialen Kollektiven ist Teil seiner soziologischen Perspektive. Die „lokale Konstruktion“ von Hybriden, die für meine Fragestellung von besonderem Interesse ist, ist auch für Latour relevant, da, wie er sagt, Gesellschaft das Ergebnis lokaler Konstruktion sei. ⁷⁰ In meiner Besprechung von Latours Arbeit werden vor allem diese Passagen diskutiert.

Latour unterscheidet in diesem Zusammenhang vier aufeinander aufbauende Weisen der technischen Vermittlung (engl. Original: *mediation*) im Verhältnis von Mensch und Technik.

Seine Abhandlung wird von vielen Beispielen begleitet. Das bekannteste ist das von Mensch und Schusswaffe. Anhand dessen veranschaulicht er die Unterscheidung zwischen der materialistischen und der instrumentalistischen Perspektive in Bezug auf Technik. Die materialistische Sichtweise sieht in einer Schusswaffe und der Möglichkeit, diese leicht zu bekommen, die Gefahr, dass „der gute Mann und gesetzestreue Bürger gefährlich“ wird. ⁷¹ Die Waffe zwingt dem Träger, oder eine Person, die Zugang zu ihr hat oder bekommen könnte, ihren Willen auf. Der Zugang zu Waffen macht

⁶⁸Latour, 2006, S. 504.

⁶⁹Latour, 2006, S. 503.

⁷⁰Latour, 2006, S. 510.

⁷¹Latour, 2006, S. 485.

gute Bürger zu Mördern.

Dagegen steht die neutrale Position, die unter andern von der NRA (*National Rifle Association*) vertreten wird. Nach dieser Position wird der Schusswaffe keine Wirksamkeit zugeschrieben, die jenseits menschlicher Akteure liegt. Demnach ist es so, dass, wenn der gute Bürger der Waffenbesitzer ist, die Waffe auch nur zum Guten eingesetzt wird und umgekehrt. Die NRA macht die Waffe zu einem „neutralen Willensträger, der der Handlung nichts hinzufügt“.⁷² Will eine Person töten, wird sie es mit oder ohne Schusswaffe tun. Mit den Fäusten oder dem Messer ist es „schlicht langsamer, schmutziger, widerlicher, [...] zu keiner Zeit verändert sich die Absicht“.⁷³ Latour ist der Meinung, dass keine der Erzählungen richtig ist. Seiner Auffassung nach liegt die Wahrheit vielmehr in dem vermittelten Zusammenwirken zwischen Mensch und Technik, in der geteilten Verantwortlichkeit, in der Wirksamkeit eines neuen hybriden Agenten, der aus Mensch und Waffe besteht.⁷⁴

Übersetzung und Komposition von Handlungsprogrammen

Im Aufeinandertreffen von Mensch und Technik entsteht ein neuer, hybrider Agent, in dem sich Ziele des Menschen und der Technik verbinden und die Ziele des neuen Agenten (meist) weder die der Technik sind noch die des Menschen. Vielmehr findet eine Modifikation bzw. Verschiebung der Ziele statt, die Latour als Übersetzung bezeichnet.

Mensch und Technik wirken zusammen, indem im Zuge einer konkreten Handlungssituation die Ziele beider Entitäten, Latour spricht von Agenten bzw. Aktanten, ineinander übersetzt werden, wobei, wie Latour sagt, beide Agenten modifiziert werden.⁷⁵ Dieses Übersetzen ist so zu verstehen, dass sich die Handlungsprogramme des einen hinsichtlich denen des anderem in den Situationen des Aufeinandertreffens aktualisieren. Er verwendet den Begriff der Handlungsprogramme, um eine, wie er sagt, neutrale Bezeichnung zu haben, unter der Funktionen von Technik und Ziele von Menschen gleichermaßen subsumiert werden können.⁷⁶

Handlungsprogramme können als strukturierte Räume möglichen Agierens verstanden werden. Beim Menschen sind die Handlungsprogramme, als Ergänzung zu Latours Ausführung, strukturierte Handlungsmöglichkeiten, die durch Gewohnheiten, Fähigkeiten und Vorlieben festgelegt sind und bei der Technik sind es die Fülle von Funktionen, nicht nur im

⁷²Latour, 2006, S. 485.

⁷³Latour, 2006, S. 486.

⁷⁴Latour, 2006, S. 488.

⁷⁵Latour, 2006, S. 487.

⁷⁶Latour, 2006, S. 488f.

Sinne von vorgegebenen Funktionen, sondern alle Möglichkeiten, die der Umgang mit Technik und Artefakten bietet.

Konkret bedeutet das, dass beispielsweise der menschliche Agent auf Grundlage dessen modifiziert wird, was die Handlungsprogramme der Technik sind, das heißt die Möglichkeiten eines Umgangs, mit denen der Mensch in ein Verhältnis eintreten will oder muss. Latour spricht davon, dass die menschlichen und nicht-menschlichen Akteure in einem Austausch von Kompetenzen stehen, indem sie einander neue Möglichkeiten, Ziele oder Funktionen anbieten.⁷⁷ Der Akteur verändert sich dahingehend, dass beispielsweise ein bisher privilegiertes Handlungsprogramm nicht mehr aktualisiert wird, sondern ein zuvor latentes Handlungsprogramm handlungsleitend wird. Der Hotelschlüssel in Latours bekanntem Beispiel ist mit einem Gewicht beschwert und wird deshalb abgegeben, weil das bevorzugte Handlungsprogramm, die volle Flexibilität beim Betreten und Verlassen des Zimmers zu haben, dadurch, dass der Schlüssel durch das Gewicht lästig geworden ist, das Handlungsprogramm, den Schlüssel immer an der Rezeption abzugeben, präferiert wird. Das Ergebnis dieser Übersetzung als Aktualisierung der verbundenen Handlungsprogramme ist das gemeinsame Agieren. Die Handlungsmöglichkeiten und konkreten Handlungen sind „Eigenschaft[en] des gesamten Gefüges“ und „nicht nur der menschlichen Aktanten“.⁷⁸

Das Beispiel mit der Schusswaffe und dem Menschen dient wieder zur Veranschaulichung. Man kann sich Situationen vorstellen, in denen Personen selbstbewusst reagieren und beispielsweise einem Streit nicht aus dem Weg gehen, wenn sie wissen, dass sie eine Waffe am Gürtel tragen. Das Handlungsprogramm der Waffe, das heißt mit Drücken des Abzugs eine Kugel abzufeuern, die töten kann, verbindet sich mit den Handlungsprogrammen der Person. Es ist ein gemeinsames Agieren von Mensch und Waffe, welches sich im Auftreten der Person im Streit zeigt, ohne dass die Waffe je gezogen worden ist. Es gibt immer explizite sowie implizite *privilegierte* Handlungsprogramme und eine Fülle von latenten Handlungsprogrammen. Diese beeinflussen das Agieren und Reagieren von Personen in bestimmten Situationen. In der beschriebenen Situation kann die Waffe am Gürtel eine vormals latente Präferenz aktualisieren, was sich in der Modifikation der konkreten Absicht bzw. die Wahl der Mittel zeigt. Ein verbaler Seitenhieb des Gegenübers kann beim Waffenträger zu einer Reaktion in Form einer Beleidigung führen, statt ignoriert zu werden, wie es die privilegierte Handlungsoption ohne Waffe gewesen wäre.

⁷⁷Latour, 2006, S. 490.

⁷⁸Latour, 2006, S. 492.

Blackboxing

Die Komposition neuer Handlungsprogramme von Menschen und Nicht-Menschen resultiert in hybriden Akteuren, deren Aktivitäten in sozio-technische Kollektive eingebettet ist. Latour stellt die Frage, warum die nicht-menschlichen Entitäten, wenn sie doch ständig in hybride Konstellationen mit Menschen eintreten, nicht in ihrer vermittelnden Rolle mit „einiger Präzision einzuschätzen“ sind.⁷⁹ Den Grund sieht Latour darin, dass das gemeinsame Handeln der hybriden Akteure, das heißt die Komposition von Handlungsprogrammen, „völlig undurchsichtig“ sei.⁸⁰ Nach Latour unterliegt die Komposition von Handlungsprogrammen einem Prozess, den er als „Blackboxing“ beschreibt.⁸¹

Um das Blackboxing zu veranschaulichen, benutzt Latour als Beispiel die Verwendung eines Tageslichtprojektors. „Er ist ein Punkt in einer Handlungsfolge (sagen wir einmal in einer Vorlesung), eine ruhige und stumme Vermittlungsinstanz, die für selbstverständlich gehalten und vollkommen von ihrer Funktion bestimmt wird“.⁸² Latour führt sein Beispiel weiter und beantwortet die Frage, was denn passiere, wenn der Projektor plötzlich nicht mehr funktioniere, mit der Ausführung, dass allererst dieses Nichtfunktionieren uns wieder an die Existenz des Projektor erinnere.⁸³ Der Projektor wird als eine Komposition von verschiedenen Teilen mit je „relativ unabhängigen Zielen“ (Funktionen) sichtbar. Die Blackbox, die das gemeinsame Agieren von Mensch und Technik betraf, hat sich zu einem bestimmten Grade aufgelöst.

Führt man die Überlegungen des Blackboxing auf die Idee der Hybridisierung zurück, ist das Blackboxing etwas, das mit der Hybridisierung einhergeht. Das Verschwinden der Technik als transparente und stumme Vermittlungsinstanz im Handlungsvollzug ist ein Symptom eines stabilen Mensch-Technik-Hybriden. Die Hybridisierung geht mit einer Übersetzungstätigkeit hinsichtlich der Handlungsprogramme der beiden Akteure einher und resultiert in einer neuen hybriden Verbindung. Der Hybrid besitzt einen eigenen strukturierten Raum an Möglichkeiten, in bestimmter Hinsicht in ein Verhältnis mit andern sozialen menschlichen oder nicht-menschlichen Akteuren zu treten, das heißt, er agiert auf Basis von Handlungsprogrammen, die, wie Latour sagt, „keinem der Handlungsprogramme der Agenten entspricht“.⁸⁴

⁷⁹ Latour, 2006, S. 491.

⁸⁰ Latour, 2006, S. 491.

⁸¹ Latour, 2006, S. 491.

⁸² Latour, 2006, S. 491.

⁸³ Latour, 2006, S. 491.

⁸⁴ Latour, 2006, S. 487.

Delegation

Unter dem Begriff *Delegation* diskutiert Latour einen weiteren Aspekt von Vermittlung. Im Modus der Delegation werden materiellen Entitäten Bedeutungen eingeschrieben („inskribiert“), das heißt, sie werden Träger von Eigenschaften der Gestalter, Auftraggeber und Herstellenden.⁸⁵

Ein Beispiel hierfür ist eine Bodenschwelle auf einem Campusgelände, „die Fahrer zwingt, auf dem Campus das Tempo zu reduzieren“.⁸⁶ „Das Ziel der Fahrer wird mittels der Bodenschwelle übersetzt von »Ich fahre langsamer, um keine Studenten zu gefährden« in »Ich fahre langsam und schütze die Federung meines Autos«“.⁸⁷ Der Fahrer, so Latour, modifiziert sein Verhalten wie von der Rektorin der Universität gewünscht, die die Bodenschwelle installieren ließ, aber nicht etwa aus moralischem Gewissen, sondern aus Eigennutz.

Das Handlungsprogramm eines Autofahrers, schnell an sein Ziel kommen zu wollen, wird weniger relevant hinsichtlich der neuen Präferenz, den Stoßdämpfer schützen zu wollen, welche stattdessen aktualisiert wird. Dass das Langsamfahren zum privilegierten Handlungsprogramm wird, ist genau das Ziel, das durch das Instrument der Bodenschwelle realisiert wird.

Wendet man Latours Überlegungen auf die Praxis der Gestaltung von Technik an – Latour würde statt von Gestaltung von einer Rekrutierung für das Kollektiv sprechen⁸⁸ –, sind alle von Menschen entworfenen und hergestellten Dinge hybride Agenten. Im Hybriden verbinden sich Eigenschaften in Form bestimmter Bedeutungen, die von Gestalter und der Rektorin, die sich einen jeder Zeit bereiten Verkehrspolizisten wünscht, stammen, mit den Eigenschaften der Materie. Die Betonschwelle vereint einige der Eigenschaften des Polizisten mit einigen des Betons, aber nicht alle.⁸⁹ Die Begrifflichkeit Rekrutierung, die Latour für die Delegation wählt, soll deutlich machen, dass gestaltete Artefakte als Techniken nicht immer an ihre Gestalter bzw. Auftraggeber gebunden sind, sondern dass sie ein neuer Akteur sind und eine eigene Rolle im Kollektiv einnehmen.

Verschiedene Autoren sehen den Aspekt der Delegation kritisch, da er für sie im Widerspruch zu dem von ihm propagierten symmetrischen Verständnis von Mensch und Technik steht.⁹⁰ Problematisch am Konzept der Delegation ist aus ihrer Perspektive der Umstand, dass die symmetrische

⁸⁵Latour, 2006 S. 494.

⁸⁶Latour, 2006 S. 494.

⁸⁷Latour, 2006 S. 494.

⁸⁸Latour, 2006, S. 492.

⁸⁹Latour, S. 497f.

⁹⁰unter anderem Verbeek, 2005, S. 160f.

Konzeption nicht aufrechterhalten wird und die technische Wirksamkeit scheinbar auf die Intentionen von Entwicklern reduziert wird.

Die Delegation stellt aber für Latours Konzeption keinen Widerspruch zu seiner symmetrischen Sichtweise dar, da auch Menschen durch andere Menschen oder nicht-menschliche Agenten für bestimmte Zwecke rekrutiert bzw. mobilisiert oder als wertvolle Mitglieder der Gesellschaft geformt werden. In gleicher Weise wie bei Menschen kann man bei Techniken im Sinne Latours davon sprechen, dass sie in ein eigenes gesellschaftliches *Leben* eintreten, wo sie in einem ursprünglich intendierten und auch nicht-intendierten Sinne wirksam werden. Mit der Zeit sind dann auch die Sprecher dieses technischen Aktes verschwunden.⁹¹

Als hybride Akteure versteht Latour sowohl operative Einheiten von Mensch und Technik, wie es beispielsweise die Verbindungen von Waffenbesitzer und Waffe, Chemiker und Pipette und die Verbindung Dozent-Projektor-Studenten zeigen, aber auch Entitäten, in denen sich im beschriebenen Sinne der Delegation Eigenschaften und Interessen von verschiedenen Akteuren materialisieren, ohne die *Ziele* des neuen hybriden Akteurs auf die Initiative von einem der Beteiligten reduzieren zu können. In diesem zweiten Verständnis geht die Konzeption über die in der Arbeit untersuchte hybride Konstellation hinaus.

Fazit: Hybride Konstellationen und deren Entstehung

In Abschnitt 3.1.2 wurde die Kritik Krämers an der avantgardistischen Position der Medientheorie im Sinne McLuhans vorgestellt, der sie vorwirft, dass sie Medien anthropomorphisiere.⁹² Auch Latour könnte von Krämer aufgrund seiner symmetrischen Anthropologie der avantgardistischen Position zugeordnet und diesbezüglich kritisiert werden. In der Beschreibungssprache, mit der Latour Technik beschreibt, finden sich zweifellos eine Reihe von Begriffen, die ausschließlich in der Sphäre menschlicher Eigenschaften oder menschlicher Vermögen zugehörig sind.

Doch ein solcher Vorwurf würde ins Leere gehen: Nicht wie beim klassischen Vorwurf einer Anthropomorphisierung, bei dem kritisiert wird, dass der menschliche Standpunkt (die Selbstkonstruktion) implizit immer leitend für eine Konzeption ist und der Standpunkt selbst nicht reflektiert

⁹¹Latour, 2006, S. 496.

⁹²Krämer, 1998a, S. 76.

wird, sondern die Begrifflichkeiten menschlichen Vermögens sind eine Beschreibungssprache für eine solche Reflexion des Standpunkts. Er *anthropomorphisiert* nicht-menschliche Akteure, aber vollzieht gleichzeitig eine *Verdinglichung* von Menschen, indem er ihr Handeln hinsichtlich ihrer (sozialen) Funktionalität – im Sinne von alltäglichen Routinen – stark macht. Und diese Routinen sind Teil eines alltäglichen gesellschaftlichen, besser kollektiven Handelns. Latour sagt: „Zweckgerichtetes Handeln und Intentionalität mögen nicht Eigenschaften von Objekten sein, aber sie sind auch keine Eigenschaften von Menschen“.⁹³ Damit meint er, dass Menschen nur in ihrer sozialen Rolle existieren und Techniken überhaupt nur Techniken sind, wenn sie in einer Austauschbeziehung stehen, die sie zur „Existenz autorisiert und befähigt“.⁹⁴ Menschliche und nicht-menschliche Operationen sind in der Form von Hantieren, Agieren, Handeln oder Funktionieren substituierbar. Durch diese Verschiebung findet er einen Weg, Mensch und Technik symmetrisch zu verbinden und trägt damit zu einem allgemeinen Verständnis von Mensch-Technik-Verhältnissen bei, das eben gerade nicht anthropomorph ist.

Hinsichtlich der bisherigen Überlegungen lässt sich sagen, dass die Hybridisierung eine – immer vorläufige – Stabilisierung von Mensch-Technik-Beziehungen ist, so dass ein selbstverständlicher Umgang mit der Technik möglich ist, der mit einem Verschwinden der Technik als Medium einhergeht. Mit Blick auf das Beispiel des Tageslichtprojektors sind der Dozent, der Projektor und die Studenten in einer hybriden Konstellation, in der die Komposition von Handlungsprogrammen, man kann sagen, die Verteilung von Handlungsanteilen, in der Praxis der Vorlesung nicht präsent sind. Der Projektor ist als stumme Vermittlungsinstanz in dieser Hinsicht in den Hintergrund der Wahrnehmung gerückt. Die Hybridisierung ist deshalb nie völlig abgeschlossen, da es immer neue Handlungssituationen geben kann, in denen sich die Medialität der Technik auf neue Weise zeigt.

3.1.5 *Trivialität* nicht-trivialer Maschinen: Heinz von Foersters Maschinenmetapher

Die Maschinenmetapher der trivialen und nicht-trivialen Maschine, die Heinz von Foerster in seiner kybernetischen Theorie entwickelt, wird im Folgenden dazu verwendet, die Hybridisierung von Menschen durch Technik aus einer weiteren Perspektive zu betrachten. Auch wenn in seiner Abhandlung hauptsächlich von Maschinen die Rede ist und auch Techniken als Beispiele dienen, handelt es sich dabei nicht um eine Maschinentheorie,

⁹³Latour, 2006, S. 503.

⁹⁴Latour, 2006, S. 503.

sondern vielmehr um ein Erklärungsmodell oder eine Metapher, die von Foerster gleichermaßen auf Technik, Menschen, Sozialsysteme oder die Natur anwendet.

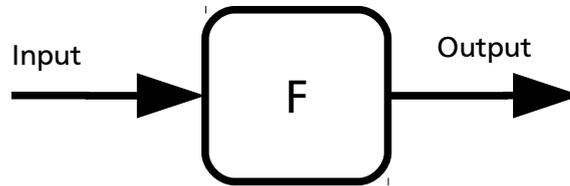


ABBILDUNG 3.1: Das Übertragungsschema der trivialen Maschine, wie es bei Heinz von Foerster dargestellt ist. Siehe Von Foerster, 2002, S. 245

Im Folgenden wird die Unterscheidung von trivialer und nicht-trivialer Maschine vorgestellt, um sie im Anschluss hinsichtlich des Umgangs von Menschen mit intelligenten Systemen zu betrachten. Die triviale Maschine ist durch ein fest gekoppeltes Ein- und Ausgabeverhältnis festgelegt. In Abbildung 3.1 ist die triviale Maschine schematisch dargestellt. Wenn man einen bestimmten Input (y) in die Maschine einbringt, reagiert die Maschine mit einem entsprechenden Output (y). Das heißt, die Übertragungsfunktion (F) der Maschine gewährleistet ein festes Input-Output-Verhältnis. Heinz von Foerster definiert die Eigenschaften einer trivialen Maschine anhand von vier Aspekten⁹⁵:

- (1) vorhersagbar
- (2) geschichtsunabhängig
- (3) synthetisch deterministisch
- (4) analytisch deterministisch

In Bezug auf eine einfache Alltagstechnik kann man sagen: Wenn man einen Knopf drückt, beispielsweise den Lichtschalter an der Schalterleiste *oben*, geht das Deckenlicht an, wenn man den Lichtschalter *unten* drückt, geht das Licht beim Bett an. Triviale Maschinen in diesem allgemeinen Verständnis begegnen uns überall im Alltag und die Trivialität ist, nach von Foerster, eine Eigenschaft, für die wir Technik schätzen. Wenn man den Lichtschalter drückt, möchte man nicht, dass man *versehentlich* beim Nachbarn klingelt oder das falsche Licht anschaltet. Wenn man auf den Knopf des Kaffeevollautomaten bei *Kaffee schwarz* drückt, dann möchte man einen schwarzen Kaffee und nicht einen Cappuccino.

⁹⁵Von Foerster, 2002, S. 246f.

Die Aktivität einer trivialen Maschine ist vorhersagbar, wenn sie nicht defekt ist und einem nach kurzer Eingewöhnung vertraut ist. Das ist sie vor allem, weil sie nicht geschichtsabhängig ist, das heißt, dass einem bestimmten Input *immer* ein bestimmter Output folgt. Synthetisch deterministisch besagt, dass es eine Logik hinter dem Input-Output-Verhältnis gibt, das heißt, dass die Maschine nicht willkürlich oder chaotisch agiert – die Logik aber potentiell von niemanden gekannt werden muss. Die letzte Eigenschaft, analytisch deterministisch, ist abhängig von der Perspektive. Analytisch deterministisch ist eine Maschine, wenn es einem Außenstehenden, der nur Zugang zu den Inputs und Outputs der Maschine hat, potentiell möglich ist, auf die Regeln der Übertragung (Übertragungsfunktion) zu schließen. Heinz von Foerster sagt zu der letzten Eigenschaft folgendes: „Ein besonders netter Zug dieser Maschinen ist, daß sie auch analytisch determinierbar sind, denn man braucht ja nur für jedes gegebene x [Input] das entsprechende y [Output] aufzuschreiben“.⁹⁶ Die Unterscheidung zwischen analytisch und synthetisch speziell in v. Foersters Ansatz kann mit Blick auf technische Systeme wie folgt erklärt werden. Aus Sicht des Gestalters einer Maschine sind die festen Regeln, die die Voraussetzung für eine synthetische Determiniertheit sind, im Idealfall die, die von ihm festgelegt sind und werden somit auch von ihm gewusst. Ein Nutzer, der keinen Einblick in die Blackbox der Maschine hat, kann dagegen nur von außen auf die Funktionsweise schließen. Ist das praktisch möglich, ist sie analytisch determiniert.

Das gegenteilige Konzept ist die nicht-triviale Maschine. Diese zeichnet sich dadurch aus, dass sie eine interne Funktion und einen internen Zustand besitzt, der sich durch Inputs verändert, dessen Funktionsweise und Wert aber nicht von außen zugänglich ist. In Abbildung 3.2 ist die nicht-triviale Maschine dargestellt. Die interne Funktion Z ist abhängig von den Inputs und verändert ihren internen Zustand z . Dieser interne Zustand beeinflusst die zukünftige Nutzung, das heißt, wie die Inputs in die Outputs gewandelt werden, und wird dabei selbst stetig verändert. Es entsteht somit eine Geschichtsabhängigkeit, das Verhalten der Maschine ist abhängig von der Historie der Nutzung. Damit ist es schwierig, Vorhersagen zu treffen, in welcher Weise bestimmte Inputs zu bestimmten Outputs führen, vorausgesetzt, man kennt die Funktionsweise der internen Funktion und/oder die Nutzungshistorie nicht. Trotz einer synthetischen Determiniertheit (durch feste interne Funktionsregeln) ist es für Außenstehende schon bei wenigen möglichen Inputs und Outputs praktisch unmöglich, auf die Verfasstheit dieses Systems zu schließen.

⁹⁶Von Foerster, 2002, S. 247.

Heinz von Foerster zeigt die praktische Unmöglichkeit, indem er eine Rechnung aufstellt, die es schon bei vier zweiwertigen Inputs und einem zweiwertigen Output 10^{10^4} Versuche benötigen würde, um sicher auf die Übertragungsfunktion schließen zu können.⁹⁷ Das ist mit heutigen und auch zukünftigen Computern nicht berechenbar. Die Eigenschaften der nicht-trivialen Maschine lassen sich wie folgt zusammenfassen.⁹⁸:

- (1) unvorhersagbar
- (2) geschichtsabhängig
- (3) synthetisch deterministisch
- (4) analytisch indeterminierbar

Dazu noch einmal das Beispiel der Lichtschalter: Es wäre vorstellbar, wenn auch unsinnig, dass man die Lichtschalter und Lichter hinter Wand und Decke so verschaltet, dass bei jedem Drücken des oberen Schalters ein Zähler mitläuft, der bei Wert X die Belegung der Lichtschalter umdreht. Man könnte sich noch weitaus kompliziertere nicht-triviale Maschinen vorstellen.

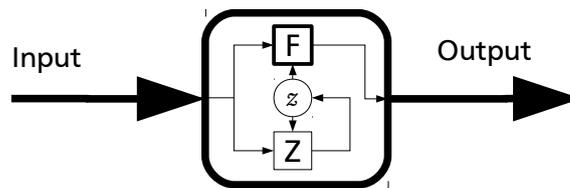


ABBILDUNG 3.2: Das Übertragungsschema der nicht-trivialen Maschine, wie es bei Heinz von Foerster dargestellt ist. Siehe Von Foerster, 2002, S. 248

Trivialitätserwartung

Wie lässt sich diese Unterscheidung hinsichtlich der Fragestellung der Arbeit fruchtbar machen? Dazu wird v. Foersterns weitergehendes Interesse an dieser Unterscheidung aufgezeigt. Von Foerster argumentiert, dass sich Menschen triviale Maschinen wünschen: „Wenn wir den Zündschlüssel des Autos drehen, dann muss das Auto starten“ und wenn es „trotzdem nicht-triviale Tendenzen zeigt, ein Auto also zum Beispiel nicht starten will usw., dann rufen wir einen Trivialisierungsspezialisten“.⁹⁹ Und an anderer Stelle: „In der Tat zielen alle Bemühungen nur darauf, triviale Maschinen zu

⁹⁷Von Foerster, 2002, 250.

⁹⁸Von Foerster, 2002, 251.

⁹⁹Von Foerster, 2002, S. 252.

erzeugen, oder dann, wenn wir auf nicht-triviale Maschinen treffen, diese in triviale Maschinen zu verwandeln“.¹⁰⁰ Nach Heinz von Foerster ist es für eine *erfolgreiche* Technik relevant, dass sie für Menschen als triviale Maschinen funktionieren. Wenn das nicht der Fall ist, wird das als Problem wahrgenommen und bei einem *Spezialisten* nach Hilfe gesucht. Andreas Kaminski, der sich mit den Konzepten Heinz von Foersters technikphilosophisch befasst, bezeichnet den Wunsch nach trivialer Technik als Trivialitätserwartung, wobei er diese Erwartung als grundlegend für die Rolle der Technik als u. a. verlässliche, stabile und entlastende Kraft auffasst.¹⁰¹ Im Folgenden wird diskutiert, wie es sich mit der Trivialitätserwartung im Umgang mit intelligenter Technik verhält.

Trivialität im Umgang mit intelligenter Technik ist deshalb klärungsbedürftig, da diese Techniken in ihrem *Normalzustand* nicht-triviale Tendenzen aufweist. Eine nicht-triviale Tendenz ist beispielsweise bei der technikbasierten *kontextsensitiven Interaktion* zu beobachten, wenn gemäß für den Nutzer unbekannter Mechanismen der Nutzer selbst und seine Umgebung gedeutet werden und darauf basierend die Möglichkeiten eines Umgangs angepasst werden (siehe den Abschnitt 2.3.1). Andreas Kaminski bezeichnet beispielsweise Individuen „als nichttriviale Maschinen per excellence“, da sie „auf kreative Weise lernen und [...] daher vergangenheitsabhängig“ seien.¹⁰² Intelligente Techniken sind zwar keine Individuen, können aber in ähnlicher Weise als geschichtsabhängig gesehen werden, da sie sich beispielsweise an die Präferenzen bzw. die Nutzungshistorie von Personen anpassen können.

Kaminski argumentiert, dass Heinz von Foerster Trivialität auf zwei Arten gebraucht, zum einen als ein Attribut des Gegenstands, demnach Trivialität eine „bloße Eigenschaft einer Sache“ ist und zum anderen als Inhalt einer Erwartung.¹⁰³ Eine Unterscheidung, die Heinz von Foerster nicht explizit vornimmt, sondern vermischt. Trivialität als Erwartung kann während der Nutzung entstehen, im „Prozess des miteinander Vertrautwerdens“.¹⁰⁴ Dies ist auch dann möglich, wenn die Technik faktisch nicht-triviales Verhalten aufweist.

Kaminski zeigt, dass sich Techniken oftmals strenggenommen nicht-trivial verhalten, beispielsweise, wenn sie temporär kaputt sind, weil Verschleißteile ausgefallen sind oder in einem Computer eine Programmschleife festgefahren ist. Handelt es sich um eine Disfunktionalität, von der bekannt ist, welche Mittel nötig sind, um die triviale Funktionalität wieder

¹⁰⁰Von Foerster, 2002, S. 207.

¹⁰¹Kaminski, 2010, S. 176.

¹⁰²Kaminski, 2010, S. 174.

¹⁰³Kaminski, 2010, S. 175.

¹⁰⁴Kaminski, 2010, S. 184.

herzustellen, beispielsweise indem Verschleißteile ausgetauscht oder der Computer neugestartet wird, kann trotzdem von einer Trivialität auf der Ebene der Erwartung gesprochen werden. Es handelt sich hierbei um „Trivialitätserwartungen zweiter Stufe“, da sich die Trivialitätserwartungen nicht auf das Input-Output-Verhältnis auf die Sachebene beziehen, sondern auf die Möglichkeiten eines zuverlässigen Managens des Funktionierens erster Stufe.¹⁰⁵

Diese Überlegungen lassen sich auf intelligente Techniken übertragen. Hier besteht, wie oben gezeigt, auch eine Nicht-Trivialität auf der Sachebene. Eine höherstufige Trivialitätserwartung (zweiter Stufe) kann es mit intelligenten Techniken geben, wenn die konkreten Resultate technischen Handelns, die eventuell schwer vorhersagbar sind, hinsichtlich ihrer alltäglichen Rolle als verlässliche und entlastende Kraft als trivial erscheinen.

Die Ausführungen zeigen, dass Trivialität und Nicht-Trivialität eine relevante Unterscheidung im Rahmen meiner Arbeit darstellt, wenn sie nicht nur als Eigenschaften der Technik aufgefasst wird, sondern als etwas, das das Verhältnis des Menschen zur Technik beschreibt. Eine Trivialisierung, die als Prozess eines Vertrautwerdens beschrieben werden kann, kann als etwas aufgefasst werden, das der Hybridisierung von Mensch und Technik entspricht. Das heißt, Prozesse der Vertrauensbildung durch Trivialisierungen unterstützen die Entwicklung hin zu einem stabilen Mensch-Technik-Verhältnis, in dem die Zielerwartungen und deren Erreichung in den relevanten Kontexten der Nutzung selbstverständlich geworden sind.

Im folgenden zweiten Teil des Kapitels werden Ansätze aus dem Bereich der Phänomenologie des Leibes für ein erweitertes allgemeines Verständnis dieses Prozesses der Hybridisierung des Menschen durch Technik behandelt.

3.2 Verleiblichung im Handeln mit Technik

Im zweiten Teil dieses Kapitels werden phänomenologische Überlegungen zum Mensch-Technik-Verhältnis vorgestellt und hinsichtlich der Konzeption einer Hybridisierung von Menschen durch Technik betrachtet.

Eine zentrale Idee der phänomenologischen Theorien ist das *embodiment*, was im Deutschen treffender als Verleiblichung bezeichnet wird. Das Konzept der Verleiblichung hat in vielen Fachdisziplinen *Karriere gemacht*: Unter anderem in der Linguistik, der Soziologie und auch in der HCI. Im Folgenden werden zwei der wichtigsten phänomenologischen Positionen zum Verhältnis von Mensch und Technik herausgegriffen, von Maurice

¹⁰⁵Kaminski, 2010, S. 176.

Merleau-Ponty und Martin Heidegger, die hinsichtlich einer Konzeption der Hybridisierung diskutiert werden. Im Anschluss werden die Überlegungen um einen Ansatz der neueren amerikanischen Tradition der Phänomenologie ergänzt.

3.2.1 Einführung

Mit der Phänomenologie, die am Anfang des zwanzigsten Jahrhunderts ihren Siegeszug begann, gewann die Körperlichkeit des Menschen für ein Verständnis menschlicher Weltbezüge stark an Bedeutung.¹⁰⁶ Der Körper wird seit je her als der Teil des Menschen gesehen, der aufgrund seiner räumlichen Präsenz in der Welt verortet ist. Über ihn wird die Welt durch die Sinne als Wahrnehmung vermittelt. Es ist der Körper in der Welt, der diese sensorisch erfasst. Die sensorischen Reize werden ihrerseits im Gehirn interpretiert, das heißt geordnet und bewertet. Diese Interpretation bietet dann die Grundlage für ein Wissen über oder die Erkenntnis der Welt und für ein bewusstes Handeln in ihr. Es gibt demnach, aus klassischer Sicht, dabei häufig mit Verweis auf René Descartes, eine Trennung zwischen dem Körper und seinen mechanisch-biologischen Aktivitäten und dem Geist mit seinen „psychisch-reflexiven Bewusstseinsvorgängen“.¹⁰⁷ Diese Position pflegt eine Hierarchie, die das Geistige priorisiert und den Körper ausschließlich als *unbelebten* Reizeempfänger und die *Hauptleistung* menschlicher Wahrnehmung in der geistigen Repräsentation und deren Interpretation sieht. René Descartes mit seiner „scharfen Trennung von Materie und Geist“, die er als *res extensa* und *res cogitans* bezeichnet, wird immer wieder als Vertreter (Strohmann) dieser, aus Sicht der Phänomenologie, überholten Ansicht angeführt.¹⁰⁸

In der Phänomenologie, wie sie bei Husserl ihren Anfang nimmt und etwas später von Heidegger und Merleau-Ponty nicht weniger prominent weitergeführt wurde, wird dieser Dualismus von Körper und Geist in Frage gestellt.¹⁰⁹ Die Körperlichkeit des Menschen gewinnt an Bedeutung und zwar als Leib. Der Leib ist das Sein am Menschen, das eine Körperlichkeit aufweist, aber eben nicht ausschließlich darauf reduzierbar ist, das heißt, nicht ausschließlich mechanisch-biologisch beschreibbar ist. Der menschliche Leib ist demnach etwas, das zwischen Körper und Geist liegt bzw.

¹⁰⁶Husserl wird üblicherweise als der Urvater verstanden, siehe Husserl, 2012.

¹⁰⁷Waldenfels, 2000, S.30.

¹⁰⁸HWPh Bd. 5 (Ritter, Gründer und Gabriel, 1980), S. 192.

¹⁰⁹Damit sind ohne Frage drei prominente Vertreter genannt, die die Phänomenologie aber nicht in ihrer Gänze repräsentieren. Sie sind in der technikphilosophischen Rezeption der Phänomenologie von besonderer Bedeutung und werden deshalb hier behandelt.

Eigenschaften aufweist, die beiden zugesprochen werden können. Erfahrungen werden ihm eingeschrieben und das nicht (nur) durch eine physische Versehrtheit (Verbrennung, Amputation), sondern als eine leibliche Prägung durch alltägliches Handeln und Wahrnehmen. Und diese Prägung wiederum macht die Medialität des Leibes aus, die im Sinne medialen Wirkens (siehe dazu der erste Teil des Kapitels) Einfluss auf zukünftige Erfahrungen hat. Das Konzept der Leiblichkeit erhebt seinen Anspruch auf Gültigkeit aus der Einsicht, dass es einen *blinden Fleck* in Erkenntnis und Handeln gibt, der nicht ausschließliche auf geistige oder auf körperliche Eigenschaften zurückzuführen ist. Unter anderem ist hier von den leiblichen Fähigkeiten oder dem leiblichen Wissen des Menschen die Rede. Diese Fähigkeiten oder dieses Wissen sind nicht geistiger Natur, weil sie nicht oder nur unvollständig geistig thematisiert (repräsentiert) werden können. Das heißt, es liegt eine große Differenz zwischen dem Können und der Erklärung (Vermittlung durch Erklärung) dieses Könnens vor. Gleichzeitig sind die leiblichen Fähigkeiten und das leibliche Wissen auch nicht rein körperlicher Natur, da sie ebenfalls nicht als reine biologische Stimulus-Reflex-Relation (im Sinne von Reflexen, siehe dazu¹¹⁰) aufgefasst werden können. Beispiele für das Erlernen von leiblichem Wissen sind das Erlernen von Tanzschritten oder Musikinstrumenten. Ebenso kann das Handhaben und Bedienen bestimmter Techniken (wobei man auch Musikinstrumente als Techniken bzw. Tanzen als Technik im Sinne einer Fähigkeit verstehen kann) unter dem Gesichtspunkt leiblichen Lernens gesehen werden.

Das Erlernen oder Erlangen leiblichen Wissens (Wissen wird klassisch meist nur der geistigen Sphäre zugeschrieben, hier aber umgedeutet) wird als Verleiblichung bezeichnet. International hat dieses Konzept unter dem Begriff des *embodiment* in vielen Disziplinen Einzug gehalten. Im Englischen gibt es keine begriffliche Unterscheidung zwischen Verkörperung und Verleiblichung. Aus diesem Grund gibt es eine Reihe von weiteren Bedeutungen, die dem *embodiment* anlasten, hier aber nicht gemeint sind. Dazu gehört beispielsweise *Verkörperung* als Umsetzung von Ideen und Konzepten in konkrete Artefakte (z. B. Prototypen), Aktualisierung von Handlungsintentionen in konkreten Handlungsvollzügen oder von einem Softwareagenten in einer physischen Entität, beispielsweise in einen Roboter. Um der Differenzierung gerecht zu werden, wird in einigen englischsprachigen Publikationen vom Leib als „living body“ oder „embodied cognition“ gesprochen.¹¹¹

¹¹⁰Waldenfels, 2000, S. 111.

¹¹¹Die Bezeichnung „living body“ ist aus meiner Sicht insofern gut gewählt, als *living* nicht nur lebend oder lebendig meint (eine Eigenschaft, die ja auch der Tier- und Pflanzenwelt zukommt), sondern im Sinne von Wohnen verwendet werden kann, so wie Martin Heidegger das menschliche *Dasein* in der Welt charakterisiert, siehe Svanæs, 2013, Kirsh, 2013.

3.2.2 Unterscheidung: Leib und Körper

Die Unterscheidung zwischen Leib und Körper wird wie folgt verstanden. Der Leib gilt als Inbegriff allen Wahrnehmens und Fühlens und ist demnach unser Zugang zur Welt. Er ist zugleich Empfindendes (alle sinnlichen Wahrnehmungen sind durch den Leib vermittelt) aber auch Empfundenes (die eine Hand kann die andere fühlen).¹¹² Dagegen wird der Körper als etwas verstanden, das im „medizinischen Sinne objektivierbar ist“.¹¹³ Er ist in der Dritten-Person-Perspektive vollständig erfassbar, unter anderem untersuchbar, messbar oder aufschneidbar, das heißt, er ist mit einem „naturwissenschaftlich-quantifizierenden“ Zugriff begreifbar.¹¹⁴ Der Leib wiederum ist – auch wenn er eine körperliche (raum-zeitliche) Dimension hat, also ein Objekt in der Welt ist – eben nicht vollständig durch einen äußeren Zugriff erfassbar. Merleau-Ponty bezeichnet den Leib als Ambiguität und meint damit nach Waldenfels den Befund, dass der Leib „weder Geist noch Natur, weder Seele noch Körper, weder Innen noch Außen“ ist.¹¹⁵

Ein wichtiger Aspekt des Leibes ist demnach der des unmittelbaren Zugangs. Das heißt, im Gegensatz zu anderen Objekten der Welt ist der eigene Körper spürbar und fühlbar und gleichzeitig ist er auch das womit bzw. durch das man fühlt und wahrnimmt. Eine besondere Rolle wird hier der taktilen Wahrnehmung zugesprochen. Doch ist sie nicht auf Körperbewegungen sowie den Tastsinn beschränkt, wie es häufig in der technischen Forschung verstanden wird.

Ein Extremfall, wie der menschliche Körper einem begegnen kann, stellt der Leichnam dar. Er ist die räumliche Hülle des Menschen, die einen Teil ihrer Funktionen eingebüßt hat. Auch wenn das Herz nicht mehr Blut durch die Venen pumpt oder das Gehirn keine Reize mehr verarbeitet, kann die wasserabweisende Funktion der Haut noch lange nach dem Tod festgestellt werden. Zudem kann aus der Untersuchung der nicht mehr funktionierenden Systeme auf ihre frühere Funktionsweise zurückgeschlossen werden.

Dem Leib kommt nach Merleau-Ponty eine eigene Art des Verstehens zu. Es werden ihm Fähigkeiten zugeschrieben, die abgerufen werden können, die gleichzeitig aber keine geistigen sind. Das ist das bereits erwähnte leibliche Wissen. Wenn das leibliche Wissen Handeln leitet, dann kann man davon sprechen, dass der Leib über eine „vorbewusste Form der Intentionalität“ verfügt.¹¹⁶ Diese Intentionalität zeigt sich in situativer Spontanität,

¹¹²Wiegerling, 2008a, S. 13.

¹¹³Wiegerling, 2008a, S. 10.

¹¹⁴Wiegerling, 2008b, S. 85.

¹¹⁵Waldenfels, 2000, S. 42.

¹¹⁶Wiegerling, 2013, S. 221.

in Gewohnheiten und Routinen und ist basaler Bestandteil unseres Orientierens in der Welt und unseres Umgangs mit Dingen, wie es beim technischen Handeln der Fall ist.

„Der Leib ist nicht im Raume, er wohnt ihm ein“.¹¹⁷ Diesen Aspekt des Wohnens finden wir auch bei Heidegger, der die Trennung von Geistding und Körperding mit dem Konzept des „In-der-Welt-seins“ aufgelöst sehen will.¹¹⁸ Der Aspekt des Wohnens zeigt sich ganz wörtlich im Begriff der Gewohnheit. Zur Erklärung des *In-der-Welt-seins* sagt Heidegger: „in« stammt von *innan-*, wohnen, *habitare*, sich aufhalten; »an« bedeutet: ich bin gewohnt, vertraut mit, ich pflege etwas; es hat die Bedeutung von *colo* im Sinne von *habito* und *diligo*“.¹¹⁹

Gewohnheiten als leibliche Phänomene sind im Folgenden von besonderem Interesse. Das Ausbilden von Gewohnheiten wird als zentrales Element einer Mensch-Technik-Hybridisierung gesehen.

3.2.3 Maurice Merleau-Ponty: Erwerb von Gewohnheiten als Verwandlung des Körperschemas

Ein zentraler Begriff bei Merleau-Ponty ist das Körperschema. Der Leib als Körperschema beschreibt die Tatsache, dass der *eigene Körper* als Leib, das heißt beim Handeln und Wahrnehmen nicht in seinen einzelnen Segmenten, beispielsweise in Form bestimmter Muskeln angesprochen wird, sondern in seiner Gesamtheit benutzt wird.

„Sitze ich an meinem Tisch und will nach dem Telephon greifen, so umschließen einander die Bewegungen meiner Hand auf den Gegenstand zu, die Aufrichtung des Rumpfes und die Kontraktion der Beinmuskeln; ich will ein bestimmtes Ergebnis und von selbst verteilen die Aufgaben sich unter den beteiligten Segmenten, wobei die verschiedenen möglichen Kombinationen zum voraus als äquivalent gegeben sind: ich kann an der Stuhllehne angelehnt sitzen bleiben, wenn ich den Arm weiter ausstrecke, oder mich vorbeugen, oder auch mich halb erheben. All diese Bewegungen stehen uns von ihrer gemeinsamen Bedeutung her zur Verfügung“.¹²⁰

Die verschiedenen Segmente des eigenen Körpers als Leib sind in ihrer gesamten Funktion bekannt und „nicht ist ihre Koordination erlernt“.¹²¹

¹¹⁷Merleau-Ponty, 1966, S. 171.

¹¹⁸Heidegger, 1967, u. a. S. 56.

¹¹⁹Heidegger, 1967, S. 54.

¹²⁰Merleau-Ponty, 1966, S. 179.

¹²¹Merleau-Ponty, 1966, S. 179.

Natürlich gibt es Situationen, beispielsweise Optimierungen von Bewegungen beim Sport, bei denen einzelne Muskelpartien gezielt angesprochen werden. Meist wird beim Handeln die Verwendung der Muskeln in einem intellektuellen Sinne aber nicht gewusst, vielmehr handelt es sich um ein Art Körperwissen. Zu dem Körperschema gehört es auch, dass die Stellung von Körperteilen zueinander immer (intuitiv) gewusst wird, ohne dass sie nochmals eine Überprüfung durch die Sinne benötigen. Beispielsweise können Personen sich mit ihren Fingerspitzen vor dem Körper berühren, ohne dass die Annäherung durch visuelle Wahrnehmung geleitet wird.

Hier liegt es erst einmal nahe, die Erfassung der Stellung von Körperteilen zueinander als ein Resultat innerer Sensorik aufzufassen. Mit einer inneren Sensorik könnten die Segmente des Körpers zueinander gefühlt und in ihrer Koordination erfasst werden. Doch die Vorhandenheit einer inneren Sensorik ist nach Merleau-Ponty eine falsche Annahme, weil sie die Phänomene nicht erklären, die sich aus der Leibese Erfahrung ergeben.

Das führt er am Beispiel des Phantomglieds aus, bei dem es sich um eine Störung leiblicher Wahrnehmung handelt. Bei Personen, die einen Körperteil, meist einen Arm oder ein Bein verloren haben, verschwindet das Körperteil nicht, zumindest scheint das Verhalten der Personen, die das Körperteil verloren hatten, dies nahezu legen. Sie nehmen es noch immer wahr und handeln in einer Weise, nach der anzunehmen wäre, dass sie glauben, sie hätten die Gliedmaße noch immer. Die Betonung liegt auf dem Handeln bzw. Wahrnehmen und nicht auf dem Wissen, denn auf Nachfrage wird der Verlust des Körperteils nicht bestritten. Was ein Phantomglied ausmacht, lässt sich weder physiologisch noch psychologisch erklären.¹²² Das Vorhandensein einer inneren Sensorik würde dazu führen, dass ein Bein sofort als fehlend erfasst würde, da ja die Nervenreize faktisch ausbleiben. Die Frage ist, „wie kommt es, daß das Phantomglied weiterhin zum Körperschema zählt, obwohl die äußeren Reize [...] nirgends mehr zu finden sind“.¹²³ Die Antwort von Merleau-Ponty lautet: Das Phantomglied gehört weiterhin zur Einheit des Leibes und ist demnach Teil der gesamten Vorstellung des Körperschemas. Patienten wissen oft, dass ihr Arm nicht mehr da ist (indem sie dem Arzt dies versichern), aber gleichzeitig *wissen* sie es auch nicht, beispielsweise wenn sie Handeln. „Die Einheit des eigenen Leibes stellt sich durch seine Tätigkeit her“.¹²⁴

¹²²Wiegerling, 2008a, S. 17.

¹²³Waldenfels, 2000, S. 114.

¹²⁴Waldenfels, 2000, S. 118.

Leibliches Lernen

Das Verständnis des Körperschemas wird besonders im Prozess des Erwerbs von Gewohnheiten klar. Ein Beispiel für den Erwerb von Gewohnheit ist der Erwerb von motorischen Gewohnheiten oder wie es Merleau-Ponty sagt, die „motorische Erfassung einer Bewegungsbedeutung“.¹²⁵ Bei Merleau-Ponty sind unter anderem Beispiele für die Erweiterungen des Körperschemas in der *Einverleibung* von Tanzbewegungen oder Bewegungen beim Erlernen eines Instruments genannt. Charakteristisch für die Prägung des Leibes ist es, dass es eine erlernte Prägung ist, die ohne eine mentale Repräsentation auskommt, das heißt, keine Reduktion auf Bewusstsein und somit auf intellektuelles Lernen erfolgt. Tanzschritte sind, wenn das Tanzen gekonnt ist, nur sehr schwer und immer unvollständig in Worten repräsentierbar. Vielmehr sind die Tanzschritte nur performativ gewusst, das heißt am besten - nicht ausschließlich - im Vormachen vermittelbar. „Der Leib ist geradezu der Inbegriff dessen, was 'ich kann', ohne daß ich es mir ausdrücklich vorstellen *muß*, und teilweise auch, ohne daß ich es mir ausdrücklich vorstellen *kann*“.¹²⁶

Merleau-Ponty betont in seiner Theorie, dass es sich beim leiblichen Lernen nicht um einen naiven Behaviorismus handelt, der sich als eine Mischung aus reinen Stimuli und Konditionierung erklären lässt. Die behavioristische Lerntheorie beschreibt „Verhaltensgewöhnung als Konditionierung“.¹²⁷ Lust und Unlust bzw. Lohn und Strafe werden dabei als Hemmung oder Verstärkung des Verhaltens eingesetzt.¹²⁸ Der Behaviorismus geht von einem mechanischen Lernen, von einem sensorischem Reiz (Wahrnehmungsreiz: Wie die Farbe Rot, die den Stier beim Stierkampf reizt) und von einer Reaktion, die reflexhaft ist, aus. Der *pawlowsche* Hund wird beispielsweise klassisch konditioniert, indem eine Körperreaktion (Speichelfluss) in bestimmten Situationen (Geräusch der Glocke) *automatisch* ausgelöst wird. Bei der Konditionierung des Hundes handelt es sich nach der Theorie von Pawlow um einen bedingten Reflex, der sich von unbedingten Reflexen, die man nicht erlernt, wie beispielsweise der Augenlidreflex, unterscheidet. Der reine Behaviorismus sieht im Lernen von Tanzbewegungen eine Reihung von Reflexhandlungen, eine „Abfolge von partikularen Reaktionen“.¹²⁹

Ein Versuch mit Kindern, den Merleau-Ponty in seinem Buch „Die

¹²⁵Merleau-Ponty, 1966, S. 172: Bedeutung ist hier im Sinne von Wissen zu verstehen.

¹²⁶Waldenfels, 2000, S. 169.

¹²⁷Waldenfels, 2000, S. 154.

¹²⁸Waldenfels, 2000, S. 155.

¹²⁹Waldenfels, 2000, S. 154.

Struktur des Verhaltens“ beschreibt, zeigt auf, wie sich menschliches Lernen von dem der Tiere unterscheidet.¹³⁰ Im Versuch sind einige Schachteln aufgebaut, in denen ein Kind nach Süßigkeiten suchen soll. Nachdem das Kind diese gefunden hat, wird eine weitere Süßigkeit der Reihe nach, immer eine Schachtel weiter weg, versteckt. Das Kind sieht am Anfang zuerst immer in der Schachtel nach, in der die Süßigkeit zuvor war, lernt aber schnell die Regel, nach der die Süßigkeiten versteckt werden. Ein ähnlicher Versuch mit Menschenaffen zeigt, dass sie nicht fähig sind, diese Regel (immer eine Schachtel weiter) zu lernen. Sie sehen immer als erstes in der Schachtel nach, in der die Süßigkeit zuvor war.¹³¹ Das Ergebnis zeigt, dass es sich beim Lernen von Gewohnheiten, wie beim Tanzen oder wie im Süßigkeitenversuch, nicht um eine Reiz-Reflex Kopplung handelt, sondern um „typische Lösungen für typische Situationen“, die je nach Aufgabe *materiell* sehr unterschiedlich aussehen können.¹³² Die Möglichkeit, eine solche Transferleistung vollziehen zu können, entwickelt sich im Laufe des Erwachsenwerdens immer weiter.

Zusammenfassend kann man sagen, dass sich Gewohnheiten, unter anderem als motorisches Können, als ein leibliches *Verstehen* festsetzen. Gewohnheiten lassen sich zwar intellektuell theoretisieren, so wie ich über Tanzen und Tanzschritte im Allgemeinen, beispielsweise im Vergleich verschiedener Stile, sprechen kann, man kann dies aber nicht in *angemessen* vollständiger Weise repräsentieren. In den Beispielen zum Tanzen und Lernen von Instrumenten wird eine Gewöhnung vor allem anhand von physischen Bewegungen des Körpers veranschaulicht. Eine Verleiblichung geht aber viel weiter: Beispielsweise erlernen Menschen bestimmte Weisen der Organisation der visuellen Wahrnehmung, das heißt, dass sich ein Wahrnehmungsfeld in unterschiedlicher Weise strukturiert, wobei bestimmte Gestalten hervortreten, die vertraut sind.

Leibliches Lernen im technischen Handeln

Die Technik spielt in den Überlegungen Merleau-Pontys zum Körperschema eine wichtige Rolle. Handeln mit Objekten und Artefakten wird als eine Erweiterung des Körperschemas gesehen. Er bespricht unter anderem drei Beispiele, die im Bereich motorischen Lernens angesiedelt sind: Den Stock als Werkzeug zur Navigation eines Blinden, das Auto als Mittel zur Fortbewegung und die Tastatur der Schreibmaschine als Werkzeug zur Textverarbeitung.

¹³⁰Im Folgenden beziehe ich mich auf Waldenfels Deutung von Merleau-Ponty, 1976.

¹³¹Waldenfels, 2000, S. 161f.

¹³²Kaminski, 2010, S. 173.

Der erfahrene Autofahrer hat die räumliche Ausdehnung seines Autos insofern einverleibt, als dass er *einfach weiß*, ob er eine Engstelle passieren kann, ohne dies explizit abschätzen oder gar eine Berechnung erstellen zu müssen. Der Stock des Blinden ist für ihn zu einem zentralen Zugang zur Welt geworden: „Ist der Stock zum vertrauten Instrument geworden, so weicht die Welt der Gegenstände zurück und beginnt nicht mehr an der Haut der Hand, sondern erst am Ende des Stockes“.¹³³ Dass der Stock ein Vermittler zwischen der Person und den Objekten der Welt ist, spielt in der Nutzung keine Rolle mehr, der Stock wird vielmehr transparent im Umgang. Wenn man den Blinden jetzt fragen würde, wie lang der Stock denn sei, den er jeden Tag nutze und wie weit die Gegenstände entfernt seien, die er damit berührt, wird er es wahrscheinlich nicht angeben können. Entfernung sind nicht relevant, sofern er implizit weiß, dass er das Bein beim nächsten Schritt etwas mehr heben muss oder ob er an einer anderen Stelle ein wenig nach links gehen muss, um an einem Tisch vorbeizukommen.

Beim Erlernen des Umgangs mit der Schreibmaschine sieht Merleau-Ponty in ähnlicher Weise eine Einverleibung der Tastatur, die er als Klaviatur bezeichnet.

„Man kann Schreibmaschine schreiben können, ohne angeben zu wissen, wo sich auf der Klaviatur die Buchstaben befinden, aus denen man die Worte zusammensetzt. Maschinenschreiben können heißt nicht, die Stelle jedes Buchstabens auf der Klaviatur kennen, noch auch für einen jeden einen bedingten Reflex sich angeeignet haben, der sich beim Hinblick auf ihn auslöste“.¹³⁴

Wenn die Tastatur im beschriebenen Sinne einverleibt worden ist, kann diese Verbindung als Ergebnis der Hybridisierung eines Menschen durch die Schreibmaschine beschrieben werden. Merleau-Ponty erklärt diese Verbindung, welches ich als hybride Mensch-Technik-Konstellationen beschreibe, anhand des Konzepts des Körperschemas. Technik wird, im Modus leiblichen Verstehens, in das Körperschema integriert bzw. führt dazu, dass sich das Körperschema verwandelt: „[D]er Leib ist es, so sagen wir, der im Erwerb einer Gewohnheit 'versteht'“.¹³⁵

Wie ist der Prozess einer Integration als Verleiblichung nun genau zu verstehen? Der Prozess der Verleiblichung von Technik beginnt mit einer Nutzung, die ich auch als Erprobung charakterisiert habe. Im Beispiel des Schreibmaschinenschreibens, welches zeitgemäß auf das Schreibenlernen

¹³³Merleau-Ponty, 1966, S. 182.

¹³⁴Merleau-Ponty, 1966, S. 173f.

¹³⁵Merleau-Ponty, 1966, S. 174.

auf einer Computertastatur übertragen lässt, beginnt das Üben bzw. Erproben mit einer Suche. Was gesucht wird, ist der Ort der Taste mit dem Buchstaben, den man schreiben möchte. Jeder, der das Schreiben auf einer Tastatur gelernt hat, weiß um diese Tätigkeit. Erst wenn die dafür nötigen Handgriffe geübt sind, entsteht ein Schreibfluss, in dem nur noch ab und zu die Stellung der Finger überprüft werden. Der Anfang einer *Verleiblichung* ist demnach durch eine intellektuelle Leistung geprägt. Erst wenn das Schreiben immer mehr zur Gewohnheit wird, rückt die intellektuelle Beteiligung – das Verortung der Taste im Aktionsraum – immer mehr in den Hintergrund. Der intellektuelle Fokus liegt dann auf dem Inhalt des Textes. Ist die Tastatur in das Körperschema integriert und zu einem wirklichen Teil des lebendigen Leibes geworden, ist eine Übereinstimmung zwischen der Intention, ein Wort zu schreiben, und dem Vollzug unmittelbar gegeben.¹³⁶ Diese Erfahrung von Übereinstimmung wird bei Merleau-Ponty als eine verstanden, die, wie zuvor besprochen, keiner intellektuellen Leistung mehr bedarf.

Es ist zu resümieren, dass die Hybridisierung eine Weise der Verleiblichung ist, eben die, die ein operatives Verbinden von Mensch und Technik betrifft. Im Folgenden werden die leibphänomenologischen Überlegungen zur Hybridisierung aus der Perspektive Heideggers diskutiert.

3.2.4 Phänomenologie technischen Handelns bei Martin Heidegger

Bei Heidegger spielt der Leib als Konzept eine weniger prominente Rolle wie es beispielsweise bei Merleau-Ponty der Fall ist. Vielmehr wird der Leib buchstäblich sogar explizit in seiner Arbeit ausgeschlossen. Er weist an einer Stelle darauf hin, dass die Leiblichkeit kein Teil seiner Arbeit ist: „Die Verräumlichung des Daseins in seiner »Leiblichkeit«, die eine eigene hier nicht zu behandelnde Problematik in sich birgt“.¹³⁷

Dagegen ist der zentrale Begriff in seinem Hauptwerk „*Sein und Zeit*“ das *In-der-Welt-sein*, wobei er das Sein als Wohnen auffasst, das sich in Gewohnheiten manifestiert. In einem Abschnitt definiert er das *In-der-Welt-sein* als geistige Eigenschaft und die Leiblichkeit als körperliche, um gleich darauf auf sein „Zusammen-vorhanden-sein“ hinzuweisen. „Das In-Sein in einer Welt ist eine geistige Eigenschaft, und die »Räumlichkeit« des Menschen ist eine Beschaffenheit seiner Leiblichkeit, die immer zugleich durch Körperlichkeit »fundiert« wird. Damit steht man wieder bei einem

¹³⁶Merleau-Ponty fasst das leiblichen Verstehen so auf, dass wir eine Übereinstimmung zwischen Intention und Vollzug erfahren. Siehe Merleau-Ponty, 1966, S. 174.

¹³⁷Heidegger, 1967, S. 108.

Zusammen-vorhanden-sein eines so beschaffenen Geistdinges mit einem Körperding“.¹³⁸

Heidegger verhandelt in seiner Arbeit, wie zuvor kurz skizziert, ähnliche Themen wie es Merleau-Ponty unter dem Begriff des Leibes tut.¹³⁹

In-der-Welt-sein und technisches Handeln

Das In-der-Welt-sein ist bei Heidegger die zentrale Metapher, die das Verhältnis des Menschen zur Welt (Umwelt, Natur, Werkzeuge) beschreibt. Wichtig ist für ihn die Unausweichlichkeit des menschlichen Seins in der Welt, das heißt die Unmöglichkeit eines Nichtseins in der Welt. Es gibt kein Ich, das irgendwie unabhängig von der Welt existiert oder sich auch nur zeitweise in einen unabhängigen Zustand versetzen kann. Damit kann sich das Ich auch nicht isoliert betrachten, indem es sich zurückzieht und reflektiert. Damit ist aber auch die Welt nicht isoliert zu erfassen, da dieses ja die Möglichkeit eines Auseintretens voraussetzt. Das Sein des Menschen, nach dem Heidegger sucht und in dem In-der-Welt-sein findet, beschreibt eine praktische Existenz, die sich im Handeln und Verwenden von Dingen äußert.

„Das In-Sein ist nach dem Gesagten keine »Eigenschaft«, die es zuweilen hat, zuweilen auch nicht, ohne die es sein könnte so gut wie mit ihr. Der Mensch »ist« nicht und hat überdies noch ein Seinsverhältnis zur »Welt«, die er sich gelegentlich zulegt.“¹⁴⁰

Hier schließt sich die Frage an, worauf sich Verstehen denn richtet, wenn nicht auf die Welt. Martin Heidegger argumentiert, dass man die perspektivische Abhängigkeit nicht einfach reflektiert, das heißt, einfach abzieht und das, was übrigbleibt, die Welt ist, sondern dass sich Erkenntnis gar nicht auf die Welt richtet, sondern *nur* das Verhältnis von Mensch und Welt, das Sein, Gegenstand des Verstehens ist. Die Erkenntnis bezieht sich auf das In-Sein und nicht auf die Welt selbst, die nie als solche vorhanden ist.

Gleichzeitig ist wichtig, dass die Welt im Sinne Heideggers kein Ergebnis einer Konstruktion ist, sie wird von Menschen geteilt, aber nicht in ihrer Dinglichkeit, sondern wie sie als Zeug vorhanden und genutzt wird. Ein Zitat aus seinem Aufsatz zum Ursprung des Kunstwerks belegt diese antikonstruktivistische Einstellung:

¹³⁸Heidegger, 1967, S. 56.

¹³⁹Mehr zur *Leibproblematik* bei Heidegger in Gaitsch, 2012.

¹⁴⁰Heidegger, 1967, S. 57.

„Welt ist aber auch nicht nur ein eingebildeter, zur Summe des Vorhandenen hinzu vorgestellter Rahmen. Welt *weltet* und ist seiender als das Greifbare und Vernehmbare, worin wir uns heimisch glauben. Welt ist nie ein Gegenstand, der vor uns steht und angeschaut werden kann“.¹⁴¹

Martin Heidegger charakterisiert das In-der-Welt-sein mit Hilfe des alltäglichen praktischen Handelns, wobei er explizit auf das technische Handeln eingeht.¹⁴² Dabei spielen die Gewohnheiten und Routinen des Gebrauchs von Dingen als „Zeug“ eine zentrale Rolle. Mit Zeugs bezeichnet Heidegger Dinge jeglicher Art, die in irgendeiner Weise so verfügbar sind, dass man sie zweckmäßig gebrauchen kann.¹⁴³ Dinge sind nicht grundsätzlich zu thematisieren, sondern nur, wenn sie Zeug sind und dann nur in ihrer Mannigfaltigkeit als Zeug.

„Ein Zeug »ist« strenggenommen nie. Zum Sein von Zeug gehört je immer ein Zeugganzes, darin es dieses Zeug sein kann, das es ist. Zeug ist wesenhaft »etwas, um zu...«“.¹⁴⁴

Wenn Heidegger davon spricht, dass das Zeug nie *ist*, meint er, dass es nie im Singular vorhanden ist, es demnach kein Zeug gibt, ohne dass es nicht in einem Verweisungszusammenhang mit anderem Zeug (Zeugganzes) steht. Das Zeugganze bezeichnet eine Struktur von Verweisungen, wozu andere Gegenstände gehören, die für einen praktischen Gebrauch nötig sind oder möglicherweise gebraucht werden können.

Der selbstverständliche Gebrauch von Dingen als Zeug wird von Heidegger unter dem Gesichtspunkt der „Zuhandenheit“ besprochen. „Die Seinsart von Zeug, in der es sich von ihm selbst her offenbart, nennen wir die Zuhandenheit“.¹⁴⁵ Demnach charakterisiert er das Zeug in seiner Funktion als Zuhandenes. Im technischen Nutzen des Hammers ist der Hammer zuhanden und nur so wird der Hammer als Technik verstanden. Der Hammer als Gegenstand (Objekt, Ding) ist nichts Technisches, ein Hammer ist nur in der möglichen Praxis des Hammerns (d. h. im Gebrauch) ein Hammer. Ein Charakterzug des Zuhandenen ist wiederum sein Verschwinden, das heißt, das Verschwinden des Hammers, der im selbstverständlichen Gebrauch nicht thematisiert wird.

¹⁴¹Heidegger, 1992, S. 45.

¹⁴²Heidegger, 1967, S. 66ff.

¹⁴³Zeug ist im Begriff des Werkzeugs zu finden, Heidegger, 1967, S. 68.

¹⁴⁴Heidegger, 1967, S. 68.

¹⁴⁵Heidegger, 1967, S. 69.

Kippfigur: Vom Zuhanden zum Vorhandenen

Werkzeuge und technische Systeme sind so lange zuhänden, bis sich eine (grundlegende) Störung einstellt. Wenn beispielsweise ein Werkzeug kaputt ist, erscheint es dem Nutzer in seiner Unzuhandenheit und zeigt sich *plötzlich* in seiner Dinglichkeit als Vorhandenes. Diesen Moment der Störung bezeichnet Heidegger als „Auffallen“, die das zuhandene Zeug als unzuhandenes Ding erscheinen lässt.¹⁴⁶ Das kaputte Werkzeug ist nicht plötzlich ein Stück Holz oder Metall, sondern wirkt in seiner vorherigen Zuhandenheit nach. Dieses Verhältnis von Zuhandenheit und Vorhandenheit wird bei Heidegger wie folgt beschrieben:

„Was aber die Unverwendbarkeit entdeckt, ist nicht das hinsehende Feststellen von Eigenschaften, sondern die Umsicht des gebrauchenden Umgangs. In solchem Entdecken der Unverwendbarkeit fällt das Zeug auf. Das Auffallen gibt das zuhandene Zeug in einer gewissen Unzuhandenheit. Darin liegt aber: das Unbrauchbare liegt nur da –, es zeigt sich als Zeugding, das so und so aussieht und in seiner Zuhandenheit als so aussehendes ständig auch vorhanden war“.¹⁴⁷

Diese Änderung des Verhältnisses kann treffend als ein „Kippeffekt“ veranschaulicht werden, der, so Kaminski, zwischen der Praxis und der Objektwahrnehmung besteht.¹⁴⁸ Man kann die Figur des Kippens mit Heideggers Worten als „Bemerken von Unzuhandenem“ beschreiben, was, wie er sagt, neben anderen Erlebnissen von Unzuhandenheit, im Modus der Auffälligkeit erfolgt.¹⁴⁹

Wichtig für das Verständnis von Heideggers Konzeption des Zuhandenen und Vorhandenen ist es, zu wissen, dass seine Erzählung mit dem Zuhandenen beginnt. In-der-Welt-sein heißt für ihn, dass Technik vorreflexiv immer schon genutzt wird. Man kann, wenn man die Entwicklung eines Säuglings bzw. Kleinkinds beobachtet, erkennen, dass es eine anfängliche Zuhandenheit gibt, die im buchstäblichen Sinne von der Hand kommt. Ein Säugling nimmt alle Dinge in die Hand, die in seiner Reichweite erscheinen und gebraucht sie, ohne diese bewusst zu reflektieren. Die Hand als erstes Werkzeug – das Werkzeug aller Werkzeuge.¹⁵⁰ Der Gebrauch eines Hammers, um wieder auf die Beispiele von Heidegger zurückzukommen, entwickelt sich aus dem Zuhandensein der Hand und verfeinert sich weiter

¹⁴⁶Heidegger, 1967, S. 73.

¹⁴⁷Heidegger, 1967, S. 73.

¹⁴⁸Kaminski, 2010, S. 141.

¹⁴⁹Heidegger, 1967, S. 73.

¹⁵⁰Vgl. eine Deutung von Aristoteles Aristoteles, 1983

im Gebrauch. Heidegger geht vom Zuhandenen aus, das heißt von einem praktischen Agieren in der Welt als das Ursprüngliche aus. Das steht einer instrumentell-technischen Perspektive entgegen, nach der wir ständig Dingen als *Vorhandene* begegnen und sie dahingehend bewerten, wie sie uns *zuhandensein* können, das heißt, als Mittel für bestimmte Zwecke praktisch gebraucht werden können.

Heidegger geht davon aus, dass Technik nur im Gebrauch verstanden wird:

„Je weniger das Hammerding nur begafft wird, je zugreifender es gebraucht wird, um so ursprünglicher wird das Verhältnis zu ihm [...] Das schärfste Nur-noch-hinsehen auf das so und so beschaffene »Aussehen« von Dingen vermag Zuhandenes nicht zu entdecken. Der nur »theoretisch« hinsehende Blick auf Dinge entbehrt des Verstehens von Zuhandenheit“.¹⁵¹

Die Überlegungen Heideggers, nach denen er das Zeug in seiner Zuhandenheit begrifflich beschreibt, entsprechen, der Systematik nach, dem *Ansatz* der Medienphilosophie, demnach die Technik als verborgen bzw. das verschwindend aufgefasst wird, wenn sie störungsfrei genutzt wird (3.1.2). Die Frage nach der Entstehung des Verhältnisses von Mensch und Technik, in der die Technik in der Zuhandenheit gebraucht wird, wird bei Heidegger nicht thematisiert, da er dieses als das ursprüngliche Verhältnis tituliert. Vielmehr ist der umgekehrte Weg, dem in seiner Theorie alle Aufmerksamkeit zukommt, eine Beschreibung dessen, wie sich die Technik als „nächstzuhandene[s] Seiende[s]“ als „unverwendbar“ herausstellt.¹⁵² Dann erst, so Heidegger, erscheint das Zuhandene als Vorhandenes.

Hinsichtlich der Konzeption Heideggers stellt sich die Frage, ob eine Analogie zu den Überlegungen zur medialen Erfahrung besteht, die in Abschnitt 3.1.3 diskutiert wurde. Störungen in der Erprobung eines Umgangs wurde unter bestimmten Umständen eine produktive Kraft zugesprochen, die es ermöglicht die Verhältnisnahme zur Technik anzupassen und somit die Verbindung enger werden lässt. Könnte das Auffallen oder Aufdrängen der Technik als Störung des Vollzugs, die das Vorhandene erscheinen lässt, nicht eine Grundlage bilden, um neue Möglichkeiten des Umgangs zu thematisieren und so in die Zuhandenheit zurückzukehren. Heidegger macht klar, dass er das nicht so versteht: Vielmehr verabschiedet sich das Zuhandene, so dass am Vorhandenen keinesfalls aufleuchtet, was das Zuhandene „fundiert“.¹⁵³ In Heideggers Konzeption verbleibt der Prozess, der zu einer

¹⁵¹Heidegger, 1967, S. 69.

¹⁵²Heidegger, 1967, S. 73.

¹⁵³Heidegger, 1967, 75.

engeren funktionalen Verbindung führt und als eine Hybridisierung verstanden werden kann, athematisch ausschließlich im Bereich des Zuhandenen. Das obige Zitat besagt, dass je zugreifender die Technik gebraucht wird, desto ursprünglicher wird das Verhältnis. Man kann davon ausgehen, dass Heidegger, *kleine* Störungen im Umgang als etwas auffasst, das von Personen nicht als Unzuhandenes wahrgenommen wird, sondern zu einer Anpassung im „gebrauchend-hantierende[n] Umgang“, das heißt in der Sphäre des Leiblichen führt und so unreflektiert im Bereich des Zuhandenen verbleibt.¹⁵⁴ Heidegger spricht hier nicht von leiblichem Verstehen, sondern davon, dass das Hantieren nicht blind sei, sondern seine eigene Sichtart hätte.¹⁵⁵

Ergänzung: Vom Werkzeug zur (intelligenten) Systemtechnik

Heidegger ist in *Sein und Zeit*, zumindest in seinen Beispielen, auf einen einfachen Werkzeuggebrauch beschränkt. Die *moderne* Systemtechnik wird erst in seinen späteren Arbeiten relevant, vor allem wenn er in „Die Technik und die Kehre“ von dem „Ge-stell“ spricht.¹⁵⁶ Für meine Arbeit sind vor allem die Überlegungen in *Sein und Zeit* relevant, die mit Blick auf die intelligente Systemtechnik aktualisiert werden sollen. Die wesenhafte Fassung moderner Technik als *Ge-stells* ist für die Arbeit von geringerem Interesse.

Bei dem Technikverständnis, wie es in *Technik und die Kehre* entwickelt wird, handelt es sich, anderes als bei *Sein und Zeit*, um eines, das aus einer globalen Perspektive erfolgt. Das *Ge-stell* sieht seine ideologische Zuspitzung, wenn auch nicht von Heidegger explizit genannt, in den kapitalistisch-technischen Gesellschaften. Zentral für seine Konzeption ist die ständige *Verpflichtung* in einer technisierten Gesellschaft, alle Dinge, die einem in der Welt begegnen, als Zeug aufzufassen, das heißt, ausschließlich hinsichtlich ihrer Verwendbarkeit zu betrachten. Heidegger bezeichnet das auf Nützlichkeit optimierte Betrachten von Dingen als ein *Stellen*. Davon ist im Speziellen auch nicht der Mensch und seine Arbeitskraft ausgeschlossen, was mit Blick auf aktuelle Strömungen beispielhaft mit dem Begriff des Humankapitals in Verbindung gebracht werden kann. Die Vorsilbe *Ge-* steht für das Versammeln von Dingen oder Aspekten, wie dies beispielsweise bei Gebirge als eine Versammlung von Bergen auf ähnliche Weise der Fall ist. Das *Ge-stell* ist demnach das Versammeln allen Stellens, also das Verfügbar-machen und Erhalten von Dingen zur Verwendung. Beispielsweise ist der Fluss nur noch ein Reservoir, das der Gewinnung von

¹⁵⁴Heidegger, 1967, S. 69.

¹⁵⁵Heidegger, 1967, S. 69.

¹⁵⁶Heidegger, 1962.

Wasserkraft dienlich ist und wird nicht auch als ein Wert an sich oder zumindest als ein ästhetischer Wert aufgefasst. Das *Ge-stell* ist für ihn das Wesen neuzeitlicher Systemtechnik, das heißt der prägende Charakter moderner Technik. Der Mensch als Teil des Systems (*Ge-stells*) kann sich aber nur schwer dazu verhalten, weil er selbst Teil des Systems ist: Zugespitzt kann man sagen, dass sein Umgang mit Problemen ausschließlich im Lösungsraum technischer Mittel erfolgen kann.

Die globale Perspektive, die Heidegger in seinen späteren Arbeiten, wie in der hier vorgestellten Arbeit zur *Technik und Kehre* einnimmt, ist eine, die hinsichtlich von Mensch-AI-Konstellationen in sozialen Systemen von einem allgemeinen Interesse ist, aber für die Beantwortung der Forschungsfrage der hier vorliegenden Arbeit zu keiner weiteren Klärung beiträgt. Für die Betrachtung der Mensch-AI Hybridisierung ist eine *lokalerer* Perspektive relevant, wie sie in *Sein und Zeit* im individuellen Werkzeuggebrauch betrachtet wurde, das heißt, in einem *bilateralen* Aktionsrahmen von Nutzern und intelligenter Technik.

Im folgenden Abschnitt wird ein neuerer Ansatz aus der amerikanischen Tradition der Phänomenologie besprochen, der sich auch auf die klassischen Positionen beziehen. Dieser Ansatz ist für meine Arbeit vor allem von Interesse, weil er explizit auf neuere technische Entwicklungen eingeht, das heißt sich mit technischen Innovationen in der Informatik befasst.

3.2.5 Technik-Mensch-Verhältnisse bei Don Ihde: Eine kritische Betrachtung

Don Ihde präsentiert in seinem Hauptwerk eine phänomenologische Konzeption des Verhältnisses von Mensch und Technik. Die Stärke von Ihdes Arbeit ist eine konkretere Bezugnahme auf aktuelle Systemtechniken, wie sie in den technischen Visionen zum Ubiquitous Computing und zur Robotik oder zum Quantified Self zu finden sind. Dabei entwickelt er, sehr anschaulich, vier verschiedene Typen von Mensch-Technik-Verhältnissen. Bei einigen der von ihm besprochenen Technologien handelt es sich um intelligente Techniken, wie sie auch in der vorliegenden Arbeit charakterisiert sind.

Wie Martin Heidegger argumentiert auch Don Ihde, dass Technik nur im Verhältnis zu Menschen zu verstehen ist. Don Ihde stellt vier Typen von Mensch-Technik-Verhältnissen vor. Der erste Typ wird als „embodied relations“ bezeichnet. Hierbei handelt es sich um ein Verhältnis, bei dem ein technisches Artefakt oder System im oben genannten Sinne verleiblicht wird. Ein sehr einfaches Beispiel für dieses Verhältnis ist bei ihm die Brille: Während eine Brille genutzt wird, das kann beim Autofahren oder Lesen sein, verschwindet diese in der Wahrnehmung des Nutzers. „My glasses

become part of the way I ordinary experience my surroundings; they 'withdraw' and are barely noticed, if at all".¹⁵⁷ Als eine andere Beispielquelle dient ihm, wie schon bei Merleau-Ponty, das Autofahren. Eine fortgeschrittene Fertigkeit, die beim Autofahren gebraucht wird, kann z.B. sein, dass bestimmte Vibrationen des Autos, die zum einen auf den Untergrund und zum anderen auf eine bestimmte Geschwindigkeit schließen lassen, beim Fahren unbewusst berücksichtigt werden, ohne dass diese taktile Wahrnehmung bewusst thematisiert wird. Den Übergang zwischen dem Menschen und der einverlebten Technik, nennt er „enigma position“, das heißt, es handelt sich um eine Grenze zwischen Nutzer und technischem System, die *verschwommen* ist, in gewisser Weise *dunkel* bleibt.¹⁵⁸

Don Ihde sagt, dass bei Technikgestaltern der Wunsch naheliegen könnte, eine „total transparency, total embodiment“ zu erreichen.¹⁵⁹ Doch hinsichtlich einer solchen Totalität der Verleiblichung, so seine Argumentation, kann man dann nicht mehr von einer *embodied relation* sprechen, sondern dies würde einer Verschmelzung von Technik und Mensch gleichkommen. Konsequenterweise kann man dann auch nicht mehr von Technik sprechen, die sich durch das Verhältnis zum Menschen definiert, welches sich in einem solchen Fall in eine Einheit auflöst. Es sagt: das wäre „equivalent to there being no technology“.¹⁶⁰

Die zweite Typisierung von Mensch-Technik-Verhältnissen nennt er „hermeneutic relations“. Hierbei handelt es sich um Techniken, zu denen Menschen in ein Verhältnis treten können, die die Welt repräsentieren, ohne dass diese Repräsentation für den Nutzer thematisierbar bzw. reflektierbar sei. Auch hier gibt es eine Position, die für den Nutzer dunkel bleibt (*enigma position*) und zwar die zwischen der Technik und der Welt.

Als Beispiel nennt er das Thermometer. Wenn es Temperaturinformationen über Orte vermittelt, die für eine Person nicht direkt zugänglich sind, dann handelt es sich um ein solches hermeneutisches Verhältnis. Ihde nennt das extreme Beispiel eines Atomreaktors, dessen Temperatur offensichtlich aus sicherheitstechnischen Gründen von außen überwacht werden

¹⁵⁷Ihde, 1990, S. 73.

¹⁵⁸U. a. Ihde, 1990, S. 86.

¹⁵⁹Ihde, 1990, S. 75.

¹⁶⁰Ihde, 1990, S. 75.

muss.¹⁶¹ Wichtig ist für dieses Verhältnis erst einmal die praktische Unmöglichkeit, einen direkten Wahrnehmungszugang zu haben.¹⁶² Die Deutungsleistung des technischen Systems bleibt, weil unzugänglich, nicht verifizierbar und deshalb dunkel. Die *hermeneutic relation* ist, so wie er sie in seinem Beispiel veranschaulicht, doch eher ein Anwendungsfall der besonderen Art. Dagegen vermitteln Thermometer, wie sie im Alltag gebraucht werden, meist Informationen über die Temperatur von Orten, an denen man gerade nicht ist, zu denen man aber Zugang haben könnte. Beispielsweise die Höhe der Außentemperatur, die von einem Thermometer im Garten in die Küche vermittelt wird. Ist man mit diesem Thermometer oder einem Thermometer als solches vertraut, wird die Anzeige zur Vermittlung selbstverständlich, das Thermometer als Instrument spezifischer Deutung der physikalischen Umgebung wurde einverleibt.

Ausgenommen der extremen Beispiele von Ihde sind technische Systeme, die physikalische Gegebenheiten repräsentieren bzw. deuten, nicht als eine eigene Kategorie zu sehen, sondern werden treffender als eine *embodied relation* beschrieben. Im Alltag spielen technische Instrumente zur Deutung von physikalischen Größen als eine Art *ausgelagerter* Sensorik eine wichtige Rolle. Ein vertrautes Instrument der Messung kann zum Bestandteil alltäglichen Handelns und Fühlens werden. Beispielsweise mit Blick auf die Techniken des Quantified Self kann man beobachten, dass die Zahlenwerte und Graphen, die die biometrischen Daten von Personen deuten, in der Wahrnehmung von Nutzern dominant werden. Das Verhältnis als bewusstes Verhältnis, das heißt der Ursprung und die Art und Weise der Deutung, *verblässen* jedoch nach einiger Zeit.

Ein dritter Typ von Verhältnissen betrifft die „alterity relations“.¹⁶³ Bei diesem Typ bezieht sich Ihde auf das Verhältnis zu Techniken, die in einer intelligenten Weise autonom agieren. Dieses Verhältnis zeichnet sich dadurch aus, dass diese Technik vom Menschen als etwas anderes bzw. als ein Anderer gesehen wird, was *alterity* wörtlich meint. Das Erscheinen der Technik als etwas anderes, das eventuell auch eigene Intentionen verfolgt, ist ein zentrales Anliegen, das er in der Entwicklung der Robotik erkennen mag. Ihde spricht bei dem Wunsch der Entwicklung von autonomen Agenten von einer „counterpart fantasy“.¹⁶⁴ Ist die Idealvorstellung eines *counterparts* aber erreicht, ist auch die Charakterisierung als Technik *passé*,

¹⁶¹Ihde, 1990, S. 85.

¹⁶²Dieses Beispiel von Ihde erinnert an den Atomunfall in Harrisburg von 1979, bei dem ein Thermometer, welches die richtige Temperatur anzeigte, aufgrund vorheriger Erfahrungen mit der Fehleranfälligkeit des Thermometers auch diesmal als Fehlerhaft interpretiert wurde und deshalb nicht die richtigen Maßnahmen getroffen wurden, um das Unglück zu verhindern

¹⁶³Ihde, 1990, S. 97ff.

¹⁶⁴Ihde, 1990, S. 106.

das heißt in seinen Worten: „extrapolates the technology into that which is not a technology“.¹⁶⁵

Bei dem vierten Typ handelt es sich, wie Ihde selbst sagt, im engeren Sinn um kein Mensch-Technik-Verhältnis. Diesen Typ bezeichnet er als „background relations“.¹⁶⁶ Die Technik befindet sich nicht wie bei den anderen Typisierungen in einem direkten Verhältnis zu Menschen, sondern agiert selbstständig als „automatic and semiautomatic machines“ im Hintergrund.¹⁶⁷ Er charakterisiert diesen Typ als „near-technological environment“ und als „no-longer-used technologies“.¹⁶⁸ Er spricht hier aber nicht von einem (subjektiven) Verschwinden von Technik während der Nutzung, wie es bei den *embodied relations* der Fall ist, sondern von einer buchstäblichen Abwesenheit¹⁶⁹, die auch keine Interventionen mehr zulässt. Nach dieser Vorstellung werden grundlegende Prozesse am Laufen gehalten, die keine Einflussnahme durch Personen mehr nötig haben und ein Einfluss auch nicht möglich ist. Bestimmte alltägliche Infrastrukturen könnte man auf diese Weise deuten, beispielsweise das Stromerzeugungs- und Transportnetz, das für Endnutzer größtenteils abwesend ist. Aus Sicht der Provider ist das Stromnetz aber sehr wohl vorhanden und funktioniert nicht ohne Überwachung und Wartung. Man kann hier höchstens von einem bestimmten Standpunkt aus (z.B. Endverbraucher) von *background relations* sprechen und nicht im generellen Sinne.

Don Ihde lotet mit seiner Typisierung von Mensch-Technik-Verhältnissen die Grenzen dieser Beziehungen aus. Von besonderem Interesse sind für ihn die Grenzfälle, in denen die Verhältnisse als *wirkliche* Verhältnisse unklar werden (Man kann eigentlich nur von einem Verhältnis von zwei Entitäten sprechen, wenn es eine Grenze gibt). Beispielsweise spricht er davon, dass Techniken bestimmte Elemente der physischen Umwelt deuten und präsentieren, ohne dass Personen Zugang zu den Ursprüngen und Mechanismen der Repräsentation haben, sowie von autonomen technischen Systemen, die als etwas Andersartiges erscheinen oder im Hintergrund völlig ohne menschliche Interventionsmöglichkeiten existieren und agieren. Er vermittelt mit seiner Typisierung eine gute Übersicht an *Tendenzen*, wie Verhältnisse von Mensch und Technik weitergedacht erscheinen können. Doch ausgenommen von den Grenzfällen, die, wie er selbst sagt, Techniken darstellen, die gar keine Techniken mehr sind, lassen sich nach meiner Ansicht alle Typen in ihrer abgeschwächten Form als Deutungs- und Repräsentationssysteme (*hermeneutic*), Roboter

¹⁶⁵Ihde, 1990, S. 106.

¹⁶⁶Ihde, 1990, S. 108ff.

¹⁶⁷Ihde, 1990, S. 108.

¹⁶⁸Ihde, 1990, S. 108.

¹⁶⁹Ihde, 1990, S. 109.

(*alterity*) und ubiquitäre Techniken (*background*) unter dem Konzept der Verleiblichung konzipieren, das heißt, diese können *auch* als eine *embodied relation* beschrieben werden.

Beispielsweise sind ein Roboter, der autonome Aufgaben in einem Haushalt übernimmt, oder ein technisches System, das im Sinne einer Smart-Home-Anwendung im Hintergrund agiert, in ihrem Einsatz meist ein Ergebnis von Entscheidungen, d. h. man beschließt, dass man ein solches System braucht und installieren möchte. Und während ihres Einsatzes bleiben die Techniken immer Gegenstand einer möglichen Reflexion ihrer Leistungsfähigkeit, indem die Resultate ihrer Unterstützung bewertet werden. Es kann einen längeren Prozess der Gewöhnung geben, bis die Technik ein selbstverständlicher Teil alltäglicher Praxis geworden ist. Bei diesem Prozess kann es sich um eine Hybridisierung handeln, die sich als Verleiblichung aktualisiert. Was die Bedingungen für die Möglichkeiten einer Mensch-AI-Hybridisierung sind, wird im folgenden dritten Teil des Kapitels diskutiert (3.3).

Die bisherigen Überlegungen aus der Medienphilosophie und der Phänomenologie des Leibes werden im Folgenden zusammengeführt und mit Bezug auf die Möglichkeiten und Grenzen einer Mensch-AI-Hybridisierung diskutiert.

3.3 Philosophie der Mensch-AI-Hybridisierung

In diesem Abschnitt werden die allgemeinen Überlegungen zur Hybridisierung von Menschen durch Technik zusammengeführt und diskutiert, welche Konsequenzen sich daraus für die Mensch-AI-Hybridisierung ergeben.

Die Hybridisierung ist ein Prozess, der seinen Ausgang bei der ersten Begegnung eines Menschen und einer bestimmten Technik nimmt und bis zu einem stabilen Zustand reicht, dem Mensch-Technik-Hybriden, der sich durch einen nachhaltigen und selbstverständlichen Umgang auszeichnet. Im Rahmen der hier geführten Diskussion der theoretischen Ansätze gab es viele Beispiele für Mensch-Technik-Verhältnisse, die hinsichtlich der Möglichkeit einer Hybridisierung diskutiert wurden. Dabei handelte es sich u. a. um den menschlichen Umgang mit Hämmern, Karten, Blindenstöcken, Tageslichtprojektoren, Schreibmaschinen oder Automobilen. Der erfahrene Autofahrer und das eingefahrene Auto bilden einen solchen Hybriden in einem vorläufig stabilen Zustand.

Es ist nicht allein aus technikphilosophischer Sicht relevant, diesen Prozess zu erklären, sondern dient auch dem interdisziplinären Interesse der Arbeit. Aus Sicht der Informatik (im Speziellen des Ubiquitous Computing) ist es relevant zu verstehen, wie Personen einzelnen Techniken *begegnen*, die für sie neu sind und zu ergründen, wie aus dieser Begegnung ein

selbstverständlicher und langanhaltender Umgang werden kann. Die philosophischen Überlegungen zur Hybridisierung können die Grundlage dafür bilden.

3.3.1 Zusammenführung: Hybridisierung von Menschen durch Technik

Die Hybridisierung von Menschen durch Technik wurde unter der Berücksichtigung medienphilosophischer Ansätze und Rückbezüge zur Phänomenologie des Leibes diskutiert.

Im medienphilosophischen Teil dieses Kapitels wurde deutlich, dass Techniken hinsichtlich ihrer medialen Wirksamkeit konzipiert werden können. Ein systematisches Verständnis des Medialen der Technik lässt sich nur im Verhältnis zum Menschen entwickeln, dessen Handeln dadurch vermittelt ist. Sybille Krämer, deren Medienphilosophie in Abschnitt 3.1.2 diskutiert wurde, spricht in ihrer Arbeit von *operativen Einheiten*, die Menschen und Medien bilden können. Die sich im Mediengebrauch bildende Einheit zeichnet sich dadurch aus, dass das Medium in seiner Rolle als Vermittelndes oder Ermöglichendes für den Menschen in seiner Wahrnehmung verschwindet. Das Verschwinden des Mediums im störungsfreien Vollzug ist eine medientheoretische Position, die Krämer von anderen Autoren übernimmt.¹⁷⁰ Die Frage, die sich für das Interesse der hier vorliegenden Arbeit daran anschloss, war, wie diese operative Einheit, die ich als hybride Konstellation verstanden sah, genauer zu beschreiben und vor allem, wie das Entstehen dieser Verbindung zu erklären ist.

Um diese hybride Konstellation genauer zu verstehen, wurde ein Ansatz von Bruno Latour diskutiert (3.1.4). In Latours soziologischen Überlegungen nimmt das hybride Zusammenwirken von Menschen und Nicht-Menschen eine zentrale Rolle ein. Bei Technik handelt es sich nach Latour um einen Spezialfall nicht-menschlicher Akteure, deren Beschreibung er viele Beispiele widmet. Hybride aus Menschen und Technik sind nach seiner Ansicht nicht nur funktionale Verbindungen, obwohl er einige Beispiele in dieser Hinsicht diskutiert. Sie sind vielmehr als der grundlegende Teil eines sozialen Netzwerks zu verstehen, in dem Eigenschaften des einen im anderen wirksam werden und sich in der wiederholten Performanz bzw. Aktivität dieser Einheit sich der Hybrid stabilisiert und als neuer sozialer Akteur mit *eigenen* Zielen etabliert.¹⁷¹ Dieser Aspekt der

¹⁷⁰U. a. Heider, 1927, Luhmann, 2001, Mersch, 2003.

¹⁷¹Siehe dazu die Überlegungen von Latour, 2006, S. 504.

Stabilisierung (oder auch Punktualisierung) findet sich in Latours Überlegungen zum *Blackboxing*, welches ein Verbinden von Menschen und Technik als einen (reversiblen) Prozess beschreibt, der die „vereinte Produktion [die das Agieren] von Akteuren und Artefakten völlig undurchsichtig macht“.¹⁷² Das *Blackboxing* ist ein erster Aspekt im Zuge der Konkretisierung der Hybridisierung von Menschen durch Technik. Beispiele für funktionale Hybride sind der Dozent und der Tageslichtprojektor oder Chemiker und Pipetten, die sich jeweils in der Praxis eines wiederholten gemeinsamen Agierens in einer stabilen Verbindung zusammenfinden.

Die Entstehung hybrider Verbindungen wurde des Weiteren unter Berücksichtigung von Christoph Hubigs medienphilosophischen Ansatz konkretisiert. Hubig erklärt die Hybridisierung von Mensch und Technik anhand des Konzepts medialer Rekonstruktion (3.1.3). Bevor die Technik als Medium im störungsfreien Vollzug verschwinden kann, muss die Medialität im Erproben bzw. Erlernen eines Umgangs erfahren bzw. erlebt werden. Diese Erfahrung ist Grundlage dafür, dass die Technik als Medium hinsichtlich ihres möglichen Wirksamwerdens Gegenstand einer partiellen Vorstellung wird.

Die Erfahrung des Mediums erklärt Hubig anhand des Konzepts von Spuren. Die Konzeption besagt, dass die Nutzung einer Technik Spuren ihrer Medialität *hinterlässt*. Diese Spuren können erfahren, das heißt gelesen werden und hinsichtlich einer praktischen Vorstellung des Mediums der Technik gedeutet werden. Was Hubig hier unter Spurenlesen versteht, lässt sich detailliert erklären, wenn man die Begegnung von Personen mit einer *neuen* Technik sequenziell als das Erproben eines Umgangs beschreibt.

Personen haben Erwartungen an den Umgang mit einer Technik, die den Einsatz der Technik für einen bestimmten Zweck betrifft. Erwartungen im Umgang mit Technik sind keine einzelnen Intentionen, die das Handeln leiten, sondern eine Fülle von Aspekten, die größtenteils nur implizit mitvorliegen und oft erst im Enttäuschungsfall den Nutzern bewusst wird. Im Umgang mit einer neuen Technik kann ein Gefühl der Enttäuschung eintreten, wenn nicht das erwartete Ergebnis erzielt wurde. Ebenso können die Erwartungen erfüllt, eventuell sogar übererfüllt werden – wenn *mehr* erreicht wurde als erwartet. Diese Erfahrungen im Umgang mit Technik nennt Hubig die „Enttäuschungs-, Überraschungs- oder Gelingenserlebnisse“.¹⁷³

Diese Erlebnisse bilden die Grundlage für Spuren, welche anhand von Enttäuschungserlebnissen weiter veranschaulicht werden. Im Zuge eines Enttäuschungserlebnisses wird die Differenz zwischen der Erwartung und

¹⁷²Latour, 2006, S. 491.

¹⁷³Hubig, 2006, S. 221.

dem beobachteten Resultat als Störung empfunden. Störungen sind Erfahrungen, die im Erproben eines Umgangs auf die Verfasstheit der Technik als Medium verweisen. Das ist aber nur der Fall, wenn die Störungen identifizierbar sind. Die identifizierte Störung – als eine Art von Differenzerfahrung – begreift Hubig als Spur. Das Spurenlesen ist eine Form der Abduktion, demnach aufgrund der Spuren auf eine *Erklärung* geschlossen wird, wie der strukturierte Möglichkeitsraum (das Medium), den die Technik aufspannt, beschaffen sein muss, so dass das vorangegangene Agieren im Umgang mit der Technik zu dem beobachteten Ergebnis und nicht dem gewünschten Ziel geführt hat. Die beobachtete Differenz wird von Hubig als ein Lesen der *Spuren von* bezeichnet, die zu einer (erweiterten) Vorstellung der Technik als einem Möglichkeitsraum führt und auf die Bedingungen möglicher (zukünftiger) Nutzung als *Spuren für* verweist.

So sind z. B. bei Werkzeugtechniken die Spuren insofern *augenscheinlich*, als die Handlungen und ihre Ergebnisse (Effekte) seitens des Handelnden oft unmittelbar miteinander verknüpft werden können. Wenn ich mir auf den Finger schlage und nicht auf den Nagel, verweist die Schramme am Finger auf die Materialität des Hammerkopfes und die Schwungkraft, die durch den Stiel mitbedingt ist und somit auf die Möglichkeit, fest zuzuschlagen. Eine Potenzialität, die bei einer Anpassung des Schwungbogens genau die gewünschte Wirkung entfalten kann. Hubig beschreibt das wie folgt: „Direktes Feedback über Zustand und Leistung des Mittels erlaubt kontinuierliche und variable Intervention zum Zwecke der Optimierung des gewünschten Effekts“.¹⁷⁴

Diese Erfahrung des Medialen auf Basis eines Spurenlesens wird in erster Linie nicht als eine bloße intellektuelle Praxis verstanden, was heißen würde, dass die Differenz zwischen Erwartungen und Resultat qualitativ reflektiert und Spuren zielgerichtet interpretiert werden. Das unterscheidet das beschriebene Lesen von Spuren von der Praxis, die Krämer darunter fasst. Bei ihr handelt es sich um eine wissenschaftlich-theoretische Auseinandersetzung mit dem Medium (3.1.2). Spuren sind in meiner Interpretation Hubigs nicht etwas, das quantifizierbar ist, obwohl die Begrifflichkeit *Differenz* dieses nahelegt, sondern das Spurenlesen ist in erster Linie eine leibliche Erfahrung, die zu einem leiblichen Verstehen führt (Vgl. dazu 3.2.3). Merleau-Ponty betont in seiner Arbeit, dass ein leibliches Verstehen keiner Berechnungen oder expliziten Vergleiche bedarf, wie er es hinsichtlich von Technik u. a. anhand der Praxis des Autofahrens veranschaulicht. Eine intellektuelle Praxis spielt zu Beginn der Nutzung sehr wohl eine Rolle, um erste Erfahrungen des Umgangs einzuordnen, wenn

¹⁷⁴Hubig, 2015, S. 135.

auch dabei meist keine Berechnungen angestellt werden. In Abschnitt 3.2.3 habe ich das anhand von Merleau-Pontys bekanntem Beispiel des Erlernen des Schreibmaschinenschreibens, darstellt.

3.3.2 Schwierigkeiten einer Mensch-AI-Hybridisierung und Möglichkeiten eines Vertraut-werdens

Hinsichtlich des Verhältnisses von Menschen und intelligenten Techniken wird im Folgenden argumentiert, dass eine Hybridisierung problematisch ist. Christoph Hubig spricht mit Bezug auf intelligente Techniken (er nennt sie auch „transklassische Technik“) von einem Verlust von Spuren.¹⁷⁵ Damit meint er, dass im Umgang mit intelligenten Techniken Nutzungsspuren nicht wahrnehmbar sind und somit auch „die Möglichkeit der Rekonstruktion medialer Voraussetzungen verloren“ ginge.¹⁷⁶

Mit dem Ziel, die Schwierigkeit der Spurenwahrnehmung systematisch zu verstehen, wird im Folgenden wieder auf die Konzeption von Hubig Bezug genommen. Das anfängliche Aufeinandertreffen von Personen und *neuer* Technik wurde als das Erproben eines Umgangs beschrieben. Wenn Personen mit Techniken allgemein in ein Verhältnis eintreten, erscheinen die Möglichkeiten und Grenzen der Technik als zunächst unklar. Die basalen Vorstellungen über die Nutzungsmöglichkeiten einer im Speziellen neuen intelligenten Technik werden u. a. von früheren Erfahrungen mit einer vergleichbaren nicht-intelligenten Technik abgeleitet. Die Erwartungen an einen Umgang sind oftmals impliziter Natur und werden den Nutzer selbst häufig erst im Moment einer Enttäuschung bewusst.

Kommt es beim Umgang mit Techniken zu einer Nichterfüllung von Erwartungen, werden diese als Enttäuschungen empfunden. Das Auffinden der Ursache für eine Enttäuschung kann, wie im Abschnitt zuvor resümiert, die Grundlage für weitere praktische Schlüsse bilden, die es in Folge ermöglichen, dass die gewünschten Ziele erreicht werden. Identifizierte Enttäuschungen sind deshalb Spuren des Mediums, weil sie auf die Möglichkeiten und Grenzen der Nutzung verweisen. Die Schwierigkeit einer Spurenwahrnehmung zeigt sich oftmals darin, dass es gar nicht zu Enttäuschungen während des Umgangs kommt, weil die Technik beispielsweise Fehlbedienungen selbstständig verbessert oder die Erwartungen der Nutzer hinsichtlich des intelligenten Agierens zu unspezifisch sind. Bei längerer Nutzungsdauer kann eine Enttäuschung dann oftmals doch eintreten, wenn sich beispielsweise verschiedene Nutzer über die Leistungsfähigkeit des Systems austauschen.

¹⁷⁵Hubig, 2006, S. 185.

¹⁷⁶Hubig, 2006, S. 184.

Bei einer Enttäuschung zeigt sich die intelligente Technik genauso wie jede Technik in ihrer Begrenzung. Nur ist die Erfahrung eines Misslingens der Zielerreichung oftmals ohne Kontext. Das heißt, es bleibt für den Nutzer weitgehend unbestimmbar, wie das eigene Handeln mit den aktuell hervorgerufenen Systemeffekten zusammenhängt, die nicht der eigenen Erwartung entsprechen, oder welchen Einfluss andere Aspekte wie z. B. ein für den Umgang mit Technik als irrelevant erachtetes Verhalten oder die bisherige Nutzungshistorie haben. Im Erproben eines Umgangs mit Technik sind es gerade diese Enttäuschungen, die es Nutzern ermöglichen, das eigene Agieren damit zu verfeinern oder Erwartungen an die tatsächliche Leistungsfähigkeit des Systems anzupassen. Im Umgang mit intelligenten Techniken erfährt der Nutzer im Fall einer Enttäuschung häufig nicht, warum die Technik nicht im erwarteten Sinne funktioniert, und hat somit keine Erfahrung gemacht, die auf die Möglichkeiten eines zukünftig gelingenden Umgang verweist. Die Erfahrung einer Differenz zwischen Erwartung und Resultat ist demnach keine Spur des Mediums, da die Differenz ohne Kontext erscheint.

Mit anderen Worten kann man sagen, dass eine Differenzenerfahrung nicht qualifizierbar ist. Das ist so, weil die Differenzen ausschließlich binäre Qualität besitzen, das heißt, sie lassen nur die Bewertung zu, ob die Erwartungen erfüllt worden sind oder nicht („Gelingen oder Misslingen der Zielrealisierung und das Funktionieren des Prozesses (ja/nein)“, sie ermöglichen darüber hinaus jedoch keine weiteren Differenzierungen.¹⁷⁷

Zum Beispiel ist es gängige Praxis, dass intelligente Nutzerinterfaces übliche Eingabe- oder Bedienungsfehler automatisch korrigieren, beispielsweise nach der Verfahrensmaxime, dass etwas in dem von der Technik gedeuteten Kontext gar nicht so gemeint sein kann. Ist es aber kein Eingabefehler, sondern ein vom Nutzer intendierter Input – z. B. weil der Nutzer Informationen hat, für die das System sensorisch *blind* ist – wird das automatische Ausbessern zu einer Störung, da die Erwartungen des Nutzers enttäuscht wurden. Besitzt der Nutzer keine Informationen, die es ihm erlaubt, die Störung zu identifizieren, bleibt die Erfahrung des misslungenen Umgangs ohne Bezugsrahmen.

In dem Abschnitt zu Heinz von Foersterns Maschinenmetapher (3.1.5) habe ich argumentiert, dass es sich bei intelligenten Techniken auf der Sachebene um nicht-triviale Maschinen handelt, da sie u. a. geschichtsabhängig sind und dies ein Grund ist, warum sie in ihrem Verhalten nicht vorhersehbar sein können. In ebendiesem Abschnitt wurde Kaminskis Beschäftigung mit von Foersterns Ansatz diskutiert. Kaminski argumentiert, dass

¹⁷⁷Hubig, 2015, S. 136.

Personen in ein Verhältnis mit der Technik treten können, in dem das Verhalten der Technik erwartbar bzw. trivial erscheint, obwohl die Technik auf der *Sachebene* zumindest zeitweise nicht-trivial prozessiert.

Kaminski argumentiert, dass auch klassische Techniken, u. a. durch zeitweise Disfunktionalitäten, häufig nicht-triviale Tendenzen aufweisen. Beispielsweise ist das bei ausgefallenen Verschleißteilen oder festgefahrebenen Programmschleifen der Fall. Solche Disfunktionalitäten werden im Alltag *akzeptiert*, wenn Nutzern bekannt ist bzw. sie eine Routine entwickelt haben, um die Trivialität wiederherzustellen (Verschleißteile wechseln oder Computer neustarten). Es handelt sich hierbei um die Erwartung eines trivialen Verhaltens, die nicht nur das Prozessieren auf der Sachebene betrifft, sondern auch die Möglichkeiten, die Trivialität erster Ordnung (Sachebene) zu sichern bzw. zu managen („Trivialitätserwartungen zweiter Stufe“).¹⁷⁸ Eine solche triviale Einstellung zur Technik muss nach Kaminski aber erst im Umgang entstehen, das heißt, das Verhältnis muss sich im Modus des miteinander Vertraut-werdens allererst entwickeln. Eine höherstufige Trivialitätserwartung (zweiter Stufe) kann es mit intelligenten Techniken geben, wenn Personen beispielsweise ein *Gefühl* dafür entwickeln, welche systemischen Deutungen in der Regel richtig und welche fehlerbehaftet sind bzw. noch einmal überprüft werden sollten. Die Möglichkeit und Dauer eines Vertraut-werdens hängt aber von der Beschaffenheit der Technik ab. Kaminski sagt: „Durch die vorab in Sachen inkorporierte Trivialität kann es schneller und leichter zur Vertrautheit mit Technik [...] kommen als bei komplexeren, weniger trivialen Gegenständen“.¹⁷⁹

Intelligente Techniken können nicht so gestaltet werden, dass sie zu trivialen Maschinen auf der Gegenstandebene werden – das heißt, das Trivialität einfach inkorporiert wird –, da sie sonst ihre Leistungsfähigkeit verlieren. Vielmehr müssen Bedingungen geschaffen werden, damit ein Vertraut-werden ermöglicht wird. Damit ist der Bogen zurück zu den Spuren und der medialen Erfahrung geschlagen, da die Hybridisierung im beschriebenen Sinne als ein Vertraut-werden gesehen werden kann.

3.3.3 Die Hybridisierung hinsichtlich der zwei Typen von Mensch-AI-Verhältnissen im Ubiquitous Computing

Die Überlegungen zu den Grenzen einer Mensch-AI-Hybridisierung werden im Folgenden anhand der zwei Typen von Mensch-AI-Verhältnissen im Ubiquitous Computing, wie sie im Abschnitt 2.3 vorgestellt wurden,

¹⁷⁸Kaminski, 2010, S. 176.

¹⁷⁹Kaminski, 2010, S. 184.

diskutiert. Dieser Diskussion werden allgemeine Überlegungen zur Hybridisierung aus Sicht des Ubiquitous Computings vorangestellt.

Die Überlegungen zur Hybridisierung hinsichtlich der Designstrategien im Ubiquitous Computing

Eine *Intelligentisierung* von Technik hat unter anderem das Ziel, die Anforderungen an den Nutzer hinsichtlich seines Umgangs mit der infragestehenden Technik möglichst gering zu halten. Das ist u. a. in der Vision des Ubiquitous Computings zu finden (2.2). Dieses Ziel wird vor allem bei Systemen verfolgt, die bei der Lösung vielschichtiger Aufgaben unterstützen sollen und die, wenn sie mithilfe *klassischer* Techniken, die per Definition nicht-intelligente Techniken sind, erledigt werden sollten, aufwendig zu bedienen sind. Propagiert wird das Ideal intuitiver Bedienung, was eine unmittelbare Selbstverständlichkeit und das bereits Erfüllen von benötigten Fähigkeiten sowie die *richtigen* Vorstellungen über die Leistungsfähigkeit und Bedienung meint. Dagegen benötigt der Umgang mit *klassischen* Techniken, beispielsweise das Schreiben auf einer Tastatur, eine feinfühlig motorische Koordination, die erst erlernt werden muss und dabei nur in geringem Maße auf ein existierendes Repertoire an Fähigkeiten zurückgegriffen werden kann.

Ein unmittelbarer, selbstverständlicher Umgang mit Technik ist eine Illusion, da die Selbstverständlichkeit immer das Ergebnis eines Gewöhnungsprozesses ist. Bei jedem neuen Technikgebrauch können Nutzer Erfahrungen machen, die dazu führen, dass ihre Erwartungen in ein konkretes Handlungsziel oder allgemeine Erwartungen an den Einsatz der Technik enttäuscht werden, das heißt, dass der Umgang gestört ist. Die Möglichkeit, eine Störung als solche zu identifizieren, das heißt, *Gründe* für ebendiese zu eruieren und gegebenenfalls Schlüsse daraus zu ziehen, wie eine solche in zukünftiger Nutzung zu vermeiden ist, beispielsweise auch durch die Anpassung der eigenen Erwartung, macht ein Vertraut-werden aus.

Die Überlegungen zur Hybridisierung können nicht nur die Problemlagen im beschriebenen Sinne erklären, sondern lösungsorientiert zu einer Designstrategie führen, um Wege zu finden, das Ziel eines langanhaltenden selbstverständlichen Umgangs zu erfüllen, ohne auf die intelligente Funktionalität zu verzichten. Ein zentrales Moment einer solchen Designstrategie ist es, dass intelligente Techniken so gestaltet werden, dass, z. B. durch die Möglichkeit, Störungen zu identifizieren, die Möglichkeit erhalten bleibt, Spuren zu lesen.

Kontextsensitive Interaktion

Eine kontextsensitive Interaktion findet mit technischen Systemen dann statt, wenn diese sich infolge von Nutzerinputs und einer Deutung der Interaktionsumgebung Handlungen anpassen. Die Forschungsagenda des Ubiquitous Computing sieht es als eines ihrer Ziele an, ursprünglich nicht-technische Umgebungen *intelligent* zu technisieren oder klassische Technik mit intelligenten Fähigkeiten zu erweitern.

Ein Beispiel für kontextsensitive Systeme sind mit einer Reihe an digitalen Sensoren vernetzte *Smart-Home*-Umgebungen, in denen Situationen und Bedürfnisse von Bewohnern automatisch gedeutet werden und darauf basierend ein bestimmtes Systemverhalten ausgewählt wird. Kontextsensitive Systeme haben u. a. die Aufgabe, sich an die alltäglichen Gewohnheiten von Nutzern anzupassen. Ein intelligentes Heizsystem hat die Aufgabe, die Heiztemperatur zu regulieren und dabei das persönliche Temperaturempfinden der Bewohner, die Zeit ihrer An- und Abwesenheit sowie bestimmte Aktivitäten der Bewohner zu berücksichtigen. Anfängliche Erwartungen der Nutzer an die Funktionsweise und Leistungsfähigkeit des Systems sind durch die bisherigen Heizgewohnheiten der Bewohner festgelegt.

Im Idealfall ist die Beziehung von Menschen zu einem solchen System als *background relation* aufzufassen, das heißt, dass es sich um ein Mensch-Technik-Verhältnis handelt, in dem die Technik aus Sicht der Nutzer im besten Fall nie in Erscheinung tritt. Die Technik hängt wie beispielsweise das kommerzielle intelligente Thermostat *Nest* an der Wand und ein Interface zeigt ausschließlich die Zieltemperatur, die aktuelle Temperatur und die Tatsache an, dass es eingeschaltet ist. Im ersten Umgang und auch darüber hinaus kann das Thermometer die Bedürfnislage der Nutzer nicht immer hinreichend einschätzen. Daher werden die Erwartungen an das System nicht *immer* erfüllt.

Solange das Heizsystem auf gewünschte Weise funktioniert, agiert es im Verborgenen und hängt im beschriebenen Sinne unbeachtet an der Wand. Erst wenn die Wohnung als zu kalt oder zu warm empfunden wird, wird es als Störung empfunden. Die aktuelle Temperatur und die Zieltemperatur, die auf dem Display des Geräts angezeigt wird, verweist zwar auf die Aktivität, die gerade ausgeführt wird (Heizen oder nicht Heizen), doch weist nichts darauf hin, warum das System auf diese Weise agiert. Eine Langzeitstudie, die in *echten* Haushalten über den Umgang mit dem genannten intelligenten Heizsystemen *Nest* durchgeführt wurde, zeigte die Probleme auf, die sich daraus ergeben, wenn Personen mit Effekten eines intelligenten Systems konfrontiert sind, sie die Effekte aber nicht hinsichtlich ihrer

Entstehung bzw. Verursachung kontextualisieren können (dazu ausführlich in Abschnitt 4.2.3). Bei den meisten Teilnehmern führten die wiederkehrenden Erlebnisse von Störungen dazu, dass der Umgang mit dem System abgebrochen oder eine Nutzungspraxis entstand, die darin bestand, das System immer wieder auf die Werkseinstellung zurückzusetzen. Auch im zweiten Fall kam es nicht dazu, dass die Nutzung des Systems im Sinne der Hybridisierung zur Selbstverständlichkeit wurde, sondern dass sich die Nutzer vielmehr mit der *Widerspenstigkeit* des Systems gewöhnt hatten.

Die beschriebenen Beobachtungen zeigen, dass es zu einem Verlust von Spuren kam, die sich daraus ergaben, dass die Ergebnisse des Technikverhaltens nur hinsichtlich eines Erfüllens und Nicht-Erfüllens von Erwartungen bewertet werden konnten, somit keine Qualifizierung der Ergebnisse jenseits von Ja/Nein möglich war. Die Störungen sind in diesem Sinne nicht identifizierbar und somit keine bloßen Irritationen, die als Spuren wirksam werden können. Vielmehr können die Störungen im Sinne eines Nicht-Erfüllens von Heizerwartung eine Reihe von nicht identifizierbaren Ursachen haben. Es könnte sich um eine Fehlfunktion des Systems handeln, an den falschen Erwartungen des Nutzers gegenüber den Leistungen des Systems liegen oder ein Fehler in der Bedienung vorliegen. Zugespielt kann man sagen: Wenn die Nutzer die Gründe für eine Störung nicht nachvollziehen können, erscheint ihnen das System entweder als *Schrott* (bei Heinz von Förster: als nicht-triviale Maschine, die nach einem Trivialisierungsspezialisten verlangt) oder manchen sogar als ein *Anderes* mit widerspenstigem Charakter (vgl. alternate Relation).

Ein anderes Beispiel für ein kontextsensitives System sind Fahrerassistenzsysteme in einem Auto. Hier werden der Fahrer, das Fahrzeug, die Umgebung des Fahrer-Auto-Systems und andere Verkehrsteilnehmer sensorisch erfasst und im Sinne der Fahrsituation gedeutet. Zur kontextuellen Deutung der anderen Fahrer gehören oftmals nicht nur deren aktuelle relative Position zum eigenen Auto und deren Geschwindigkeit, sondern auch Vorhersagen hinsichtlich zukünftiger Situationen basierend auf den angenommenen Intentionen der Fahrer. Eine mögliche situationsbedingte Aktion des intelligenten Assistenzsystems kann es sein, in Situationen, die als gefährlich eingeschätzt werden, die Bremskraftverstärkung zu erhöhen, um den Fahrer bei einer gemäß den vorgenommenen Berechnungen anstehenden Bremsung zu unterstützen. Wenn sich das Auto selbstständig in einen solchen *Alarmzustand* versetzt, ändern sich die Bedingungen der Interaktion. Schon ein leichter Bremsdruck, der sonst nur eine leichte Bremsung verursachen würde, wird jetzt zu einer Vollbremsung. Wenn nach dem Bremsvorgang, der zwar durch den Fahrer ausgelöst wurde, bei dem das System bei der Durchführung allerdings auch mitgewirkt hat, die

Anteile des Systems nicht vermittelt werden, kann es zu falschen Zuschreibungen hinsichtlich der Ursache von Handlungen kommen. Eine Konsequenz kann sein, dass der Fahrer aufgrund seiner Erfahrungen von seinen schnellen Reflexen überzeugt ist und gefährliche Situationen auch in Zukunft nicht vermeidet.

Wie im ersten Beispiel gezeigt, kann ein Thermostat an der Wand hängen und bei einer Irritation ist an ihm nichts erkennbar, was auf die Ursache dieser Irritation schließen lässt. Damit weist auch nichts auf die Bedingungen einer gelungenen Nutzung hin, das heißt, es erscheint nicht im Lichte des Zuhandenen. Die Spuren vergangener Nutzung, die als gelungen oder nicht gelungen bewertet wurden, verweisen immer auf die Bedingungen möglicher Nutzung.

Self-Tracking-Techniken

Der zweite Typ von intelligenten Systemen im Ubiquitous Computing, die in Abschnitt 2.3 besprochen wurden, sind die Self-Tracking-Techniken. Der praktische Teil der Arbeit, vgl. dazu Kapitel 5, ist im Feld der Self-Tracking-Techniken verortet. Aus diesem Grund ist die Diskussion an dieser Stelle im Vergleich zur Diskussion der kontextsensitiven Systeme kürzer gehalten. Als Self-Tracking-Techniken werden technische Systeme verstanden, die personenbezogene Informationen aufzeichnen, diese hinsichtlich bestimmter Kriterien deuten und diese Deutung an die Personen, deren Informationen gedeutet wurden, zurückmelden. Die Rückmeldung kann unter Mithilfe von graphischen Visualisierungen auf einem Display Personen ermöglichen ihre alltäglichen Gewohnheiten besser zu verstehen und gegebenenfalls ihr Verhalten zu ändern.

Intelligente Self-Tracking-Systeme zeichnen sich dadurch aus, dass sie nicht nur personenbezogene Daten sensorisch erfassen – wie bei der Aufzeichnung von Vitalwerten (Herzschlag, Blutdruck) – sondern auch schon eine bestimmte Deutung auf Basis eines maschinellen Lernansatzes vornehmen. Dies ist unter anderem bei der Technik der Aktivitätserkennung der Fall, die der hier vorliegenden Arbeit als Anwendungsbeispiel dient. Bei der Technik der Aktivitätserkennung werden meist Bewegungssensoren (u. a. Beschleunigungssensoren), die am Körper getragen werden, als Grundlage für die Deutung bestimmter Handlungen verwendet. Die Kriterien für diese Deutung werden durch die Auswahl der Algorithmen und der Features in den Sensordaten und vor allem durch die Auswahl der Trainingsdaten bzw. der Settings zur Erzeugung der Trainingsdaten festgelegt. Beispiele für ein solches System ist die Deutung von Sport- und Freizeitaktivitäten. Die Motivation, die hinter der Entwicklung vieler solcher Systeme steht, ist ein Gewohnheitsmanagement im Sinne eines Ausbildens, Änderns und

Analysierens von Gewohnheiten. Bei Geräten der Aktivitätserkennung wie beispielsweise Fitnessarmbändern oder Anwendungen auf Smartwatches wird das alltägliche Fortbewegungsverhalten aufgezeichnet und als gelaufene Kilometer oder Schritte gedeutet bzw. angezeigt. Bei Anwendungen im Bereich der Medizin bzw. bei pseudo-medizinischen Geräten wird das gedeutete Verhalten zusätzlich als gesund oder ungesund klassifiziert.

Mit Blick auf die Hybridisierung taucht die Frage auf, ob der Umgang mit der Self-Tracking-Technik selbst zur Gewohnheit wird. Eine Gewöhnung im Sinne einer Hybridisierung setzt ein, wenn die technisch-vermittelte Perspektive auf die eigene Aktivität selbstverständlich wird. So wie die Brille beim Sehen verschwindet oder der Stock eines Blinden zum transparenten Zugang zur Welt wird, kann auch die automatische Deutung der eigenen Aktivität von einem vermittelten Zugang zu einem unmittelbaren werden, wenn das Gerät hinsichtlich seiner Medialität, das heißt seine Vermittlungsleistung, verschwindet. Genauso wie bei der kontextsensitiven Interaktion ist es für die Möglichkeiten einer Hybridisierung relevant, dass das maschinelle Deutungssystem hinsichtlich der Möglichkeiten und Grenzen erfahren wird.

Kapitel 4

Verstehbarkeit intelligenter Systeme: Kandidaten zur Vermittlung von Systemmechanismen

Das Verbinden von Mensch und Technik in einem funktionalen Verhältnis wurde im vorherigen Kapitel hinsichtlich der Hybridisierung erörtert. In diesem Kapitel wird die Hybridisierung von Menschen durch Technik, im Speziellen durch intelligente Technik, aus Perspektive der Technikgestaltung betrachtet.

Die *gelungene* Hybridisierung von Mensch und Technik stellt, so meine Annahme, ein für das Design von technischen Systemen mit intelligenten Befähigungen erstrebenswertes Ziel dar. Die Hybridisierung, verstanden als eine selbstverständliche und fortwährende Integration der Technik in den Alltag der Nutzer, deckt sich in zentralen Aspekten mit Designstrategien in der HCI und des Ubiquitous Computing, wie es in Abschnitt 2.2 gezeigt wurde.¹

Im vorherigen Kapitel wurde diskutiert, warum eine Verbindung im Sinne einer Hybridisierung von Menschen durch intelligente Systeme nicht möglich oder zumindest problematisch ist (siehe dazu ausführlich in Abschnitt 3.3). Unter anderem war das Konzept der Spur dazu geeignet, die Möglichkeiten und Grenzen einer Hybridisierung zu erklären. Spuren werden als *Hinterlassenschaften* einer Nutzung von Techniken verstanden, die, wenn sie gelesen werden, es Personen in der Rolle des Nutzers ermöglichen, das technische Systems hinsichtlich seiner Medialität zu *erfahren*. Spuren sind dabei aber keine konkreten Effekte, die wahrgenommen werden können, sondern Erfahrungen einer Differenz von Erwartungen und

¹Die Überlegungen in der HCI-Forschung und im Ubiquitous Computing zur „seamless integration“ und zur „embodied interaction“ beinhalten diesen Aspekt, siehe dazu Abschnitt 2.2.

Resultat einer Nutzung. Eine Störung im Umgang mit der Technik verursacht beispielsweise eine solche Differenz und verweist auf das Medium als Handlungsraum, in dem ein bestimmtes Agieren möglich oder nicht möglich ist. Ein mediales Verständnis des technischen Systems ist die Grundlage dafür, dass Nutzer in ein *stabiles* Verhältnis zur Technik eintreten, das heißt, dass sie u. a. praktisch wissen, worin die mögliche Wirksamkeit in verschiedenen Anwendungskontexten besteht und wie mit bestimmten Störungen umzugehen ist. Ich sprach bei dem Prozess der Hybridisierung von einem Vertraut-werden, der zu einer Selbstverständlichkeit des Umgangs führt.

Aufgrund dieser Konzeption einer Hybridisierung ergibt sich die Frage aus Sicht der technischen Gestaltung, wie intelligente Techniken gestaltet werden müssen, damit ein Lesen von Spuren als Grundlage für die mediale Erfahrung möglich bleibt. Das heißt, wie ist eine Hybridisierung möglich, ohne auf die intelligenten Befähigungen zu verzichten.

Im Umgang mit intelligenten Techniken kann es, wie bei herkömmlicher Technik, zu Differenzerfahrungen kommen, die als Störungen empfunden werden. Zum Beispiel macht ein intelligentes Assistenz, beispielsweise ein *recomendation system* für Restaurants nicht die Vorschläge, die im Erwartungsraum möglicher Antworten des Nutzers liegt oder ein intelligentes Heizsystem verändert die Raumtemperatur auf nichtgewünschte Weise. Der Unterschied liegt in der Möglichkeit einer Kontextualisierung der Differenzerfahrung. Die Kontextualisierung bzw. Identifizierung ist schwierig, da die Technik so beschaffen ist, dass die Mechanismen, das heißt, die Grundlagen ihrer Ergebnisproduktion, sich einer Zugänglichkeit entziehen.

Diese Unzugänglichkeit der Mechanismen des Systems, das heißt, die Prozesse konkreten Wirkens, müssen bis zu einem bestimmten Grade aufgehoben werden. Dabei ist es aber nicht das Ziel, die Systeme in ihrer Funktionalität, die zweifelsohne determiniert ist und ich deshalb von einem (digitalen) Mechanismus spreche, für den Nutzer *vollständig* verständlich bzw. nachvollziehbar zu machen, indem z. B. Pläne zur Sensorik und Verschaltung und der gesamte Code in Textform offengelegt werden. Vielmehr soll nach Aspekten der Mechanismen gesucht werden, das heißt nach Aspekten, die den Weg der Ergebnisproduktion (die maschinelle Deutung und daran anschließende Aktionen der Technik) in einer konkreten Nutzungsinstanz betreffen. Für eine systematische Untersuchung wird auf eine Diskussion im Feld in der Informatik Bezug genommen, die sich mit der Diskussion zur Verstehbarkeit (im engl. *intelligibility*) von Lernalgorithmen befasst.

In diesem Kapitel wird die Verstehbarkeit von intelligenten Systemen

auf Basis maschineller Lernansätze aus zwei Perspektiven diskutiert. Begonnen wird mit einer Diskussion der Verstehbarkeit von Ansätzen maschinellen Lernens aus der Perspektive der technischen Informatik, was exemplarisch anhand von zwei Lernansätzen erfolgt (4.1). Anschließend werden Konzepte aus dem Forschungsfeld der HCI diskutiert, in dem sich anhand von Konzepten und Nutzerstudien mit der Verstehbarkeit von intelligenten Systemen aus Sicht der Nutzer befasst wird (4.2). Im letzten Teil des Kapitels wird ein Framework entwickelt, welches die bestehenden Ansätze hinsichtlich des Interesses und den Voraussetzungen meiner Arbeit erweitert (4.3). Das Framework versammelt eine Reihe von Kandidaten zur Vermittlung von Systemmechanismen, die im folgenden Kapitel die Grundlage für die Nutzerstudien bilden.

4.1 Verstehbarkeit intelligenter Systeme im maschinellen Lernen

Das maschinelle Lernen ist ein weites Forschungs- und Anwendungsfeld der Informatik, welches auch in vielen anderen Disziplinen Einfluss erlangt hat. In Kapitel 2 wurde sich mit dem maschinellen Lernen im Rahmen einer Auseinandersetzung mit der Intelligenz technischer Systeme befasst. In Abschnitt 2.1.2 wurde auf die grundlegende Unterscheidung hingewiesen, nach der Ansätze maschinellen Lernens üblicherweise unterteilt werden. Demnach wird im maschinellen Lernen zwischen dem überwachten (*supervised*), unüberwachten (*unsupervised*) und dem bestärkenden (*reinforcement*) Lernen unterschieden.²

Im Folgenden wird auf das überwachte Lernen fokussiert, da ausschließlich dieser Typus für den späteren praktischen Teil von Interesse ist. Die grundlegenden Überlegungen aus diesem Kapitel lassen sich auch zum Teil auf die anderen Typen maschinellen Lernens übertragen, da viele Algorithmen in modifizierter Weise auch in den anderen Bereichen Anwendung finden.

4.1.1 Ein Datensatz und zwei gängige Lernstrategien

Es existiert eine nahezu unendliche Anzahl von Lernstrategien, die sich auf jeder Fachkonferenz erweitern. In diesem Abschnitt werden zwei sehr unterschiedliche Lernstrategien, zum einen die Entscheidungsbäume (*decision trees*) und zum anderen die künstlichen neuronalen Netze (*neural networks*) vorgestellt. Die Fülle von Ansätze im maschinellen Lernen können in dieser Arbeit nicht präsentiert werden, so dass sich auf diese verschiedenartigen

²Eine Einführung, siehe beispielsweise Alpaydın, 2008.

Strategien beschränkt wird. Grundlage dieser Vorstellung ist eine Arbeit von Kotsiantis³, der das heterogene Feld in einem *survey* strukturiert hat. Ein Ergebnis seiner Arbeit ist eine Übersicht in tabellarischer Form, die die unterschiedlichen Lernstrategien bezüglich ihrer Eigenschaften sortiert. Es ist für die Fragestellung dieses Kapitels von besonderem Interesse, wie die Lernansätze hinsichtlich ihrer Verstehbarkeit einzuschätzen sind. Kotsiantis bezeichnet diese Eigenschaft von maschinellen Lernansätzen als „Explanation ability/transparency of knowledge/classifications“.⁴ Die vollständige Tabelle ist in Abbildung 4.1 dargestellt. Die beiden Lernstrategien befinden sich hinsichtlich der Verstehbarkeit am oberen und unteren Ende der Skala. Die Bewertung der Eigenschaften folgt einer relativen Bewertungsskala von einem Stern bis zu vier Sternen.

	Decision Trees	Neural Networks	Naïve Bayes	kNN	SVM	Rule-learners
Accuracy in general	**	***	*	**	****	**
Speed of learning with respect to number of attributes and the number of instances	***	*	****	****	*	**
Speed of classification	****	****	****	*	****	****
Tolerance to missing values	***	*	****	*	**	**
Tolerance to irrelevant attributes	***	*	**	**	****	**
Tolerance to redundant attributes	**	**	*	**	***	**
Tolerance to highly interdependent attributes (e.g. parity problems)	**	***	*	*	***	**
Dealing with discrete/binary/continuous attributes	****	***(not discrete)	***(not continuous)	***(not directly discrete)	** (not discrete)	*** (not directly continuous)
Tolerance to noise	**	**	***	*	**	*
Dealing with danger of overfitting	**	*	***	***	**	**
Attempts for incremental learning	**	***	****	****	**	*
Explanation ability/transparency of knowledge/classifications	****	*	****	**	*	****
Model parameter handling	***	*	****	***	*	***

Table 4. Comparing learning algorithms (**** stars represent the best and * star the worst performance)

ABBILDUNG 4.1: Eigenschaften unterschiedlicher Ansätze maschinellen Lernens, übernommen von Kotsiantis, 2007, (S. 263)

Die Diskussion der beiden Lernansätze erfolgt anhand eines Anwendungsbeispiels botanischer Pflanzenbestimmung. Das Beispiel wurde ausgewählt, da es sich auf anschauliche Weise dafür eignet, die Ansätze zu

³Kotsiantis, 2007.

⁴Kotsiantis, 2007, S. 263.

vergleichen. Ein überwachter Lernansatz ist dadurch charakterisiert, dass das Lernen anhand eines Trainingsdatensatz mit gelabelten Daten erfolgt. Der Datenansatz, der die Grundlage für das Anwendungsbeispiel bildet, ist das *iris dataset*. Dabei handelt es sich um einen klassischen Datensatz im maschinellen Lernen, der zu Test- und Demonstrationszwecken Bekanntheit erlangt hat.

Der Datensatz basiert auf Pflanzenbeobachtungen des Biologen Ronald Fischer auf der Halbinsel Gaspésie, dessen Aufzeichnungen die Basis für einen 1936 veröffentlichten Artikel bildete.⁵ Dieser Datensatz dokumentiert Funde der Pflanzengattung *Iris* in drei Arten (*Iris Setosa*, *Iris Virginica* and *Iris Versicolor*). Das Datensatz besteht aus 150 Stichproben (Samples), die neben dem Namen der Art auch die Länge und Breite der Kelchblüte (biol. *Sepalum* oder äußere Blütenhülle) und des Kronblattes (biol. *Petalum* oder innere Blütenhülle) beinhaltet. Insgesamt besteht der Datensatz aus vier Merkmalen (Features), die genannten Längen und Breiten der inneren und äußeren Blütenhülle, und drei Klassen, die durch die drei Arten der *Iris* gebildet werden.

Länge Sepalum	Breite Sepalum	Länge Petalum	Breite Petalum	Art
4.6	3.6	1.0	0.2	I. Setosa
4.4	3.0	1.3	0.2	I. Setosa
...	I. Setosa
5.1	2.5	3.0	1.1	I. Versicolor
5.6	2.5	3.9	1.1	I. Versicolor
...	I. Versicolor
6.2	2.8	4.8	1.8	I. Virginica
6.4	2.8	5.6	2.1	I. Virginica
...	I. Virginica

TABELLE 4.1: Ausschnitt aus dem klassischen Iris Datensatz. Insgesamt gibt es 150 Samples. Der Datensatz kann u. a. hier gefunden werden: (IrisDataset, 1937)

4.1.2 Entscheidungsbäume

Ein Entscheidungsbaum besteht aus einer Reihe an Entscheidungsknoten, bei denen in der Laufzeit des Algorithmus (Klassifikation) ein Sample an einem Knoten anhand eines bestimmten Merkmals (*feature value*) geprüft wird und je nach Ergebnis dieser Prüfung eine Verzweigung gewählt wird, die zu einem weiteren Knoten führt, der ein weiteres Merkmal prüft oder

⁵Fisher, 1936.

ein Klassifikationsergebnis darstellt (Blattknoten). Der Prozess beginnt immer bei der Wurzel (*root node*) und wird so lange fortgesetzt, bis ein Blattknoten erreicht wird. Ist der Blattknoten erreicht, ist eindeutig bestimmt, welcher Kategorie das Sample zugehörig ist. Das Merkmal, anhand dessen im Knoten entschieden wird, kann ein bestimmter Schwellenwert (*thresholds*) sein oder eine Frage, die mit ja oder nein beantwortet werden kann bzw. mehr Auswahlmöglichkeiten hat.

Hinsichtlich der Erstellung eines Entscheidungsbaums, das heißt dem Training, wird zwischen manuellen und automatischen Ansätzen unterschieden. Für einfache Datensätze können manuelle Bäume erstellt werden, das heißt, die Entwickler legen die gesamte Struktur, das heißt die Knoten, Verzweigungen und die Entscheidungsmerkmale fest. Bei manuellen Bäumen handelt es sich im engeren Sinne nicht um einen maschinellen Lernansatz, sondern ist vielmehr als *hard-coding* zu verstehen. Ein System auf Basis eines manuellen Baumes erfüllt damit nicht das Charakteristikum eines intelligenten Systems, wie es schon in der Vorarbeit definiert wurde (2.1.2). Das ist nicht der Fall, wenn die Baumstruktur das Ergebnis eines automatischen Prozesses ist. In der Trainingsphase wird demnach ein automatischer Prozess durchlaufen, der zu einer geeigneten Architektur von Knoten und Merkmalen führt. Der automatische Prozess kann anhand von unterschiedlichen Strategien ablaufen, die im Feld des maschinellen Lernens als Heuristiken bezeichnet werden. Die gängigste Heuristik ist die sogenannte „Teile- und-Herrsche-Strategie“: Das Ziel dieser Strategie ist es, nach Features zu suchen, die die Trainingsdaten an jedem Knoten am besten in zwei gleichgroße Gruppen aufteilt.⁶ Das Feature, das die Menge aller Trainingsdaten am besten in zwei Teile aufteilt, bildet die Wurzel des Baums. Im nächsten Schritt wird in den jeweiligen Hälften wieder nach dem Feature gesucht, welches die Daten dieses Strangs am besten aufteilt. Dieser Prozess wird fortgesetzt, bis alle Samples eindeutig ihrer Klasse zugeordnet worden sind.

Beispiel

Das *iris dataset* dient als Grundlage, um die Implementierung eines Entscheidungsbaums zu veranschaulichen. Wollte man eine Baumstruktur manuell erstellen, könnte man die Tabelle dahingehend untersuchen, welche Merkmale oder Merkmalskombinationen charakteristisch für eine Art der

⁶Siehe beispielsweise Alpaydın, 2008, S. 183.

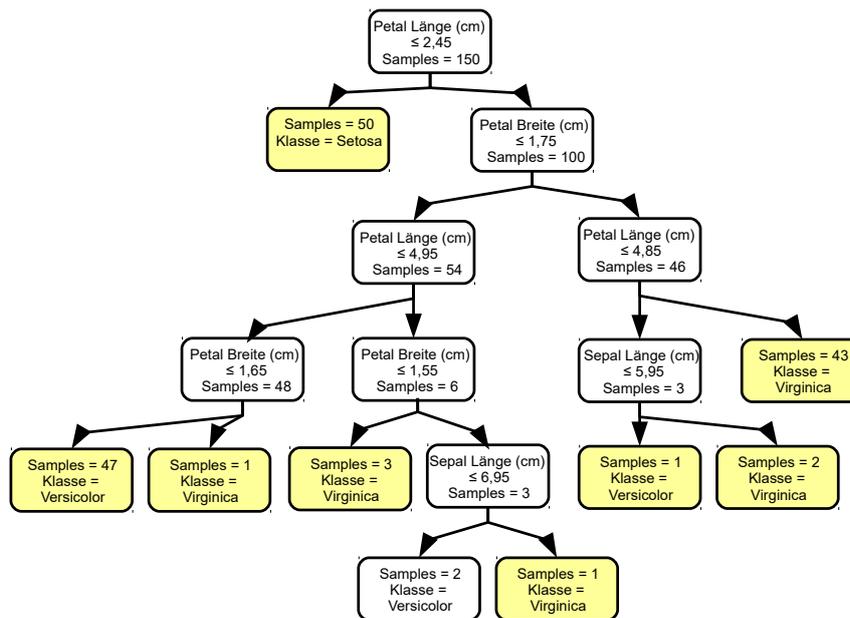


ABBILDUNG 4.2: Hierbei handelt es sich um einen Entscheidungsbaum, der aus dem Iris Datensatz automatisch erstellt wurde. Die gelben Felder markieren die Blattknoten des Entscheidungsbaumes.

Gattung Iris sind. Beispielsweise hat die *Iris Virginica* vergleichsweise lange Blüten und die *Iris Setosa* vergleichsweise kurze. Eine zuverlässige Festlegung der Grenzwerte in den Knoten ist manuell schwierig. Eine automatische Erstellung, die nach der oben beschriebenen *Teile- und-Herrsche*-Strategie abläuft, ist besser geeignet. Das Ergebnis eines solchen automatischen Prozesses ist in Abbildung 4.2 dargestellt.⁷ Man kann sehen, dass die Länge des Petalum ein geeignetes Feature (oberster Knoten) ist, um zwischen der *Iris Setosa* und den anderen beiden Gattungen zu unterscheiden. Die Unterscheidung zwischen den beiden übriggebliebenen Gattungen ist dann schwieriger und hängt von einer Kombination von Merkmalen ab.

Das Ziel des Trainings ist ein Algorithmus, der dem Anspruch einer Generalisierung gerecht wird, das heißt, dass das System nicht nur hinsichtlich der Trainingsdaten, sondern auch bei unbekanntem Samples aus dem Bereich eine zuverlässige Einschätzung bzw. Klassifikation ermöglicht. Es besteht der Anspruch, dass, wenn man dem im Beispiel gezeigten Klassifikationssystem die Werte von Messungen der Blüten und Blätter von aktuellen Funden übergibt, das System richtig einschätzt, um welche Art der Pflanze es sich handelt. Voraussetzung ist, dass die Funde, die in

⁷Die Abbildung orientiert sich an einer Darstellung in ScikitLearn, 2017.

den Trainingsdaten versammelt wurden, repräsentativ sind, und die relevanten, weil diskriminierenden, Merkmale dokumentiert wurden.

Verstehbarkeit

Entscheidungs bäume werden als Algorithmen gesehen, die einfach zu verstehen sind, dass heißt, die Entscheidungen des Algorithmus gelten als potentiell einfach nachvollziehbar (höchste Bewertung in der Kategorie „Explanation ability“ in der Abbildung 4.1). Der Grund liegt darin, dass es sich um eine hierarchische Entscheidungsfindung handelt, die Schritt für Schritt nachverfolgt werden kann. Deshalb wird dieser Ansatz da, wo Verstehbarkeit wichtig ist, Verfahren mit höherer Genauigkeit vorgezogen.⁸

Bei Entscheidungs bäumen mit einer sehr großen Anzahl von Merkmalen gilt diese Charakterisierung nur bedingt. Die Schritte der Klassifikation nachzuverfolgen kann zu einem langwierigen Unterfangen werden. Es hängt hier, wie allgemein, wenn es um eine Verstehbarkeit von Algorithmen geht, davon ab, in welchem Detailgrad ein Verständnis ermöglicht werden soll. Auch bei Systemen mit vielen relevanten Merkmalen ist der Entscheidungsbaum nur bedingt geeignet, detailliert jedes Klassifikationsergebnis nachzuvollziehen, kann aber geeignet sein, die wurzelnahen und damit die in besonderer Weise diskriminierenden Merkmale zu nutzen, um ein erstes Verständnis über die Mechanismen, die zu einem bestimmten Klassifikationsergebnis führen, zu erlangen.

Es ist Ziel dieses Kapitels zu diskutieren, wie Aspekte von Mechanismen, die einem intelligenten System zugrundeliegen, vermittelt werden können, damit ein Spurenlesen und damit eine mediale Erfahrung möglich ist. Auf Grundlage des *iris datasets* wäre es vorstellbar, ein *intelligentes* System auf Basis eines Entscheidungsbaums für Endnutzer zu entwickeln. Bei dem System könnte es sich um eine Smartphone-Applikation handeln, die Wissenschaftler oder Laien dazu nutzen können, die Arten der Gattung Iris zu bestimmen. Nach Eingabe der Messdaten, d. h. der Länge und Breite der Blätter und Blüten, würde die Benennung der zugehörigen Irisart auf dem Bildschirm erscheinen. Eine zusätzliche Vermittlung von Aspekten der Baumstruktur, die für die Kategorisierung der Gattung ausschlaggebend war, könnte zum Verständnis des Klassifikationsergebnisse beitragen.

4.1.3 Künstliche neuronale Netze

Eine Einführung in die künstlichen neuronalen Netze beginnt meist mit einem Verweis auf das menschliche Gehirn. Das Konzept, wonach das menschliche Gehirn als neuronales Netze zu erklären versucht, diene als

⁸Alpaydın, 2008, S. 184.

Inspiration für die Entwicklung dieses maschinellen Lernansatzes. Die künstlichen neuronalen Netze sind ein eigenes Forschungs- und Anwendungsfeld, welches sich unabhängig von der Hirnforschung entwickelt hat. Im Folgenden wird sich ausschließlich mit künstlichen neuronalen Netzen als effizienten Lernalgorithmen befasst, welche in verkürzter Weise als neuronale Netze bezeichnet werden.

Eine zentrale Eigenschaft neuronaler Netze ist es, dass sie in elementarer Weise von einem Autoadaptionsprozess abhängen. Das heißt, es gibt eine Struktur, die sich während der Trainingsphase in einer Vielzahl von Optimierungsschritten selbst ausbildet.

Ein neuronales Netz besteht aus sogenannten Perzeptronen, wobei es sich dabei um Knoten handelt, die netzförmig angeordnet sind: „A neural network consists of large number of units (neurons) joined together in a pattern of connections“.⁹ Die Perzeptronen sind dabei in Schichten angeordnet, wobei jedes Perzeptron in einer Schicht mit allen Perzeptronen der vorherigen sowie der nächsten Schicht verbunden ist. Bei einem zweilagigen Netzwerk gibt es die Inputs, die selbst nicht als Schicht gesehen werden, die Perzeptronen, die die versteckte Schicht (*hidden layer*) bilden und die Schicht von Outputs.

Jedes Perzeptron im Netz besitzt einen Wert. Der Wert eines Perzeptrons ist sein Gewicht und berechnet sich aus der Summe aller Eingangswerte („activation values“), das heißt den Gewichten der Perzeptronen der vorherigen Schicht anhand einer Schwellwertfunktion („activation function“).¹⁰ Der so berechnete Wert eines Perzeptrons ist der Eingangswert für die Berechnung der Perzeptronen in der nächsten Schicht oder der Ausgangsperzeptronen der letzten Schicht. Die Ausgangsperzeptronen repräsentieren das Klassifikationsergebnis. Die Schwellwertfunktion wird vorher festgelegt und ist für die Berechnung aller Perzeptronenwerte in einem Netz gleich. Die Schwellwertfunktion kann unterschiedlich aussehen: Es kann sich beispielsweise um eine binäre Funktion handeln, die den Wert 1 zurückgibt, wenn ein bestimmter Schwellwert überschritten wurde und 0 wenn er nicht überschritten wurde. Wenn ein Ausgangsperzeptron für ein bestimmtes Klassifikationsergebnis steht, ist dies der Fall, wenn der Schwellwert überschritten wurde und nicht der Fall, wenn er nicht überschritten wurde. Eine weitere gängige Schwellwertfunktion ist die sogenannte *Sigmoid* Funktion. Die Schwellwertfunktion, die Auswahl der relevanten Parameter aus den Trainingsdaten (Eingangswerten) und die Anzahl der Schichten werden in der Initialisierung festgelegt. Dazu gibt es viele Publikationen, die die Vor- und Nachteile von Schichtanzahl und der

⁹Kotsiantis, 2007, S. 255.

¹⁰Kotsiantis, 2007, S. 255.

Schwellenwertfunktionen diskutieren. Diese Architektur des Netzes bildet die Grundlage für das Training. Ziel des Trainings ist es, die Gewichte in den Knoten so anzupassen, dass das System für die Eingangsgrößen des Trainingsdatensatzes die zugehörigen Ausgangswerte besitzt, das heißt, die bestmögliche Klassifikation für alle Trainingsdaten zu erreichen. In jedem Schritt des Prozesses der Autoadaptation werden die Gewichte in den Perzeptronen angepasst. Die Anpassung erfolgt aufgrund der Evaluierung der Ergebnisse des aktuellen Netzes. Da sich die Gewichte erst während dieses Prozesses einstellen und man zu Beginn also noch nichts über sie weiß, müssen Startwerte eingestellt werden. Eine gängige Praxis für die Festlegung der Startwerte ist es, kleine zufällige Werte im Intervall $[-0,01; 0,01]$ anzunehmen.¹¹

Die Autoadaptation kann unterschiedlichen Strategien folgen, die gängigste ist die *Backpropagation*. Wie diese abläuft, wird im Folgenden skizziert.

Ein Schritt im Prozess der Autoadaptation wird als *Epoche* bezeichnet. Folgt man der Beschreibung von Kotsiantis, hat eine Epoche sechs Schritte, die sich wiederholen, bis eine als ausreichend erachtete Qualität der Klassifikation erreicht wird.¹² Was für ausreichend gehalten wird, hängt von der Anwendung ab, beispielsweise der Qualität der Trainingsdaten und den Anforderungen an die Leistungen des Systems.

1. Ein Sample der Trainingsdaten wird initialisiert.
2. Der Output des Netzes wird mit dem *wahren* Label verglichen. Für jeden Ausgangswert (Outputperzeptron) wird der Fehler berechnet. Handelt es sich bei der Anwendung um eine Klassifikation, ist das Ergebnis richtig oder falsch, das heißt der Fehler ist entweder eins oder null.
3. Für jeden Ausgangswert wird die Differenz zwischen dem vom System klassifizierten und dem wahren Wert berechnet (Skalierwert). Dabei handelt es sich um den sogenannten lokalen Fehler.
4. Die Gewichte, die zu jedem Ausgangsneuron führen, werden so angepasst, dass sich der Fehler verringert. Das heißt, die Gewichte in der vorherigen Schicht werden so verändert, dass sich für das aktuelle Sample das zugehörige Klassifikationsergebnis einstellt.
5. Mit Blick auf die vorherige Ebene, deren Fehler man nicht direkt berechnen kann (es gibt ja keinen *wahren Wert*), werden deren Gewichte

¹¹Vgl. Alpaydm, 2008, S. 263.

¹²Kotsiantis, 2007.

über das Konzept der *Verantwortlichkeit* („blame“) für die Werte der Ausgabeschicht berechnet.

6. Der Vorgang wird auf der vorherigen Ebene wiederholt, wobei die *Verantwortlichkeit* hier als lokaler Fehler angenommen wird.

Die genannten Schritte werden in einem Autoadaptionprozess vielfach durchlaufen, bis zu einem Zeitpunkt, bei dem der Klassifikationsfehler als klein genug angenommen wird.

Beispiel

Das *iris dataset* kann auch anhand eines neuronalen Netzwerks repräsentiert werden. Auch wenn das neuronale Netz genauso wie ein Entscheidungsbaum Knoten besitzt, die bestimmte Schwellwerte haben, ist die Funktionsweise eine ganz andere. Das Netz ist nicht hierarchisch strukturiert und ist zudem komplexer, da es mehr Knoten und Verbindungen gibt, die alle einen Einfluss auf das Gesamtergebnis haben.

In einem Autoadaptionprozess, wie er im Abschnitt zuvor beschrieben wurde, könnte ausgehend von zufälligen Gewichten in den Knoten und mithilfe eines Samples im Datensatz eine erste Klassifikation erfolgen. Für jedes Sample im Trainingsdatensatz würden die lokalen Fehler berechnet und die Gewichte so angepasst, dass die Fehler möglichst verschwinden. Schritt für Schritt werden die Knoten so optimiert, dass ein Klassifikator auf Basis eines neuronalen Netzes entsteht.

Um das Potential von neuronalen Netzen aufzuzeigen, wird ein weiteres Beispiel skizziert, das auf einem möglichen anderen Iris-Datensatz beruhen könnte. Man könnte sich vorstellen, dass es einen Iris-Datensatz gibt, der nicht auf den genannten Abmessungen von Blüten und Blättern basiert, sondern auf einer Reihe von Fotos der Gattung Iris. Fotos als Grundlage für eine Klassifikation von Inhalten auf dem Foto ist ein klassisches Anwendungsgebiet der *computer vision*. Die Features sind dann nicht die Zahlenwerte, die für die Abmessungen der Blüten und Blätter der Pflanze stehen, sondern die Pixel der Fotos. Damit sind es nicht nur vier Eingangsfeatures, wie beim klassischen *iris dataset*, sondern je nach Bildauflösung und Farbbandbreite viele Millionen. Je nach verwendeten Lernalgorithmus kann es üblich sein, die Features in den Bildern zu reduzieren, indem Pixel zusammengefasst, nur die Farbverläufe verwendet (Farbhistogramm) oder das Bild auf seine Kanten reduziert werden. Aktuelle leistungsstarke neuronale Netze könnten aus den Bildern lernen, welche Art der Iris auf dem Foto dargestellt ist. Die Art der Aufnahmen legen fest, welche Merkmale in den Bildern überhaupt für den Algorithmus zugänglich sind. Sind die Fotos alle aus einem bestimmten Winkel aufgenommen, der die Längen von

Blüten und Blättern in der richtigen Relation darstellt, kann es sein, dass auch auf Basis dieses Datensatzes die Abmessungen der Blüten und Blätter die diskriminierten Merkmale sind, welches der neuronale Klassifikationsalgorithmus *vermittelt* als Grundlage nutzt, um die Arten der Iris auf den Bildern zu unterscheiden.

Verstehbarkeit

Algorithmen auf Basis von neuronalen Netzen eignen sich nur wenig für einen verstehenden Zugriff und somit sind auch Aspekte des Mechanismus, die die Entscheidungsfindung betreffen, schwer darstellbar. Werden die Perzeptronen und ihre Gewichte in einer Grafik visualisiert, ist schon bei nur vier Features, wie es im klassischen *iris dataset* der Fall ist, nicht abzulesen, welche Features für ein bestimmtes Ergebnis ausschlaggebend sind. Da jedes Perzeptron in einer Schicht mit allen anderen der nächsten Schicht verbunden ist, beeinflusst jedes Perzeptron das Ergebnis und macht den Zusammenhang zwischen Features und Ergebnis nur schwer nachvollziehbar.¹³

4.1.4 Fazit: Verstehbarkeit maschineller Lernalgorithmen

Lernende Algorithmen hinsichtlich ihres Wirkens anschaulich darzustellen und somit zu ermöglichen, dass das algorithmische Prozessieren in der Anwendung verstehbar bzw. nachvollziehbar ist, ist vor allem für die Gestalten solcher Systeme von Interesse.

Ein Verständnis darüber, wie der Prozess der Klassifikation abläuft, kann helfen aufzuzeigen, welche Merkmale in den Daten die Grundlagen für die Klassifikation bilden. Ein solches Wissen kann wichtig sein, wenn es bestimmte Eigenschaften in den Trainingsdaten gibt, die nicht repräsentativ für das ganze System sind und deshalb auch nicht den resultierenden Algorithmus determinieren sollen. Beispielsweise könnten die Fotos für den alternativen Iris-Datensatz mit unterschiedlichen Kameras produziert worden sein und das Lernsystem nicht eine Eigenschaft der Gattung als Grundlage für die Klassifikation verwendet haben, sondern in Wirklichkeit die Eigenschaft der Kamera. Kameras haben Charakteristiken, die in die Bilder eingehen (Kameras produzieren beispielsweise einen bestimmten *Look*). So

¹³Es gibt Methoden, neuronalen Netzen zu *entlocken* welche Feature in den Sensordaten für das Funktionieren des neuronalen Netzes von besonderer Relevanz sind. Beispielsweise kann eine solche Analyse ergeben, dass ein System, das Fotos dahingehend klassifiziert, welche Sportart auf dem Bild praktiziert wird, vor allem die Hintergrundfarben verwendet, um zwischen den Sportarten zu unterscheiden. Für solche Methoden siehe Zeiler und Fergus, 2014, Grün et al., 2016.

kann es passieren, dass ein Algorithmus, der auf Basis dieser Bilder trainiert wird, fälschlicherweise nicht Merkmale des Dargestellten als Diskriminiertem auffasst, sondern Merkmale der Kamera verwendet. Wenn die verschiedenen Spezies beispielsweise mit verschiedenen Kameras aufgenommen wurden, könnte es sein, dass das System scheinbar optimal zwischen den Gattungen der Iris unterscheiden kann. Wenn die Gestalter die diskriminierenden Features als *falsch* erkennen, können sie Schritte unternehmen, die ein Training auf Basis der Kameraeigenschaften verhindert, beispielsweise durch ein Filtern der Daten oder ein gezieltes *Verrauschen* der Daten. Letzteres wäre zum Beispiel hilfreich, wenn die Kameras große Qualitätsunterschiede haben und die Fotos durch die Rauschschaltung auf ein niedrigeres, aber einheitliches Level gesetzt werden können.

In das Framework, das ich im letzten Abschnitt dieses Kapitels entwerfe, werden die Überlegungen zur Verstehbarkeit von Algorithmen einfließen. Dass die Endnutzer eines intelligenten Systems dieses vollständig verstehen, ist praktisch nicht möglich. Trotzdem werden hinsichtlich des Frameworks Überlegungen angestellt, wie einzelne Aspekte der maschinellen Klassifikation für den Nutzer in bestimmten Handlungssituationen darzustellen sind.

Die Verschiebung hin zu einer Nutzerperspektive wird in dem folgenden Abschnitt (4.2) vorbereitet. Auch in der HCI-Forschung sind die Probleme, die sich aus einem nicht vorhandenen Zugang zu den Mechanismen für eine erfolgreiche Nutzbarkeit von intelligenten Systemen ergeben, thematisiert worden. Es gibt eine Reihe von Ansätzen, die sich konzeptuell und empirisch mit dieser Problemlage befassen. Im Folgenden werden diese Konzepte diskutiert, um im letzten Teil des Kapitels in eine eigene Konzeption, die sich im Framework (4.3) kondensiert, vorzustellen.

4.2 Verstehbarkeit intelligenter Systeme in der HCI

Im Forschungsbereich der HCI gibt es eine Reihe von Ansätzen, die sich mit den Herausforderungen auseinandersetzen, die angesichts von intelligenten Systemen, die für alltägliche Nutzer als Black-Boxes erscheinen, entstehen. Als Herausforderungen werden u. a. ein mangelndes Vertrauen und eine geringe Akzeptanz genannt, wie das folgende Zitat zeigt.

„Context-aware intelligent systems employ implicit inputs, and make decisions based on complex rules and machine learning models that are rarely clear to users“. Such lack of system intelligibility can lead to loss of user trust, satisfaction and acceptance

of these systems“.¹⁴

In der Forschung werden Designstrategien diskutiert, die es möglich machen sollen, dass das intelligente Systemverhalten für die Nutzer nachvollziehbar wird. Der Fokus liegt dabei vor allem auf zusätzlichen, erklärenden Informationen, die den Nutzern während der Interaktion vermittelt werden. Die testweise Realisierung der Designansätze findet im Bereich sogenannter *webservices* statt, wozu Suchmaschinen, Spamfilter sowie *recommendation systems* gehören.

Im Folgenden werden zwei Ansätze präsentiert, die sowohl konzeptuelle Vorschläge unterbreiten als auch empirische Untersuchungen durchführen, um die Wirksamkeit ihrer Konzepte aufzuzeigen.

4.2.1 Verstehbarkeit durch Erklärungen

Eines der ersten Artikel zu diesem Thema ist das von Lim et al., aus dem auch das eingangs verwendete Zitat stammt.¹⁵ Sie entwickeln einen Ansatz, der es ermöglichen soll, dass Nutzer während des Umgangs mit einem intelligenten System ein weitergehendes Verständnis der Systeme entwickeln. Dieses soll durch die Vermittlung von Erklärungen, die das System als Ganzes bzw. das aktuelle Verhalten betreffen, während der Laufzeit des Systems, erreicht werden. Das intelligente System, das sie in der Veranschaulichung und dem Testen des Ansatzes verwenden, stammt aus dem Bereich kontextsensitiver Systeme. Genauer gesagt handelt es sich um ein einfaches System zur Aktivitätserkennung, das aufgrund bestimmter Messwerte erkennen kann, ob eine Person gerade sportliche Übungen macht oder nicht. Der zugrundeliegende maschinelle Lernansatz ist ein Entscheidungsbaum.

Im konzeptuellen Teil der Arbeit wird die theoretische Dimension von Verständlichkeit (sie sprechen von einer „intelligibility“) durch die Vermittlung von Erklärungen aufgespannt. Dabei beziehen sie sich auf Konzepte aus dem Bereich des Wissensmanagements und der Expertensysteme.¹⁶ Die Autoren nennen mehrere Kategorien zur Systematik von Erklärungen, wobei sie für ihren Ansatz den Typus „reasoning trace“ als besonders relevant erachten. *Reasoning trace* meint eine Abfolge von Ursachen und Wirkungen bzw. Schlüssen auf Basis bestimmter Kriterien, die letztendlich zu einem Ergebnis führen. Die Erklärungen, die diese Abfolge nachvollziehbar machen sollen, sind die Antworten auf die sogenannten *W-Fragen*. Das heißt, *wie* und *warum* kommt das System zu einem bestimmten Ergebnis

¹⁴Lim, Dey und Avrahami, 2009, S. 2119.

¹⁵Lim, Dey und Avrahami, 2009.

¹⁶Siehe Lim, Dey und Avrahami, 2009, S. 2120.

(Deutung oder Handlungsentscheidung). Zudem werden Erklärungen zu einem alternativen Verhalten des Systems gegeben, die sich in *warum-nicht*-Erklärungen äußern und Erklärungen, die die Randbedingungen betreffen, das heißt, wenn ein bestimmtes Ergebnis der Fall wäre, dann müssten die Eingangsgrößen so beschaffen sein.

Auf Basis dieses Ansatzes, der ein Verständnis durch verschiedene Erklärungen erzeugen soll, präsentieren die Autoren eine Nutzerstudie. Die Studie basiert auf dem eingangs erwähnten System zur Klassifizierung von Aktivitäten in einem binären Klassifikationssystem. Das System klassifiziert menschliche Aktivität entweder als eine Sportübung („*exercises*“) oder als keine Sportübung („*no exercises*“).¹⁷ Grundlage für die maschinelle Kategorisierung sind drei Typen von Inputs und zwar Körpertemperatur, Puls und *Tempo* (Als *Tempo* wird die Geschwindigkeit von Körperbewegungen verstanden). Die Klassifikation im Testszenario erfolgte auf Basis eines einfachen Entscheidungsbaums, der aus vier Knoten aufgebaut war.

In der Studie konnten die Teilnehmer das Verhalten des Systems beobachten, das aus den Eingangswerten, die Klassifikation *Sportübung* oder *keine Sportübung* vornahm. Zusätzlich zu den Eingangs- und Ausgangsgrößen wurden den Nutzern Erklärungen präsentiert, die den konkreten Nutzungsfall betrafen. Bei den Erklärungen wurde zwischen verschiedenen Erklärungstypen unterschieden, wobei jeder Typus eine Antworten auf eine der *W-Fragen* lieferte. Zudem gab es eine Testgruppe, die keine Erklärung zu sehen bekam. Ein Beispiel für eine solche Erklärung ist: Die Aktivität wurde als *keine Sportübung* eingeschätzt, weil die Körpertemperatur unter X Grad lag und das Tempo kleiner als Y war.

Am Ende wurde evaluiert, ob ein Verständnis vermittelt wurde und ob sich durch Vermittlung das Vertrauen in das System verändert hatte. Ein Verständnis des Systems wurde auf zwei Arten getestet: Zum einen mussten die Teilnehmer fehlende Werte ergänzen und zum anderen bei einem „reasoning test“ die Klassifikationen des Systems begründen.

Die Auswertung zeigte, dass es eine klare Steigerung des Verständnisses gab, wobei die *Warum-* und *Warum-nicht*-Erklärungen die beste Unterstützung lieferten. Zudem gab es eine klare Steigerung im Empfinden von Vertrauen, wenn den Personen Erklärungen vermittelt wurden: „The Why and Why Not explanations improved participants understanding, increased their trust in the system, and their task performance“.¹⁸

¹⁷Lim, Dey und Avrahami, 2009, S. 2120.

¹⁸Lim, Dey und Avrahami, 2009, S. 2126.

Die Arbeit von Lim et al. wird in diesem Abschnitt skizziert, weil sie eine grundlegende Arbeit zur Vermittlung von Systemmechanismen zur Verstehbarkeit aus Sicht der HCI ist. In der Arbeit von Lim et al. werden Ansätze entwickelt, die im praktischen Teil der hier vorliegenden Arbeit von Relevanz sind (Kapitel 5). Den Grenzen bzw. Schwächen ihres Ansatzes, die in den nächsten Paragraphen besprochen werden, gilt es, in meinem Framework zu begegnen.

Die Grenzen ihres Ansatzes liegen zum einen in der konzeptuellen Schwäche, vor allem der Einschränkung auf Erklärungen als Vermittlungsinstanz. Zum anderen ist die Aussagekraft ihres empirischen Teils begrenzt, da, wie die Autoren selbst immer wieder betonen, es sich bei ihrer Studie *nur* um eine Laborstudie handelt, der ein vereinfachtes Testszenario zugrunde liegt. Das heißt, dass das System, mit dem die Teilnehmer der Studie interagierten, nur wenige einfache Fälle abbilden kann. In dem Abschnitt ihres Artikels „From the Lab to the Real World“ werden diese Einschränkungen behandelt und in diesem Zuge die Grenzen der Aussagekraft der Ergebnisse besprochen:

„Real context-aware applications are more complex and several issues would have to be handled regarding the provision of intelligibility type explanations. Firstly, applications that use decision tree models tend to have much larger trees learned from possibly hundreds of features, and it would not be scalable to generate explanations from them“.¹⁹

Reale maschinelle Algorithmen sind meist komplexer als in der oben genannten Studie dargestellt und folgen beispielsweise einer autoadaptiven Lernstrategie, die nicht den Anspruch auf Verstehbarkeit erhebt, sondern auf die Qualität der Vorhersage optimiert ist. Erklärungen, wie sie bei Lim et al. vorgestellt wurden, sind dafür nicht geeignet, da eine verständliche sprachliche Vermittlung, die dem oben genannten Schema folgt, bei komplexeren zugrundeliegenden Algorithmen schnell an ihre Grenzen kommt.

Das Framework, das Aspekte von Systemmechanismen versammelt, die für den Nutzer vermittelt werden können, wird in Abschnitt 4.3 entworfen. Das Framework soll sich für reale Systeme mit *State-of-the-Art*-Algorithmen eignen und geht damit über die Überlegungen in diesem Abschnitt hinaus. Einen Schritt in diese Richtung gehen die Autoren der Arbeiten, die im Folgenden diskutiert werden.

¹⁹Lim, Dey und Avrahami, 2009, S. 2127.

4.2.2 Verstehbarkeit im Zuge praktischer Erfahrung

Eine Weiterführung des ersten Ansatzes wurde von Todd Kulesza und Kollegen in die Diskussion eingebracht²⁰ Sie argumentieren, dass ein Verständnis eines intelligenten Systems vor allem in der Interaktion in einem iterativen Prozess entsteht. Ein weiterer Aspekt in ihrem Ansatz ist es, dass ein Verständnis des Systems durch eine aktive Teilhabe, das heißt, durch die Möglichkeiten, auf die Leistungen und Aktionsmöglichkeiten des Systems Einfluss zu nehmen, indem beispielsweise die Kategorien des Klassifizierungssystems angepasst werden bzw. verbessert werden.

Kaluesza et al. binden ihre Überlegungen an ein Konzept von mentalen Modellen. Dabei handelt es sich um explizite und implizite Vorstellungen über den adäquaten Umgang mit technischen Systemen. Damit ist mehr als ein explizites Wissen über die Funktionen gemeint, sondern auch ein praktisch implizites Aktionswissen.

„Mental models are internal representations that people build based on real world experiences. These models allow people to understand, explain, and predict phenomena, and to then act accordingly“.²¹

Auch hier spielen Erklärungen als eine Vermittlungsinstanz eine Rolle, es wird aber das Konzept von *in-situ explanations* verwendet, das heißt, die Erklärungen haben nicht die Struktur von ausführlichen Argumentationsketten, sondern sie werden, wie sie sagen, *häppcheweise* („bit by bit“) während der Interaktion vermittelt.

Sie präsentieren in ihrem Ansatz ein konzeptuelles Framework, das der Gestaltung hinsichtlich der Vermittlung von Systemmechanismen folgen soll. Diese werden als Prinzipien bezeichnet und im Folgenden vorgestellt:

- Prinzip 1: Be Iterative

Die Erklärungen zu den System-Mechanismen sollen während der Interaktion wiederholt in kleinen *Happen* präsentiert werden („easily consumable ‘bites’ of information“²²).

- Prinzip 2: Be Sound

Nach diesem Prinzip sollen die Erklärungen nicht auf wesentliche Weise weniger komplex sein, als sie in Wirklichkeit sind. „Sound“

²⁰Kulesza et al., 2013 und auch Kulesza et al., 2015, auf die ich mich hier vor allem beziehen werde.

²¹Kulesza et al., 2013, S. 3.

²²Kulesza et al., 2015, S. 127.

meint, dass die Erklärungen aussagekräftig und auch wahrheitsgemäß sind. „Soundness means that everything an explanation says is true“.²³

- Prinzip 3: Be Complete

„Be Complete“ zielt darauf ab, dass die gesamte Struktur des Systems vermittelt werden soll und nicht nur Teilbereiche. „A complete explanation does not omit important information about the model“.²⁴

- Prinzip 4: But Don't Overwhelm

Das vierte Prinzip schränkt den absoluten Anspruch der vorherigen Prinzipien ein: die drei Prinzipien sollen nur soweit eingehalten werden, dass die Nutzer im Umgang mit dem System durch zusätzliche Informationen nicht überfordern werden.

Kuleza et al. präsentieren zwei Studien, anhand derer sie die Prinzipien veranschaulichen. Eine davon, die sich mit einem intelligenten Sortierungssystem von E-Mails befasst, wird kurz rezipiert (Kulesza et al., 2015): Das Klassifikationssystem scannt den Inhalt und Betreff von eingehenden E-Mails und klassifiziert die E-Mails anhand dessen in bestimmte Kategorien. Beispielsweise wird eine E-Mail dem Themenbereich Baseball zugeordnet, weil der Name einer Baseballmannschaft und der Name eines Baseballwettbewerbs im Betreff vorkommt. Das zugrundeliegende maschinelle Lernsystem zur Sortierung von E-Mails ist ein *Naive Bayes Classifier*, wie er auch in der Abbildung 4.1 aufgeführt ist und sich durch eine gute Verstehbarkeit auszeichnet.

Um die oben genannten Prinzipien zu evaluieren, werden zusätzliche Informationen an die Nutzer vermittelt, um das zugrundeliegende Systemverhalten zu verstehen. Dazu werden neben den Ergebnissen der Sortierung auch die Features angezeigt, die für die Klassifizierung ausschlaggebend waren. Eine Prozentzahl zeigt zudem für jede sortierte E-Mail an, mit welcher *Überzeugung* (in diesem Zusammenhang wird von *confidence* gesprochen) das System die Sortierung vorgenommen hat. Diese Information wird mit einem Hinweis auf alternative Klassifizierungen, die aufgrund geringerer Wahrscheinlichkeit nicht berücksichtigt wurden, kombiniert.

Ihre Prinzipien sahen die Autoren wie folgt umgesetzt. Die Erklärungen waren „sound“, weil alle Features, die der Klassifikator nutzt, offengelegt wurden („by accurately disclosing all of the features the classifier used to make its prediction“), „complete“, weil immer auch auf alternative Möglichkeiten der Zuordnung verwiesen wurde und „iterative“, weil Nutzer

²³Kulesza et al., 2015, S. 127.

²⁴Kulesza et al., 2015, S. 128.

aktiv die Darstellung anpassen konnten und die zusätzlichen Informationen immer, wenn eine Klassifikation stattfand, präsentiert wurden. Dabei war das Design so gewählt, dass eine Überforderung der Nutzer nicht eintrat. Wenn die Anzahl der darzustellenden Features zu groß wurde, wurden die, die für die Klassifikation kaum eine Rolle spielten, weggelassen.

Bei der Durchführung der Studie wurde den Teilnehmern Klassifikationen der E-Mails und die Informationen über die Gründe präsentiert. Eine Kontrollgruppe bekam keine zusätzlichen Informationen. Das Verständnis der Nutzer über das System wurde in einem Fragebogen getestet.²⁵ Auch in dieser Studie gab es eine deutliche Steigerung des Verständnisses des Systems.

Als Ergänzungen zu den konzeptuellen Arbeiten, die den Herausforderungen eines Umgangs mit intelligenten Systemen begegnen wollen, wird im Folgenden eine Studie vorgestellt, die die Problemlage in einer qualitativen Studie unter *realen* Bedingungen zeigt und diskutiert.

4.2.3 *In the Wild*: Interaktion mit einem intelligenten Heizsystem

In diesem Teil der Rezeption des aktuellen Forschungsstand in der HCI wird eine Studie vorgestellt, die sich mit den Herausforderungen eines Umgangs mit alltäglichen intelligenten Systemen befasst.²⁶ In der Diskussion der Ergebnisse dieser Studie wird auch der Bedarf einer Vermittlung von Systemmechanismen für eine Verständlichkeit des Systems besprochen und auf die zuvor diskutierten Ansätze verwiesen.

Bei der Studie handelt es sich um eine Langzeitbeobachtung, die drei Wochen umfasste und in einem realen *Smart-Home*-Setting stattfand, in dem Sinne, wie Kulesza et al. von einem „real world“ Szenario sprechen. Die Studie wird von den Autoren als eine *wild study* charakterisieren (siehe Abschnitt 5.0.5). Die Studie veranschaulicht die Herausforderungen, die schon mit einem einfachen lernenden Systemen in einem alltäglichen *Smart-Home*-Kontext auftreten können.

In der Studie wurde der Umgang mit einem lernenden Thermostat, *Nest*, („first mass-market thermostat, feature machine learning“) untersucht, welches schon als kommerzielles Produkt in Privathaushalten im Gebrauch ist.²⁷ Im Rahmen der Studie wurde das Thermostat für drei Wochen in Privathaushalten installiert. Die Bewohner führten in dieser Zeit ein Tagebuch über ihre Erfahrungen mit dem Thermostat, zudem gab es mehrere Interviews der Wissenschaftler mit den Bewohnern.

²⁵Kulesza et al., 2015, S. 133.

²⁶Yang und Newman, 2013, S. 93.

²⁷Yang und Newman, 2013, S. 93.

Das intelligente Thermostat hat nach Herstellerangaben das Ziel, sich an das Verhalten (die Präferenzen) ihrer Nutzer anzupassen, um optimal heizen zu können. Dabei kann es die Umgebung mit einem Bewegungssensor erfassen und gegebenenfalls auf die An- und Abwesenheit der Bewohner schließen. Zudem registriert es die Temperaturvorgaben der Nutzer, aus denen es Präferenzmuster schließt. Nach einer Initialisierungszeit passt sich das System, so die Idee, automatisch an die Wünsche bzw. an die Gewohnheiten der Nutzer an. Die Basis für die Anpassung ist ein interner Algorithmus, der weder für den Nutzer noch für die Wissenschaftler der Studie zugänglich war. Der Anspruch des Algorithmus ist es, dass er die Temperatur so regelt, dass eine *Nachregulierung* des Nutzers in der Regel nicht nötig ist und gleichzeitig so viel Energie wie möglich gespart wird.

In der Studie wurde untersucht, wie Personen über einen längeren Zeitraum mit einem intelligenten System umgehen. Dazu wurde das Gerät in zehn Haushalten installiert und die Erfahrungen in Interviews und täglichen Tagebüchern dokumentiert.

Die Ergebnisse zeigten, dass die meisten Teilnehmer große Probleme mit dem System hatten, auch wenn sie größtenteils eine positive Grundhaltung mitbrachten. Hauptkritikpunkt waren die Irritationen, die immer wieder während der Nutzung auftraten und nicht oder nur schwer aufgelöst werden konnten: „Participants had trouble understanding how the Nest interpreted their input when creating a schedule and how the Nest sensed their movement or occupancy“.²⁸

Ein Teilnehmer beschreibt einen Fall, in dem eine Temperaturanpassung nur in einem spezifischen Fall vorgenommen wurde, vom System allerdings als Präferenz für jeden Tag aufgefasst wurde („had a pregnant daughter [visiting], and she doesn't like hot weather, so we turned it down for her“²⁹). Daraufhin zog es der Teilnehmer in Erwägung, das Gerät abzubauen: „It makes assumptions, and I don't like the assumptions, and I can't train it to make different assumptions“.³⁰ Ein anderer Kommentar zeigt, dass die Person Probleme hatte nachzuvollziehen, wann und wie das System feststellt, dass niemand mehr im Haus ist.

Der Umgang mit Irritationen erfolgte bei einem Teilnehmer durch das Zurücksetzen des gesamten Systems, damit der Lernprozess von neuem starten konnte („I erased the whole schedule and we started again“³¹). Um

²⁸Yang und Newman, 2013, S. 97.

²⁹Yang und Newman, 2013, S. 97.

³⁰Yang und Newman, 2013, S. 97.

³¹Yang und Newman, 2013, S. 98.

das Lernen danach in seinem Sinne zu gestalten, passte er sein eigenes Verhalten an, indem er Temperaturänderungen nicht mehr so frequentiert anpasste, wie es seinem üblichen Umgang mit einem Thermostat entsprach.

In der Diskussion der Ergebnisse schlagen die Autoren mögliche *Lösungen* vor, wie Systeme gestaltet werden sollen, dass es keine Missverständnisse über die Leistungen des Systems und ein Nachvollziehen des Systemverhaltens möglich ist. Dazu wird auf die in den Abschnitten zuvor besprochenen Ansätze zur Vermittlung von Informationen über das Systemverhalten für eine bessere Verständlichkeit verwiesen.

Die Studie zeigt und veranschaulicht einige Aspekte, die sich unter anderem mit den Überlegungen zum Verlust von Spuren und zum Irritationsmanagement decken.³² Das in der Studie untersuchte intelligente Thermostat ist eines, welches nur zwei Inputvariablen besitzt (Temperaturänderung und der Bewegungssensor). Es ist naheliegend, dass es in Zukunft komplexere Systeme geben wird. Doch schon bei diesem Produkt war es vielen der Teilnehmenden nur schwer oder gar nicht möglich, die Unklarheiten über die Anpassung oder das Nichterfüllen von Erwartungen, die sich im Umgang mit dem System einstellten, hinsichtlich ihrer Verursachung bzw. Gründe einzuordnen.

4.2.4 Fazit: Verstehbarkeit intelligenter Systeme

Die Auseinandersetzung mit den beiden konzeptuellen Ansätzen und der *Smart-Home*-Studie zeigte, dass es in der HCI-Forschung ein Bewusstsein dafür gibt, dass eine Unzugänglichkeit von Systemmechanismen intelligenter Systeme zu einer Verringerung der Nutzungsqualität bis hin zu einer Ablehnung führen. Ein Lösungsvorschlag besteht darin, Nutzern während des Umgangs Erklärungen über das aktuelle Systemverhalten anzubieten.

Die Autoren, die diesen Lösungsvorschlag erarbeiteten, konnten die Wirksamkeit ihrer Ansätze zeigen, aber nur in einem Laborsetting mit kontrollierten Randbedingungen und auf Basis *einfacher* lernender Algorithmen. Explizite Erklärungen des Systemverhaltens sind beispielsweise nicht mehr wirkungsvoll, wenn es sich um komplexere maschinelle Klassifikationssysteme handelt, da längere Beschreibungen nötig wären, um das Geschehene in Worte zu fassen. Kaluesza et al. weisen selbst darauf hin, indem sie die Grenzen ihres Konzepts mit Bezug auf „multiclass-classifier“- und „deep-learning“-Ansätzen hervorheben.³³ Da es das Ziel der Arbeit ist, eine gelungene Hybridisierung von Menschen durch intelligente Technik auf

³²Siehe Abschnitt 3.3.

³³Yang und Newman, 2013, S. 135.

Basis aktueller komplexer Lernsysteme zu ermöglichen, sind Erklärungen nicht der richtige Ansatz.

4.3 Kandidaten zur Vermittlung von Systemmechanismen

In den vorherigen Abschnitten wurde diskutiert, wie den Herausforderungen einer Nichtverstehbarkeit intelligenter Systeme durch ein besseres Systemdesign begegnet werden kann. Ein Lösungsvorschlag bestand darin, dem Nutzer während des Gebrauchs Erklärungen über das algorithmische Prozessieren anzubieten (4.2). Eine der aus den Diskussionsergebnissen gezogene Konsequenzen ist die, dass der *Erklärungsansatz*, der schon bei einfachen Lernalgorithmen an seine praktische Grenze kam, für aktuelle komplexe maschinelle Lernansätze, wie sie im ersten Teil des Kapitels (4.1) besprochen wurden, nicht geeignet ist.

Der Anspruch der besprochenen *Erklärungs-Ansätze* war es, ein weitgehend vollständiges Verständnis des Systemverhaltens beim Nutzer zu erzeugen. Von diesem Anspruch, so meine Auffassung, muss sich verabschiedet werden, vielmehr muss es Ziel sein, einzelne Aspekte des Mechanismus für den Nutzer in bestimmten Handlungssituationen darzustellen, um Schritt für Schritt eine Vorstellung über die Möglichkeiten und Grenzen der Technik zu entwickeln.

In diesem Abschnitt wird ein Framework vorgestellt, das diesem *neuen* Anspruch gerecht werden soll. Das Framework beinhaltet eine Typisierung von Elementen lernender Algorithmen, die sich je nach Anwendung dazu eignen können, den Nutzern im Umgang mit der Technik visualisiert zu werden. Die Aspekte werden mit Bezug auf die Gestaltungspraxis als Kandidaten bezeichnet und im Folgenden anhand von vier Gruppen typisiert. Je nach Anwendung kann es sinnvoll sein, nur einen Kandidatentypus zu verwenden oder eine Kombination. Die vier Kategorien orientieren sich dabei vor allem an den Überlegungen zum überwachten Lernen, wie sie in Abschnitt 4.1 diskutiert wurden.

Die folgenden Kategorien sortieren die Kandidaten zur Vermittlung von Systemmechanismen. Die Beispiele beziehen sich auf lernende Algorithmen, die dahingehend gelernt wurden, bestimmte Sensordaten anhand von festgelegten Klassen zu deuten. Man spricht von einer sensorbasierten maschinellen Klassifikation. Diese Zusammenstellung wird als Kandidaten-Framework bezeichnet. Das Kandidaten-Framework bildet die Grundlage für ein Systemdesign, das in den Studien im folgenden Kapitel (5) Verwendung findet.

- Schritte der maschinellen Klassifikation - Chronologie

- Quellen
- Metadaten
- Semantische Modelle: Domainwissen

4.3.1 Schritte der maschinellen Klassifikation - Chronologie

Der erste Kandidat betrifft die Schritte einer Klassifikation in ihrer Chronologie. Ein Klassifikator durchläuft in seinem Entscheidungsprozess meist eine Reihe von Schritten. Dieser Prozess kann dabei je nach dem zugrundeliegenden Ansatz systematisch sehr unterschiedlich aussehen. Mit Systematik ist vor allem gemeint, welches Paradigma der Optimierung als Basis dient, beispielsweise ein probabilistischer Ansatz oder ein Konzept neuronaler Netze. Dabei eignet sich nicht jeder Schritt für eine Vermittlung.

Die Klassifikation ist zu einem Ergebnis gekommen und hat auf dem Weg mehrere algorithmische Entscheidungsinstanzen (Stationen) durchlaufen. Diese Stationen können Gegenstand einer Visualisierung für den Nutzer werden, nicht allgemein, sondern individuell, abhängig vom Klassifikationsergebnis. Hinsichtlich eines Entscheidungsbaums können es die relevanten Knoten und die geprüften Merkmale sein (siehe 4.1.2). Es ist dabei wichtig, dass diese Information über das Durchlaufen von Entscheidungsinstanzen im maschinellen Deutungsprozess nicht für sich allein, das heißt, anhand einer Klassifikation für einen potentiellen Nutzer, der diese Schritte zu sehen bekommt, vollständig verstehbar sein muss, sondern wird als ein Indikator verstanden, der immer nur unvollständig auf die Funktionsweise des Gesamtsystems verweist.

Ein Entscheidungsbaum ist ein Beispiel für eine Klassifikation in mehreren Schritten, welche tendenziell in einem hohen Maße nachvollziehbar sein kann. Doch sind Entscheidungsbäume aufgrund sonstiger schlechter Leistungsfähigkeit für viele Anwendungen nicht geeignet (siehe die Bewertungen von Entscheidungsbäumen in der Tabelle 4.1). Wird stattdessen ein *Random-Forest*-Ansatz verwendet, bei dem es eine große Anzahl von parallel laufenden Bäumen gibt, deren Parameter der Instanziierung variieren.³⁴ Jeder Baum kommt zu einem individuellen Klassifikationsergebnis. Die Klassifikation mit den meisten *Votes* wird dann als Gesamtergebnis übernommen. Ein Zwischenschritt der Klassifikation ist demnach das Ergebnis des Votings der einzelnen Bäume, das auf den Systemmechanismus verweist und beispielsweise auch als Indikator für die Verlässlichkeit der Ergebnisse dienen kann.

³⁴Siehe Abschnitt 4.1.2.

Ein *Random-Forest*-Ansatz ist ein Beispiel für *ensemble methods*, bei denen mehrere Klassifikatoren, die parallel laufen, eingesetzt werden, um die Ergebnisse zu optimieren.³⁵ Dabei muss es sich nicht um die gleichen Typen von Lernalgorithmen mit unterschiedlicher Parametrisierung handeln, sondern es können auch mehrere ganz andere Ansätze gewählt werden. Bei dem *Dense-Motifs*-Ansatz, der im praktischen Teil dieser Arbeit verwendet wird, handelt es sich auch um eine *ensemble method*, bei der auch ein *Voting*-Zwischenschritt stattfindet (siehe 5.2.2).

4.3.2 Quellen

Ein zweiter Bereich von Kandidaten wird als Quelle bezeichnet. Gegenstand der Vermittlung ist der Ursprung dessen, auf dem die Klassifikation des Systems fußt. Je nach Anwendung können die Quellen eine Reihe von Sensorquellen sein sowie externes Wissen, das mit den Sensordaten kombiniert wird. Die Quellen könnte man auch als den ersten Schritt der Klassifikation bezeichnen, das heißt der Ausgangspunkt der Chronologie.

Die Quellen repräsentieren den Wahrnehmungsraum eines Systems und sind somit das Element, das die Möglichkeiten und Grenzen des Wirkens grundlegend festlegt. Die Quelldaten sind ein wichtiger Indikator für den Nutzer, seine eigene Erfahrung der Situation mit der des Systems abzugleichen. Die Möglichkeit eines zumindest teilweisen Abgleichs ist wichtig für das Akzeptieren vor Ergebnissen einer Klassifikation und den darauf basierenden Aktionen des Systems.

Bei vielen Systemen maschinellen Lernens im Bereich der HCI liegen unterschiedliche Datenquellen, das heißt Sensoren oder externe Informationen, zu Grunde, die zu bestimmten Entscheidungen (Klassifikationen) des Systems führen. So kann eine Darstellung nützlich sein, die anzeigt, welcher Sensor in einer bestimmten Anwendung, beispielsweise dem adaptiven Agieren des Systems, gerade aktiv ist oder aktiv war. Bei Smartphones mit einem Android Betriebssystem erscheint ein Symbol, wenn eine Anwendung auf GPS-Informationen zugreift. Das kann helfen, das Verhalten der Anwendung zu verstehen.

So ist beides vorstellbar, zum einen zu vermitteln, dass eine Informationsquelle gerade verwendet wird bzw. verwendet wurde, als die Entscheidung getroffen wurde, zum anderen den Inhalt der Informationsquelle selbst darzustellen.

³⁵Siehe unter anderem in der Dokumentation des scikit von python: <http://scikit-learn.org/stable/modules/ensemble.html>.

4.3.3 Metadaten

Die dritte Kategorie von Kandidaten wird unter der Bezeichnung Metadaten zusammengefasst. Hierbei wird eine Unterteilung in zwei Typen vorgenommen. Zum einen Metadaten, die das Training betreffen und zum anderen solche, die während der Laufzeit des Systems anfallen. Beim Training fallen beispielsweise Metadaten an, die die Qualität der Klassifizierung natürlich nur hinsichtlich der Trainingsdaten (Lern- und Testdaten) betreffen. Klassische Bewertungskonzepte von maschinellen Lernansätzen hinsichtlich ihrer Qualität sind unter anderem *precision* und *recall* oder bei autoadaptiven Lernsystemen die *convergence time*, das heißt, die Zeit bzw. die Anzahl der Iterationen, bis ein System zu einem ausreichend optimierten Zustand kommt. Die Qualität der Erkennung ist eine Art von Kontextinformation, die es ermöglicht, die Anwendung des Systems in verschiedenen Nutzungssituationen bezüglich der Verwendbarkeit einzuordnen. Metadaten, die während der Laufzeit des Systems anfallen, sind Daten der individuellen Nutzung im Sinne einer persönlichen Nutzungshistorie.

Im Folgenden wird eine Zusammenfassung der „performance measures for classification tasks“ präsentiert.³⁶ Übliche Mechanismen zur Bewertung der Genauigkeit im maschinellen Lernen sind, wie oben angesprochen, *precision* und *recall*, welche häufig als *F1 score* kombiniert werden. Unterschiedliche Performanzskalen sagen etwas anderes über die Leistungsfähigkeit des Systems in bestimmten Kontexten oder Aufgaben aus. Diese Performanzdaten sind Aspekte des Systemmechanismus, die für den Nutzer vermittelt werden können.

Bei der Bewertung der Performanz eines Systems berechnet man mit einer Formel einen Prozentsatz aus den richtig und falsch klassifizierten Daten. Hierfür gibt es eine vierfache Unterscheidung: Zum einen gibt es Samples, die einer Klasse richtig zugeordnet worden sind (*true positives*), zum anderen gibt es Samples, die richtigerweise der Klasse nicht zugeordnet wurden (*true negatives*). Zudem gibt es Samples, die einer Klasse zugeordnet wurden, aber nicht dazu gehören (*false positiv*) und schlussendlich Samples, die einer Klasse nicht zugeordnet worden sind, obwohl sie eigentlich dort hinein gehören (*false negatives*).

Beispielsweise ein System, das ein Foto dahingehend klassifiziert, ob sich ein Mann, eine Frau oder keins von beiden auf dem Bild befindet. Ein Bild mit einer Frau der Klasse *Frau* zuzuordnen, ist ein *true positiv*, einen Mann nicht der Klasse *Frau* zuzuordnen, ist ein *true negativ*. Einen Mann der Klasse *Frau* zuzuordnen, ist *false positiv* und eine Frau nicht der Klasse *Frau* zuzuordnen, ist *false negativ*.

³⁶Sokolova und Lapalme, 2009.

Die Bewertung von Samples hinsichtlich ihrer richtigen und falschen Zuordnung kann unterschiedlich gewichtet werden, so dass man zu anderen Genauigkeitswerten kommen kann. Ein Artikel von Sokolova und Lapalme vergleicht 24 Bewertungskonzepte. Hier sollen nur vier gängige vorgestellt werden und hinsichtlich ihrer Bedeutung verglichen werden.³⁷

Das Naheliegendste ist die sogenannte *accuracy*, bei der die Samples, die richtig klassifiziert wurden, und alle, die richtigerweise nicht als Teil der Klasse gesehen wurden, durch alle stattgefundenen Klassifikationen geteilt werden. Ein anderer wichtiger Messmodus ist *precision*, wobei hier nur die richtig klassifizierten Samples durch die Summe der richtigen und falschen Klassifikationen geteilt werden. Alle anderen Samples, die fälschlicherweise nicht klassifiziert worden sind, beeinflussen das Ergebnis nicht. Ist dieser Wert hoch, möchte man sicherstellen, dass alle Samples, die positiv sind, auch hier mit dem Kompromiss zugeordnet werden, dass auch einige Samples fälschlicherweise in diese Kategorie gelangen. Das könnte sinnvoll sein, wenn man einen Klassifikator hat, der Frauen erkennen soll und sichergehen möchte, dass man kein Bild mit einer Frau verpasst, auch wenn die Folge ist, dass es einige Fotos gibt, die fälschlicherweise als Frauen erkannt worden sind. Ein Algorithmus, der hinsichtlich des Qualitätsmaßes *Precision* optimiert wurde, soll dagegen nur wirklich das klassifizieren, bei dem er ein hohes Maß an Sicherheit besitzt. Das kann bei sicherheitsrelevanten Fragen besonders relevant sein, beispielsweise wenn die Klassifikation die Grundlage für eine automatisch Entscheidungen bildet, die, wenn sie falsch ist, großen Schaden anrichten kann. Eine weitere gängige Bewertungseinheit ist der *F1 score*, der eine Kombination aus den beiden Vorherigen ist, wobei der Fokus auf der Ausgeglichenheit der beiden Aspekte liegt.

Bewertungsskalen werden üblicherweise genutzt, um in der Autooptimierung die Ergebnisse zu bewerten und die Parameter anzupassen. Je nach Anwendung möchte man am Ende den einen oder anderen Wert möglichst hoch haben oder eine Ausgeglichenheit erreichen. Es sollte demnach auch klar sein, dass ein solcher Wert als eine Vermittlungsinformation für Nutzer dienen kann, um Entscheidungen und Deutungen des Systems zu kontextualisieren.

Es gibt auch noch andere Größen zur Messung der Performance. Dazu gehören beispielsweise die *convergence time*, das ist die Zeit oder besser gesagt die Menge der Durchläufe (*number of iterations*), bis sich beispielsweise ein neuronales Netz einstellt.

³⁷Sokolova und Lapalme, 2009.

4.3.4 Semantische Modelle: Domainwissen

Der vierte Typ von Kandidaten betrifft das Bereichswissen, welches bei der Implementierung eines Lernsystems häufig als Grundlage dient. Dieses kann, wenn es vorhanden und Grundlage des maschinellen Lernansatzes ist, Gegenstand einer Vermittlung werden.

Im maschinellen Lernen wird zwischen domain-abhängigen und domain-unabhängigen Ansätzen unterschieden.³⁸ Letzteres sind rein statistische Lernverfahren, bei denen kein Wissen über den Anwendungsbereich in Form von Modellen eingebracht wurde, sondern der Algorithmus ausschließlich die Beschaffenheit der Daten berücksichtigt. Der Vorteil solcher Ansätze ist es, dass sie relativ einfach von einem Anwendungsbereich in einen anderen übertragen werden können.

Mit Blick auf das grundlegende Schema des überwachten Lernens handelt es sich um einen domain-unabhängigen Ansatz, wenn die gelabelten Daten, das heißt, beispielsweise Sensordaten als Input (Features) und die Labels als gewünschte Outputs, dem System vorliegen und ausschließlich basierend auf diesen Daten der Algorithmus implementiert wird. Wenn zusätzlich zu den Trainingsdaten Informationen über das Verhältnis von Eingangsdaten und Ausgangsdaten vorliegen, unter anderem Gewichtungen der verschiedenen Inputs oder eine Hierarchie der Labels, dann ist die Lösung domain-abhängig.

Beispielsweise haben Berlin und Learhoven einen Lernalgorithmus zur Erkennung menschlicher Aktivität auf die Erkennung von Zugtypen, aufgrund ihrer charakteristischen Vibration beim Fahren auf einem Gleis, übertragen.³⁹ Die Trainingsdaten für den Anwendungsfall der Aktivitätserkennung sind üblicherweise die Daten eines Sensors, beispielsweise Daten eines Beschleunigungssensors, wie es das technische Setup für den praktischen Teil dieser Arbeit ist, zusammen mit den zugehörigen Labels, die besagen, in welchem Zeitraum eine Aktivität *wirklich* stattgefunden hat. Bestimmte Features in den Sensordaten dienen als Indikatoren dafür, ob eine Aktivität stattgefunden hat. Nur aus den Trainingsdaten ergibt sich die Assoziation zwischen Sensordaten und den zugehörigen Aktivitäten. Mit dem gleichen Ansatz, nur mit anderen Trainingsdaten, können die Sensordaten als Indikatoren für einen Zugtypen dienen, dessen Vibration von einem Beschleunigungssensor am Gleis, über den der Zug fährt, aufgezeichnet wurden.

³⁸Z.B. in dem Artikel von Ye et al. wird argumentiert, dass ein Domainwissen oder Expertenwissen, das in den Algorithmus eingebracht wird, die Qualität verbessert: „Learning-based approaches can benefit if the more accurate local domain knowledge applies“ (Ye, Dobson und McKeever, 2012).

³⁹Berlin und Laerhoven, 2012.

Bei den domain-abhängigen Ansätzen werden Modelle, das heißt, schon existierendes Wissen über den Anwendungsbereich mit eingebracht, um die Qualität zu verbessern. Grundlage für ein Bereichswissen können andere Wissenschaften oder das Domainwissen der Entwickler selbst sein, wobei man allgemein von einem *expert knowledge* spricht. Beispielsweise ist es naheliegend, Wissen zur Physiologie des menschlichen Körpers in eine Aktivitätserkennung einzubringen, um die Qualität der Deutung zu verbessern. Dabei wird beispielsweise das Wissen über das Verhältnis von Aktivitäten in einem bestimmten Kontext, beispielsweise dass bestimmte Subaktivitäten zu einer höherstufen Aktivität gehören für die Aktivitätserkennung genutzt.⁴⁰ So kann dieses Wissen über einen bestimmten Zusammenhang von bestimmten Subaktivitäten, die wiederum Teil einer höherstufigen Aktivität sind („complex activity“), die Erkennung verbessern.⁴¹

Ist ein solches *semantisches* Modell (im Gegensatz zu einem statistischen Modell) Grundlage eines Lernalgorithmus, kann es nützlich sein, Elemente dieses Modells an die Nutzer zu vermitteln.

⁴⁰Siehe beispielsweise bei Helaooui, Riboni und Stuckenschmidt, 2013, bei der eine *Ontologie* für allgemeine Haushaltstätigkeiten in die Aktivitätserkennung eingebracht wurde.

⁴¹Ein anderes prominentes Beispiel ist die maschinelle Spracherkennung, bei der grammatikalisches Wissen und Bedeutungswissen in die Erkennung mit eingebracht wird.

Kapitel 5

Konsequenzen für die Gestaltung von Self-Tracking-Techniken

In diesem Kapitel werden die bisherigen Überlegungen hinsichtlich ihrer Konsequenzen für eine technische Gestaltung intelligenter Self-Tracking-Techniken diskutiert.

Die beiden Studien sind im Anwendungsfeld der Self-Tracking-Techniken verortet. Bei den Systemen, die zum Zweck des Self-Trackings zum Einsatz kommen, handelt es sich um Techniken der sogenannten Aktivitätserkennung.¹ Beide Studien folgen jeweils eigenen *Subzielen*, die Fragestellungen im Forschungsbereich der HCI behandeln. Die Überlegungen zur Vermittlung der Systemmechanismen spielen aber im Untersuchungsdesign eine zentrale Rolle. Am Ende des Kapitels werden die theoretischen Konzepte zur Hybridisierung anhand der Ergebnisse der Studien reflektiert.

Die erste Studie (5.2) befasste sich als *Subziel* mit der Systemgestaltung zur möglichen Nutzung der Technik der Aktivitätserkennung in der psychologischen Therapie. Die zweite Studie (5.3) befasst sich als *Subziel* mit der Frage nach den Möglichkeiten und Grenzen einer Selbstreflexion mit Self-Tracking-Techniken. Die *Subziele* werden in den jeweiligen Einführungen zu den Studien besprochen, wobei in diesem Zusammenhang auch der aktuelle Forschungsstand (*Related Work*) vermittelt wird.

5.0.5 Methodische Ausrichtung

Die beiden Studien unterscheiden sich hinsichtlich ihrer Untersuchungsmethode. Bei der ersten handelt es sich um eine *lab-based* Studie und bei der zweiten um eine *wild-study*. Beide Ansätze sind gängige Methoden im

¹Eine medienphilosophische Betrachtung der Technik der Aktivitätserkennung ist hier zu finden: Dietrich und Laerhoven, 2016a.

Forschungsbereich der HCI mit ihren Vor- und Nachteilen. Studien in einem Laborsetting (*lab-based*) entsprechen dem klassischen experimentellen Design. Die experimentelle Umgebung ist kontrolliert, so dass die Bedingungen für alle Teilnehmenden so weit wie möglich gleich sind und nur einzelne zu untersuchende Größen variiert werden. Die *wild-studies* sind unter dem Einfluss der Ethnomethodologie entstanden. Die Motivation für diesen alternativen Ansatz lag in der Künstlichkeit einer Laborumgebung, die, so die gängige Kritik, wenig Aussagekraft bezüglich des realen Einsatzes von computertechnischen Systemen in alltäglichen Anwendungen hätte. Demnach versuchen die *wild-studies*, die Versuchsanordnung vom Labor in alltägliche Nutzungsumgebungen zu verlegen. Ein Beispiel für ein solches Studiendesign ist eine Untersuchung der Nutzung eines technischen Systems über mehrere Wochen im Haushalt einer Familie.² Bei den *wild-studies* werden keine detaillierten Hypothesen aufgestellt, die dahingehend untersucht werden, ob sie zutreffen oder nicht. Vielmehr sind es unerwartete Ergebnisse, die bewusst angestrebt werden.

„lab-based research often do not map onto the messy human-computer interactions in the real world. People are much more unpredictable—for example, they get distracted and are constantly interrupted or interrupt their own activities by talking to others, taking breaks, starting new activities, resuming others, and so on“.³

Selbstverständlich können auch in einer klassischen Versuchsanordnung unerwartete Ergebnisse auftreten, die die Angemessenheit der Hypothese beispielsweise gänzlich in Frage stellt. Bei *wild-studies* sind die Hypothesen von Anfang an vage, so dass von den Untersuchungsergebnissen nicht erwartet wird, dass sie die Hypothese als wahr oder falsch herausstellen, sondern immer nur mögliche Erklärungen und Hinweise und so Ideen zum Weiterdenken liefern. Der Nachteil liegt aber auf der Hand; die erschwerte Reproduzierbarkeit der Ergebnisse und die begrenzten Möglichkeiten der Verallgemeinerung.

In der ersten Studie war eine Laborstudie geeignet, weil die Datenbasis, anhand derer die Teilnehmenden verschiedene Visualisierungen von Aktivitätsdaten betrachteten, schon vorlag und damit für alle gleich sein konnte. In der zweiten Studie wurde ein *wild*-Ansatz gewählt, weil es für die Untersuchung der Möglichkeiten und Grenzen der Selbstreflexion wichtig war, dass die Teilnehmenden sehr individuelle Erfahrungen auf Basis ihres eigenen Verhaltens machten.

²Ein Beispiel für eine solche Langzeitstudie in einem Privathaushalt: Houben et al., 2016.

³Rogers, 2011, S. 62.

Bei den beiden Studien handelt es sich um Anwendungen im Feld des sogenannten *long-term activity recognition*. Das technische Setup besteht aus einem Sensorsystem mit dem Namen *Hedgehog*, das die Grundlage für die Deutung der Aktivitäten bildete. Im folgenden Abschnitt wird das technische Setup ausführlich beschrieben, bevor im Anschluss die erste Studie präsentiert wird.

In den Studien wird im Gegensatz zum Rest der Arbeit eine andere Sprachregelung gewählt, die die weibliche Form als die festlegt, die sowohl für die männlichen als auch die weiblichen Teilnehmenden gilt.

5.1 Technisches Setup

In diesem Abschnitt werden das Sensorsystem und der maschinelle Lernansatz vorgestellt, die die technische Grundlage für beide Studien bilden. Teil dieser Vorstellung ist eine kurze Einführung in die Technik der Aktivitätserkennung mit Beschleunigungssensoren, die an die grundlegende Besprechung der Ansätze maschinellen Lernens aus Abschnitt 4.1 anknüpft.

5.1.1 Das Sensorsystem: Der *Hedgehog*

Das Sensorsystem, das in meinen Studien Verwendung findet, basiert auf einem schon existierenden System mit dem Namen *Hedgehog*, welches maßgeblich in der *Embedded Sensing Systems Group* an der Technischen Universität Darmstadt entwickelt wurde. Das Sensorsystem befindet sich in einem Kunststoffgehäuse und kann mit einem Armband am Handgelenk befestigt werden. Siehe dazu Abbildung 5.1.

Das Sensorsystem besteht aus einem Beschleunigungssensor (ADXL345), der Beschleunigungen in drei Achsen und in einer Frequenz von bis zu 2 kHz aufzeichnen kann. Die Beschleunigungsdaten werden mit einem *low-power microcontroller* (Microchip 18F46j50) verarbeitet, der im Wesentlichen die Aufgabe hat, die Sensordaten zu komprimieren und anschließend auf einer micro-SD-Karte, die sich im Gerät befindet, abzuspeichern. Das Sensorsystem kann über einen mini-USB-Anschluss mit einem Computer verbunden werden, um die Daten herunterzuladen. Der USB-Anschluss dient zudem dazu, den internen 180mAh Akku aufzuladen.

Der *Hedgehog* ist als Bauplan für das Gehäuse, als Schaltplan für die Platinen und der Software zum Aufzeichnen und Visualisieren der Daten für die Forschungscommunity auf *Github* veröffentlicht worden.⁴ Das Anwendungsgebiet, für den das Gerät entwickelt wurde, liegt im Bereich des

⁴<http://kristofvl.github.io/HedgeHog/>.



ABBILDUNG 5.1: Das Sensorsystem *Hedgehog*, das wie eine Armbanduhr am Handgelenk getragen werden kann. Es kann Beschleunigungsdaten mit einer Frequenz von bis zu 2 kHz aufzeichnen.

long-term activity trackings. Aufgrund dessen ist das Sensorsystem auf eine lange Laufzeit von bis zu 2 Wochen optimiert, ohne dass ein Aufladen nötig ist. In den Studien ist eine kontinuierliche Aufzeichnung der Sensordaten von einer Woche nötig.

In beiden Studien wird das Sensorsystem am Handgelenk getragen. Die physischen Bewegungen des Arms dienen als Indikatoren für eine Vielzahl von alltäglichen Aktivitäten wie Sport- und Freizeitaktivitäten. Die Vorstellung der Grundlagen der Technik der Aktivitätserkennung auf Basis von Beschleunigungssensoren erfolgt im folgenden Abschnitt.

5.1.2 Die Technik der Aktivitätserkennung auf Basis von Beschleunigungssensoren

Der Beschleunigungssensor, der im Sensorsystem verwendet wird, kann in jeder Achse bis zu 2000 Datenpunkte pro Sekunde (2 kHz) aufzeichnen, wobei in den Studien eine geringere Aufzeichnungsfrequenz (100 Hz) verwendet wurde.

Der Beschleunigungssensor gehört zu den sogenannten inertialen Sensoren. Wenn mehrere inertielle Sensoren in einer Messeinheit zusammenwirken, spricht man von IMUs (*inertial measurement units*), wozu neben dem Beschleunigungssensor meist ein Gyroskop und seltener ein Magnetometer gehört. IMUs sind oft schon als eine Sensoreinheit günstig verfügbar, so dass im Unterschied zu dem hier vorliegenden System die ganze Einheit Verwendung findet und nicht nur die Beschleunigungssensoren alleine.

Die Technik der Aktivitätserkennung mit Beschleunigungssensoren wird klassisch wie folgt charakterisiert:

„We formulate activity recognition as a classification problem where classes correspond to activities and a test data instance is a set of acceleration values collected over a time interval“..⁵

Was im Zitat als *classification problem* bezeichnet wird, bezieht sich auf die Anwendung überwachten Lernens für die Deutung von Aktivität (siehe Abschnitt 4.1). Ein zentrales Element der Deutung von Aktivität ist die automatische Selektion von *Features* in den Rohdaten, wobei es sich um einen hochfrequenten und diskreten Datenstream (*time-series*) handelt. Die *Features* bilden die Grundlage für die Implementierung eines Klassifikators.

Viele wissenschaftliche Arbeiten haben sich mit der Auswahl von *Features* in diskreten Datenstreams befasst. Das Isolieren von *Features* geht meist mit einem *preprocessing* einher, welches das Ziel hat, die Datenmenge zu verringern, ohne dass die charakteristischen Eigenschaften des vollständigen Datenstreams verloren gehen. Nur durch das *preprocessing* der Daten ist es in der Folge überhaupt möglich, dass ein Klassifikator auf die Daten angewandt wird und dieser in vertretbarer Zeit ein Klassifikationsergebnis liefert.⁶

Für das *preprocessing* eignet sich unter anderem der *sliding window* Ansatz. Ziel dieses Ansatzes ist es, Datenpunkte innerhalb eines Fensters zusammenzufassen, wobei der Begriff *sliding* besagt, dass das Fenster in Bewegung ist. Jedes Fenster, das eine Reihe von aufeinanderfolgenden Datenpunkten zusammenfasst, endet nicht da, wo das Fenster zuvor aufgehört hat, sondern es findet immer eine Überschneidung statt. So zeigt sich beispielsweise mit Blick auf eine Arbeit von Ravi et al., dass sich ein „window size of 256 with 128 samples overlapping between consecutive windows“ bewährt hat und zwar unter der Voraussetzung, dass eine Messrate von 50 Hz gewählt wurde.⁷

Die Bewertung der zusammengefassten Datenpunkte in einem Fenster erfolgt aufgrund statistischer Größen wie Mittelwert und Varianz. Die Ergebnisse dieser Bewertung bilden im engeren Sinne den Feature-Raum. Hinsichtlich des Verständnisses dieser Bewertungseinheiten für die Deutung von Aktivität kann man sich vorstellen, dass es Aktivitäten gibt, die

⁵Ravi et al., 2005, S. 1543.

⁶Eine Übersicht zu den Verfahren des *preprocessing* bzw. der Featureselektion in der Technik der Aktivitätserkennung ist in der Dissertation von Berlin, 2014 zu finden.

⁷Ravi et al., 2005, S. 1542.

sich durch spezifische Mittelwerte, Varianzen und Frequenzen der Bewegungen (beispielsweise eines Armes) auszeichnen.⁸

Wie zuvor angesprochen spielt das *preprocessing* und die Auswahl der Features eine wichtige Rolle für die Performance der Erkennung. Die Klassifikation erfolgt anhand der Eigenschaften des Feature-Raumes. Sehr viele der klassischen Konzepte kommen auch hier erfolgreich zur Anwendung. Dazu gehören Naïve Bayes, Bayesian Networks, Hidden Markov Models (HMMs) oder Support Vector Machines (SVMs).⁹

Die Auswahl des geeigneten Klassifikators hängt unter anderem von der Art der Aktivitäten ab und auch von der Qualität der Trainingsdaten. In den Studien spielt die Deutung von Aktivitäten, die in einer längeren Zeitspanne stattfinden, eine zentrale Rolle, das heißt Freizeit- und Sportaktivitäten sowie grundlegende Aktivitäten der Fortbewegung (Laufen, Radfahren, Bus- und Bahnfahren). Eine Liste von Aktivitäten, die in der ersten Studie Teil der Aktivitätserkennung sind, befindet sich in Tabelle 5.1.

Der Ansatz maschinellen Lernens, der in der ersten Studie Verwendung findet und im Speziellen für längere Aktivitäten geeignet ist, ist die „dense motif discovery“.¹⁰ Es ist ein Ansatz, der in der Aktivitätserkennung weniger präsent ist und ursprünglich aus der DNA-Analyse stammt, sich aber für den hier vorliegenden Anwendungsfall der Klassifikation von alltäglichen Aktivitäten gut eignet.

5.2 Studie 1: Intelligentes Self-Tracking-System zur möglichen Anwendung in der psychologischen Therapie

In der ersten Studie wird ein intelligentes Self-Tracking-System implementiert und die Möglichkeiten und Grenzen der Vermittlung von Systemmechanismen untersucht. Bei der Vermittlung der Aktivitätsinformation wird eine Strategie verfolgt, die sich an dem Framework der Kandidaten aus dem vorherigen Kapitel orientiert (4.3). Als maschineller Lernansatz, der die Klassifikation von Aktivitäten übernimmt, wird die sogenannte *dense motif discovery* verwendet.

⁸Huynh und Schiele, 2005, Lester et al., 2005.

⁹Philipose et al., 2004, Oliver und Horvitz, 2005, Stiefmeier et al., 2008, Holleczeck et al., 2010, Partridge und Begole, 2011, Amft, Junker und Troster, 2005, Chang, Chen und Canny, 2007, Lester et al., 2005, Mahdavian und Choudhury, 2008, Donald J. Patterson und Philipose, 2005.

¹⁰Berlin und Van Laerhoven, 2012.

Die Gestaltung der Studie orientiert sich an einem potentiellen Anwendungsszenario für die Technik der Aktivitätserkennung, das im Folgenden vorgestellt wird.

5.2.1 Szenario: Aktivitätserkennung für die psychologische Therapie

Die Studie befasst sich mit der Aktivitätserkennung in ihrer möglichen Anwendung in der psychologischen Therapie. Informationen über alltägliche Aktivitäten als Routinen und Gewohnheiten dienen in der Therapie häufig als Indikatoren für Psychiater/Psychologen, um psychische Erkrankungen zu diagnostizieren und zu behandeln. Berlin und Laerhoven haben eine Reihe an Kriterien aufgestellt, nach denen ein System zur Therapieunterstützung gestaltet werden könnte.¹¹

Demnach ist es bisher gängige Praxis, dass Patienten *self-reports* über bestimmte Aktivitäten in dem Zeitraum zwischen zwei Therapiesitzungen anfertigen, die dem Therapeuten in den Gesprächen als Grundlage dienen können. Eine automatische Deutung der Aktivität mit Hilfe eines Sensorsystems könnte als eine weitere Informationsquelle dienen, die ein Bild über den Zustand des Patienten vermittelt. Dabei ist nicht der Umstand wichtig, dass ein umfassendes Aktivitätsprofil erstellt wird, sondern dass Patient und Therapeut bestimmte *Schlüsselaktivitäten* auswählen, die von besonderer Aussagekraft bezüglich der aktuellen Verfassung des Patienten sind. Die Regelmäßigkeit hinsichtlich der Durchführung einer bestimmten Sportaktivität kann viel darüber aussagen, in welcher psychischen Verfassung sich eine Person, beispielsweise mit einer Bipolaren Störung befindet.¹²

Bei den *Schlüsselaktivitäten* handelt es sich vor allem um Sport- und Freizeitaktivitäten. Sind die Aktivitäten bekannt, kann das System gezielt auf diese Aktivitäten trainiert werden. Eine solche gezielte Auswahl macht es möglich das System hinsichtlich der *Erkennung* zu optimieren und verringert das Gefühl umfassender Überwachung. In einer zukünftigen Implementierung könnten die *self-reports* als Labels für die Trainingsdaten dienen, die ohnehin Teil der gängigen Therapiepraxis sind.

Wie weiter oben erwähnt, handelt es sich bei der Studie um keine Untersuchung, die zeigen soll, ob sich die Technik der Aktivitätserkennung

¹¹Berlin und Van Laerhoven, 2012, S. 251.

¹²Personen mit Bipolarer Störung haben phasenweise depressive Episoden, die sich mit manischen Episoden, die sich meist durch ein Stimmungshoch und erhöhte Aktivität auszeichnen, abwechseln. „Bipolar disorder, also known as manic-depressive illness, is a brain disorder that causes unusual shifts in mood, energy, activity levels, and the ability to carry out day-to-day tasks“ (von der Homepage des „National Institute of Mental Health“ aus den USA: <https://www.nimh.nih.gov/health/topics/bipolar-disorder/index.shtml>).

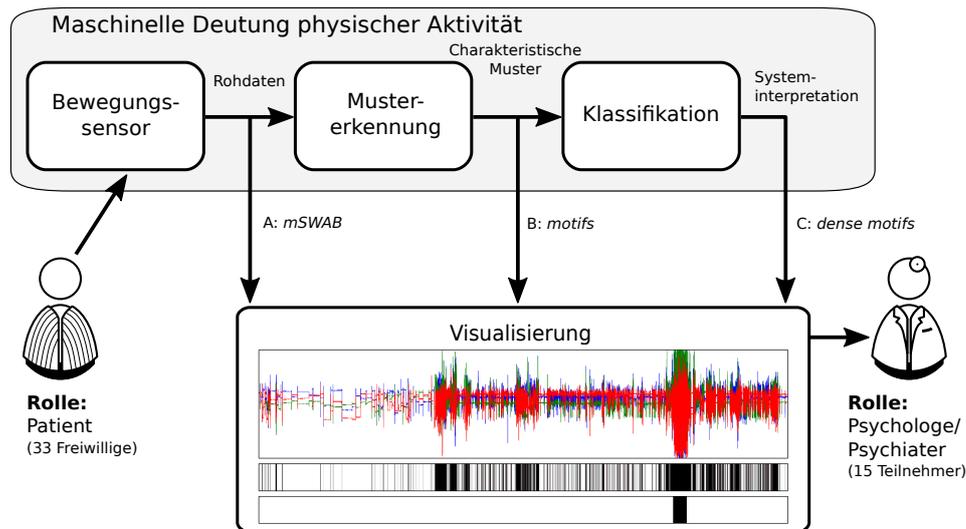


ABBILDUNG 5.2: Übersicht über die grundlegende Struktur der Studie

für eine psychologische Therapie eignet. Eine solche Frage seriös zu beantworten wäre nur möglich, wenn eine umfassende Kooperation mit Psychologen/Psychiatern eingegangen würde, was hier nicht der Fall war. Untersucht wurden vielmehr nur die Möglichkeiten der Darstellung von Aktivitätsinformationen auf Basis der maschinellen Deutung und wie Personen mit dem Dargestellten umgehen können. Das ist ein Aspekt der, so die Annahme, zentral für eine mögliche Nutzung in der psychologischen Therapie gesehen wird. Mit Bezug auf dieses Anwendungsszenario ist es wichtig für das Studiendesign, dass die Teilnehmenden der Studie keine Experten in der Analyse solcher Daten sind.

In der Studie wird untersucht, inwiefern Personen (in der Rolle von Psychiatern/Psychologen) die Aktivitäten anderer Personen (Patientendatensatz) anhand verschiedener Typen von Visualisierungen interpretieren können. Dabei werden neben der Ergebnisdarstellung der Aktivitätserkennung auch bestimmte Zwischenschritte der Erkennung vermittelt. Eine Übersicht hinsichtlich der Gestaltung der Studie ist in Abbildung 5.2 dargestellt. Das in der Studie verwendete Datenset (Patientendaten), welches 33 einwöchige Aufzeichnung von Bewegungsdaten beinhaltet, wird in Abschnitt 5.2.4 (Studiendesign) detailliert vorgestellt.

5.2.2 Lernansatz: *Dense Motif Discovery*

Der maschinelle Lernansatz mit dem Namen *dense motif discovery* folgt in seinem Prozess der Deutung von Beschleunigungsdaten als Aktivitäten drei wesentlichen Abstraktionsschritten. Der erste Schritt ist das sogenannte *preprocessing*, bei dem die Menge der Rohdaten reduziert werden, ohne

dass die Gestalt des Graphen verloren geht. In einem zweiten Schritt werden die so vorbereiteten Rohdaten nach charakteristischen Mustern (diese werden auch als *Motifs* bezeichnet) durchsucht. Charakteristische Muster sind typische Substrings der Aktivität, die dem System im Training beigebracht wurden. Im letzten Schritt findet die eigentliche Klassifikation statt, in dem untersucht wird, ob in einem bestimmten Zeitabschnitt die Dichte (*Density*) von charakteristischen Mustern über einem bestimmten Grenzwert liegt. Ist dies der Fall, wird angenommen, dass die Aktivität in diesem Abschnitt stattgefunden hat. Eine detaillierte Beschreibung der Schritte findet im Folgenden statt.

Mit Blick auf die Typen von Kandidaten zur Vermittlung von Systemmechanismen (4.3), bieten sich die Ergebnisse der Zwischenschritte, wobei der erste Zwischenschritt auch als Quelle gesehen werden kann, und das Ergebnis der Klassifikation als Grundlage für eine Visualisierung für den Nutzer an. Demnach gibt es drei Typen von Visualisierungen (siehe Abbildung 5.2).

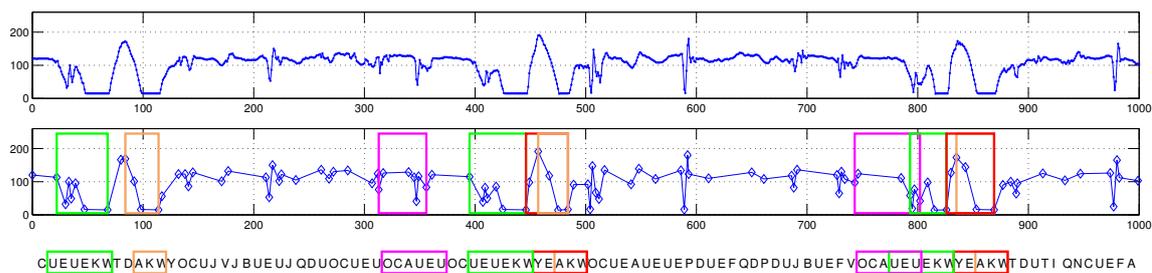


ABBILDUNG 5.3: Die Rohdaten des Beschleunigungssensors (oberer Graph) wurden durch eine schrittweise Linearisierung (*sliding window*) in Segmente transformiert (mittlerer Graph), wobei die Gestalt bewahrt wurde. Die Segmente wurden dann abschnittsweise in diskrete Symbole abstrahiert (Buchstaben, unten)

Vorverarbeitete Rohdaten (*mSWAB*)

Das Sensorsystem zeichnet die Beschleunigungsdaten mit 100 Hz auf, was eine große Menge an Daten zur Folge hat. Da eine so große Datenmenge als Eingangsgröße für die folgenden Schritte zu umfangreich ist, müssen die Daten vorverarbeitet werden. Die Herausforderung einer Vorverarbeitung besteht darin, die Menge der Daten zu verringern, aber dabei die Gestalt des Datengraphen und damit den Informationsgehalt möglichst zu bewahren. Für eine gestaltbewahrende Vorverarbeitung bietet sich der *sliding-window*-Ansatz an, wie er zuvor in Abschnitt 5.1.2 beschrieben wurde. Dieser Ansatz wurde in der Studie mit einem *Bottom-Up*-Algorithmus

kombiniert, was zusammen unter der Bezeichnung *mSWAB* zu finden ist.¹³ Das Ergebnis dieses *preprocessing* bietet die Grundlage für den ersten Typen von Visualisierungen, der als Rohdatengraph bezeichnet wird. Die Bezeichnung Rohdaten ist, wie gezeigt wurde, nicht ganz richtig, da ja eine Bearbeitung stattfand. Das ist aber insofern gerechtfertigt, weil sich die Daten erst durch die Verringerung der Datenmenge im Bereich des sinnvoll Darstellbaren befinden und weil die Gestalt des Graphen im Rahmen der Möglichkeiten bewahrt wurde.

Erkennung charakteristischer Muster *Motifs*

Nach der Vorverarbeitung der Rohdaten werden die so entstandenen reduzierten Datenpunkte symbolischen Repräsentationen zugeordnet. Bestimmte Ketten von symbolischen Repräsentationen, als *Words* bezeichnet, bilden die charakteristischen Muster. Die Zuordnung zu den symbolischen Mustern basiert auf dem Winkel der Steigung zwischen einem Datenpunkt zum nächsten. Die Symbole der Repräsentation sind Buchstaben (eine zugehörige Tabelle strukturiert, welche Steigung welchen Buchstaben hat), die in ihrer Reihenfolge bestimmte Wörter bilden und damit die charakteristischen Muster repräsentieren. Abbildung 5.3 zeigt, wie aus der Zeitreihe der Rohdaten durch eine lineare Segmentierung eine Vereinfachung vorgenommen wird und wie diese schlussendlich zu der symbolischen Repräsentation in Form von Buchstaben kommt, die sich aus den Steigungen zwischen den Datenpunkten ergeben.

Die charakteristischen Muster bilden die Grundlage für den zweiten Typus von Visualisierungen. Ein schwarzer Balken zeigt, ob an dieser Stelle ein charakteristisches Muster für eine gesuchte Aktivität gefunden wurde.

Klassifikation: *Dense Motifs Discovery*

In der Trainingsphase lernt der Algorithmus, welche Ketten von Symbolen (Muster) charakteristisch für eine Aktivität sind, das heißt am besten hinsichtlich ihrer Unterscheidbarkeit zu anderen Aktivitäten ist. Ein *bag of words classifier* durchsucht die symbolische Repräsentation in ihrer Chronologie nach bestimmten Wörtern. Ist die Dichte von charakteristischen Mustern als Wörter in einem Fenster hoch genug – auch hier wurde der *sliding-window*-Ansatz mit einer Fenstergröße von 10 Minuten verwendet – wird die entsprechende Aktivität diesem Zeitfenster zugeordnet.

¹³Van Laerhoven, Berlin und Schiele, 2009.

5.2.3 Visualisierung der Zwischenschritte

Die drei Abstraktionsschritte (oder Zwischenschritte) bieten die Grundlage für die Typen von Visualisierungen. Bei den drei Schritten handelt es sich um die Rohdaten als Quelle (die vorverarbeitet wurde, aber unter Einhaltung der grundlegenden Gestalt), den charakteristischen Mustern als Zwischenschritte der Erkennung sowie um die Klassifikation als Ergebnisse der maschinellen Deutung.

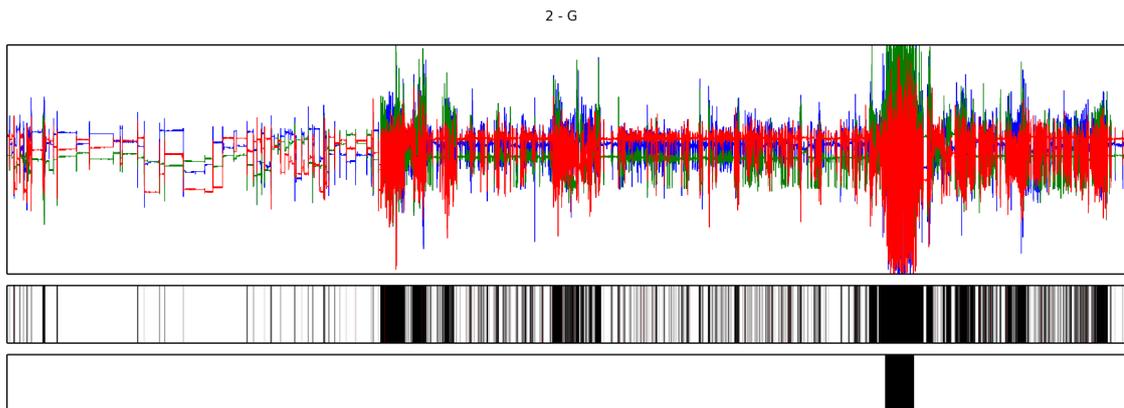


ABBILDUNG 5.4: Die drei grundlegenden Typen von Visualisierungen, die jeweils alleine und in Kombinationen den Studienteilnehmern gezeigt wurden: (oben, A) *vorverarbeitete* Rohdaten des dreiachsigen Beschleunigungssensors, (Mitte, B) Auftreten von charakteristischen Mustern der Aktivität, (unten, C) maschinelle Klassifizierung der Aktivität im schwarz dargestellten Zeitabschnitt. Das Beispiel in der Abbildung zeigt das Beispiel Badminton in einem Zeitfenster von ca. 1,5 Stunden

Für die Studie wurde eine Darstellungsweise gewählt, die auf wesentliche Aspekte reduziert ist. Zum einen wurde auf eine Achsenbeschriftung verzichtet und die Daten in einer 24-Stunden-Ansicht vereinheitlicht. Die drei Typen von Visualisierungen sind in Abbildung 5.4 dargestellt, das heißt die Rohdaten, die sich oben befinden, in der Mitte die charakteristischen Muster und unten die Ergebnisse. In der Studie werden die Typen von Visualisierungen in der gleichen Reihenfolge mit den Buchstaben A, B und C bezeichnet. Die Typen können in der Studie alleine oder in Kombination gezeigt werden, wobei sie bei letzterem mit einem + verbunden werden (beispielsweise A+C, wenn die Rohdaten und die Ergebnisse der Klassifikation gezeigt werden).

5.2.4 Übersicht: Studiendesign

Das Szenario in der Studie orientiert sich an einer möglichen Verwendung der Self-Tracking-Technik in der psychologischen Therapie. Dabei wird angenommen, dass Psychiater/Psychologen keine Erfahrung mit der Technik der Aktivitätserkennung und der Interpretation von komplexen Zeitreihen haben. Basierend auf dieser Annahme wurden die Teilnehmenden so ausgewählt, dass sie diesen Voraussetzungen möglichst entsprachen und weder einen technischen noch einen naturwissenschaftlichen Hintergrund hatten. Den Teilnehmenden der Studie, insgesamt 15 Personen, wurde einer der Typen von Visualisierung oder eine Kombination vorgelegt und sie mussten einschätzen, wo im Graphen eine ihnen genannte Aktivität stattfand.

Die Daten, die die Grundlage für die Visualisierungen bildeten, stammten aus dem sogenannten Patientendatensatz. Dieser Datensatz basierte auf den Daten von 33 Freiwilligen, die keine Überschneidung mit den 15 Teilnehmenden der Studie hatten und entstand zum Teil im Zusammenhang mit einer anderen Studie, die von Berlin und Laerhoven angefertigt wurde.¹⁴ Der Patientendatensatz wurde erzeugt, indem die 33 Freiwilligen eine Freizeittätigkeit, die sie regelmäßig ausübten (*Schlüsselaktivität*), mehrmals in einer Woche durchführten. In dieser Woche trugen sie das Sensorsystem 24 Stunden am Tag. Ein Aktivitätsprotokoll, in das die Freiwilligen Start und Endpunkt ihrer *Schlüsselaktivität* eintrugen, diente als Label für die Sensordaten. Die Idee der Erkennung einer einzelnen *Schlüsselaktivität*, die die Personen regelmäßig ausübten, deckte sich mit den Annahmen in dem Szenario eines Einsatzes in der Therapie, wie es zuvor beschrieben wurde.

Tabelle 5.1 zeigt eine Übersicht des Patientendatensatzes, indem neben Alter, Geschlecht und Art der Aktivität auch die durchschnittliche Qualität angegeben wird, mit der die Aktivitäten mit dem zuvor vorgestellten Lernansatz gedeutet wurden.¹⁵ Die Qualität wird dabei in Prozent hinsichtlich drei gängiger Messungsmethoden der Performance von Klassifikatoren angegeben. Eine Erklärung dieser Methoden wurde im Abschnitt 4.3 vorgestellt. Insgesamt hatte das System eine durchschnittliche Erkennungsrate von 76 %.

¹⁴Berlin und Van Laerhoven, 2012.

¹⁵Ein Teil des Patientendatensatzes diente immer als Trainingsdatensatz und der andere als Testdatensatz.

TABELLE 5.1: Für jeden Freiwilligen im Patientendatensatz wurde eine *Schlüsselaktivität* ausgewählt, die eine Woche lang einmal am Tag durchgeführt wurde. Die Qualität der Aktivitätserkennung mit dem *Dense-Mmotifs*-Ansatz ist in drei gängigen Bewertungsmethoden *precision*, *recall*, und *F1-score* (in %) angegeben. Die Labels der Aktivitäten im Datensatz wurden einheitlich in englischer Sprache festgelegt und in der Studie auch so genutzt

Teilnehm.	Geschlecht	Alter	<i>Schlüsselaktivität</i>	<i>precision</i>	<i>recall</i>	<i>F1</i>
1	männlich	30	badminton	95.2	90.0	92.5
2	männlich	32	badminton	94.4	89.6	91.9
3	männlich	31	basketball	96.2	92.4	94.3
4	weiblich	26	canoeing	82.5	76.5	79.4
5	männlich	32	cooking	38.9	42.0	40.4
6	männlich	35	cycling	87.8	89.0	88.4
7	männlich	30	dancing	92.2	92.0	92.1
8	weiblich	14	dancing	84.6	86.4	85.5
9	weiblich	16	dancing	44.9	63.5	52.6
10	männlich	20	drums	91.5	97.6	94.4
11	männlich	31	fishing	62.6	77.0	69.1
12	männlich	53	fishing	47.0	81.1	59.5
13	weiblich	26	flamenco	62.7	61.1	61.9
14	männlich	27	guitar	83.7	79.8	81.7
15	weiblich	27	guitar	94.2	91.0	92.6
16	männlich	23	guitar	77.6	73.3	75.3
17	männlich	28	gym	61.6	67.6	64.5
18	männlich	32	gym	86.1	77.5	81.6
19	männlich	30	gym	59.6	62.7	61.1
20	weiblich	28	gym	76.4	52.2	62.0
21	männlich	31	ironing	93.4	84.5	88.7
22	weiblich	27	keyboard	91.7	85.6	88.6
23	weiblich	28	knitting	58.8	73.7	65.4
24	männlich	30	lunch	26.7	30.1	28.3
25	männlich	25	soccer	98.1	93.2	95.6
26	weiblich	25	squash	92.0	74.1	82.1
27	männlich	27	squash	90.4	77.4	83.4
28	männlich	29	streetdance	66.4	65.8	66.1
29	weiblich	30	streetdance	58.4	69.1	63.3
30	männlich	32	washing car	81.3	79.9	80.6
31	weiblich	28	xbox	95.5	96.2	95.9
32	weiblich	28	yoga	66.6	43.2	52.4
33	weiblich	30	zumba	96.8	97.6	97.2
Durchschnitt				76.8	76.1	76.0

Demographie und Methodologie

Insgesamt 15 Personen (10 Frauen, 5 Männer) nahmen an der Studie teil. Der größte Teil der Teilnehmenden wurde an der Universität und ausschließlich an nicht-technischen Fakultäten rekrutiert. Das Alter lag zwischen 23 und 57 Jahren, im Durchschnitt bei 32 Jahren.

Die Studie lief wie folgt ab: Jedem Teilnehmenden wurde wiederholt ein einzelner Typ von Visualisierung oder eine Kombination gezeigt und die Teilnehmenden mussten anhand dessen einschätzen, wo in dem Dargestellten eine ihnen genannte Aktivität zu finden ist. Die Visualisierungen

wurden auf Basis des Patientendatensatzes erzeugt.

Der Ablauf, bei dem die Typen von Visualisierungen oder eine Kombination gezeigt wurden, verlief in Phasen. Die erste Phase beinhaltete nur einzelne Typen von Visualisierungen, die zweite eine Kombination von zwei Typen und in der dritten Phase die Kombination von allen drei Typen. Innerhalb der Phase war die Reihenfolge der Typen zufällig ausgewählt, wobei im Laufe des Versuchs jeder Typ von Visualisierung und jede Kombination den Teilnehmenden jeweils zweimal gezeigt wurde. Insgesamt musste jede Teilnehmende 14 Visualisierungen bearbeiten.

Zur Vorbereitung der Studie wurde jedem Teilnehmenden erklärt, worum es sich bei dem Dargestellten handelt. Bei der Erklärung handelte es sich nicht um eine ausführliche Erläuterung des Lernansatzes, wie sie in den vorherigen Absätzen durchgeführt wurde, sondern hauptsächlich um die Benennung der Visualisierung als Rohdaten als charakteristische Muster und Klassifikationsergebnisse. Zudem wurden sie informiert, dass die durchschnittliche Erkennungsrate bei 76 % liegt.

Die zentrale Aufgabe für die Teilnehmenden bestand darin, einzuschätzen, wo in den ihnen vorliegenden Visualisierungen eine bestimmte Aktivität zu finden ist. Der Name der Aktivität und die ungefähre Dauer der Präsentation wurde den Teilnehmenden genannt, wobei die Dauer zwischen einer halben Stunde und vier Stunden lag, im Durchschnitt ca. 1,5 Stunden. Dazu wurden die Visualisierungen ausgedruckt und die Teilnehmenden hatten die Aufgabe, direkt einzuzichnen, wo sie die Aktivität vermuteten. Die Einschätzungen der Teilnehmenden wurden am Ende mit den Labels der Trainingsdaten verglichen. Wenn das eingezeichnete Zeitfenster mit dem Zeitfenster aus den Trainingsdaten eine Überschneidung von 50 % hatte, wurde angenommen, dass die Einschätzung richtig war, falls nicht, wurde sie als falsch gewertet.¹⁶ Nachdem die Teilnehmenden eine Aktivität eingezeichnet hatten, wurden sie gefragt, wie sicher sie mit ihrer Einschätzung sind und für wie intensiv sie die eingezeichnete Aktivität einschätzen. Diese Bewertung erfolgte anhand einer Skala von *sehr sicher/sehr intensiv* bis *sehr unsicher/wenig intensiv* in sieben Abstufungen, einer sogenannten *7-point Likert-scale*.

Nachdem die Teilnehmenden alle 14 Visualisierungen bearbeitet hatten, wurde eine Reihe von allgemeinen Fragen zur Bewertung der Typen von Visualisierungen gestellt. Es wurde gefragt, wie nützlich sie die Typen von Visualisierungen und ihre Kombinationen fanden. Die Frage nach der Nützlichkeit bezog sich auf die Möglichkeit, etwas über die Aktivitäten der Personen zu erfahren, die die Daten repräsentierten. Zudem wurde

¹⁶Dieser großzügige Grenzwert wurde verwendet, da auf den ausgedruckten Visualisierungen keine sehr präzisen Einzeichnungen möglich waren.

gefragt, wie sehr sie den Typen bzw. den Kombinationen vertrauten. Für beide Fragen mussten die Teilnehmenden die Visualisierungen anhand einer Rangfolge von 1 bis 3 – vom Besten zum Drittbesten – bewerten. Den Abschluss machte eine freie Frage, in der die Teilnehmenden einschätzen sollten, welche Arten von Aktivitäten ihrer Meinung nach von ihnen am besten erkannt worden sind. Alle Fragen wurden sehr allgemein gestellt, das heißt, es wurden sowohl Nützlichkeit als auch Vertrauen nicht weiter charakterisiert. Damit ist anzunehmen, dass teilweise ein unterschiedliches Verständnis der Begriffe bei der Beantwortung vorgeherrscht hat. Auf eine weitere Ausdifferenzierung der Begriffe wurde verzichtet, da diese die Komplexität erhöht hätte.

5.2.5 Ergebnisse der Studie

Die Ergebnisse der Studie sind zum einen die quantitative Auswertung der eingezeichneten Schätzungen der Teilnehmenden, deren Selbsteinschätzung hinsichtlich der Qualität ihrer Einzeichnungen und das abschließende Ranking der Typen von Visualisierungen und ihrer Kombinationen. Zum anderen die qualitativen Ergebnisse, die sich aus den Aussagen der Teilnehmenden ergab, die während der Durchführung gemacht wurden. Eine weitere qualitative Quelle waren die freien Antworten am Ende des Fragebogens.

Die Ergebnisse sind in den folgenden fünf Stichpunkten zusammengefasst und basieren auf den quantitativen Auswertungen, die in den Abbildungen 5.5 bis 5.8 dargestellt sind.

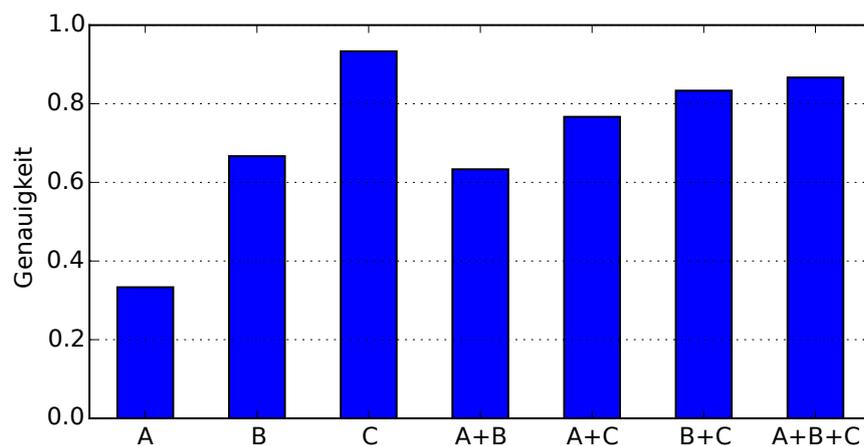


ABBILDUNG 5.5: Die Genauigkeit (*accuracy*), mit der die Teilnehmenden die Aktivitäten schätzten und einzeichneten. Eine vergleichende Darstellung hinsichtlich der Typen von Visualisierungen (A-C) und deren Kombinationen

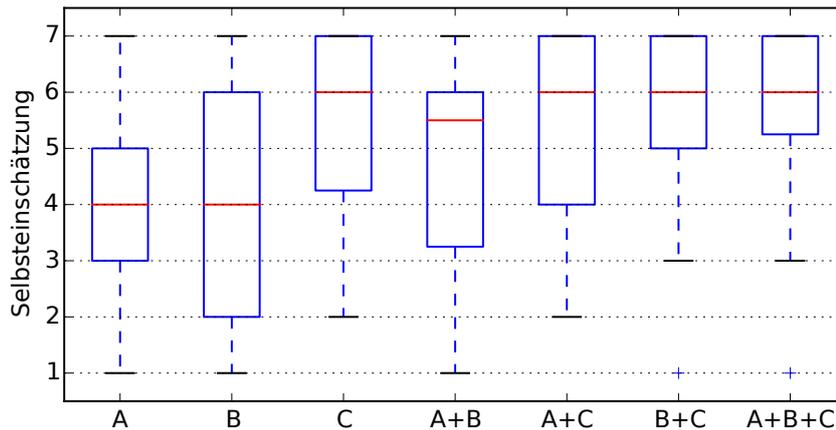


ABBILDUNG 5.6: Ein Vergleich der Selbsteinschätzungen der Teilnehmenden hinsichtlich ihrer Einzeichnungen. Die Selbsteinschätzung erfolgte anhand einer 7-point Likert-scale

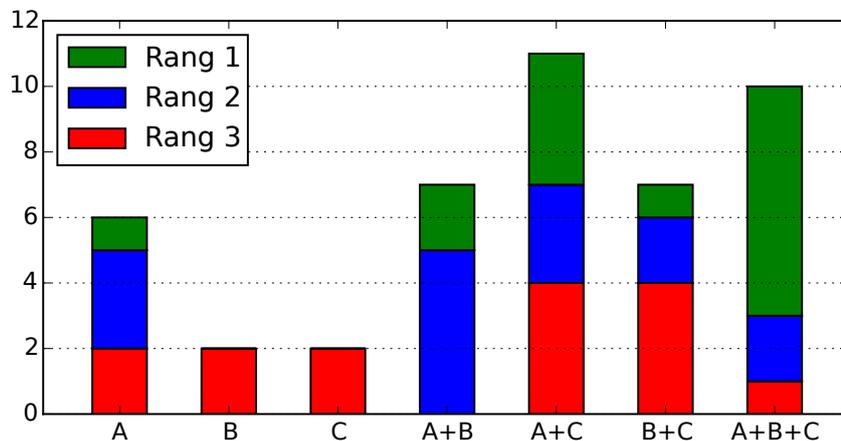


ABBILDUNG 5.7: Die Top-3 Visualisierungen, die von den Teilnehmenden als besonders nützlich benannt worden sind, um die Aktivitäten zu interpretieren

- **Die Visualisierung auf Basis der Systemklassifikation bietet die beste Schätzgrundlage:**

Die Auswertung der eingezeichneten Schätzungen, die von den Teilnehmenden vorgenommen wurden, zeigt, dass die Visualisierung bei der es sich ausschließlich um die Klassifikation des Systems (C) handelt, am besten abgeschnitten hat. Abbildung 5.5 zeigt das Ergebnis dieser Auswertung. Auf dem zweiten Platz liegt die Kombination von allen drei Typen von Visualisierungen (A+B+C), gefolgt von der Kombination der Systemergebnisse zusammen mit den charakteristischen Mustern (B+C). Alle Einschätzungen, in denen die Rohdaten (A) eine Rolle spielten, haben relativ zu den anderen Ergebnissen schlecht abgeschnitten.

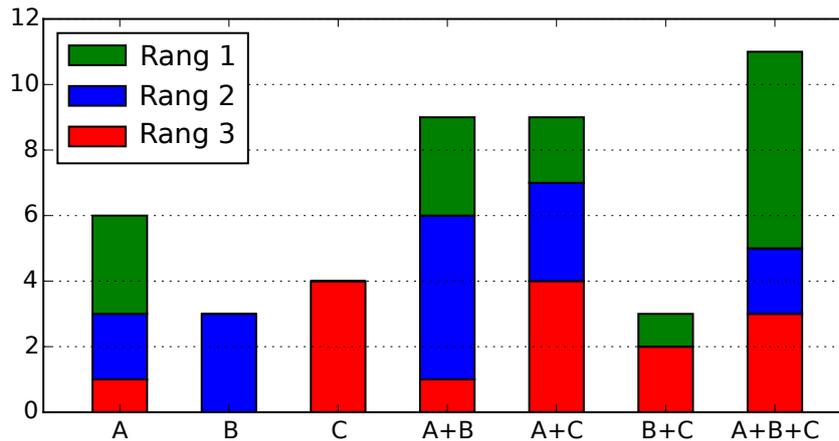


ABBILDUNG 5.8: Die Top-3 Visualisierungen, die von den Teilnehmenden als besonders vertrauenswert herausgestellt worden sind

- **Die Selbsteinschätzung der Teilnehmenden stimmte größtenteils mit den Ergebnissen des Einzeichnens überein:**

Die Beantwortung der Frage, wie sicher sich die Teilnehmenden mit ihrer Einzeichnung sind, ergab einen ähnlichen Trend. Abbildung 5.6 zeigt die Ergebnisse der Auswertung der *7-point Likert-scale*. Die Selbsteinschätzung war in dem Sinne gut, da auch hier die drei Topplatzierten mit denen der Auswertung der Einzeichnung übereinstimmten. Ein bemerkenswerter Unterschied ist, dass die Reihenfolge in der *Top 3* dagegen eine andere ist und die Kombination von allen drei Visualisierungstypen (A+B+C) am besten abschneidet. Auch sind die Visualisierungen, die den Typen Rohdaten (A) beinhalten, besser eingeschätzt worden als sie tatsächlich waren. Aufgrund der geringen Teilnehmerzahl ist der Unterschied hier aber nicht so groß, so dass keine Signifikanz festzustellen ist.

- **Die Teilnehmenden favorisieren Rohdaten und charakteristische Muster:**

Nach den Durchläufen, in denen die Teilnehmenden die Aktivitäten einzeichnen mussten, wurden sie anhand eines Fragebogens gebeten, ihre drei favorisierten Typen von Visualisierungen beziehungsweise Kombinationen in Form eines Rankings auszuwählen. Dieses Ranking betraf zum einen die Frage, welche Visualisierungen sie am nützlichsten fanden und zum anderen, wie sie das Dargestellte hinsichtlich des Vertrauens bewerten. Die Abbildungen 5.7 und 5.8 zeigen das Ergebnis in einem *gestapelten Balkendiagramm (stacked bar plot)*. Das Ergebnis kann wie folgt zusammengefasst werden: Die Rohdaten zusammen mit den Klassifikationen des Systems (A+C) und die Kombination aus allen drei Visualisierungstypen (A+B+C) wurden als am

nützlichsten eingeschätzt, gefolgt von den Rohdaten zusammen mit den charakteristischen Mustern (A+B). Es ist auffallend, dass das, was im Ranking als am nützlichsten genannt wurde, den vorherigen Ergebnissen entgegenstand, das heißt, in Relation zu der Qualität der eingezeichneten Aktivitäten und wie die Teilnehmenden die Qualität des Eingezeichneten selbst bewertet haben. Vor allem, da die Systemklassifikationen als am wenigsten nützlich bewertet wurden, wo sie doch die beste Qualität bei den eingezeichneten Schätzungen produziert hatten. Eine weitere wichtige Beobachtung ist es, dass alle Kombinationen, die Rohdaten beinhalten (A, A+B, A+B+C), sowohl bei der Nützlichkeit als auch beim Vertrauen gute Ergebnisse erzielten, vor allem relativ zu den Ergebnissen der vorangegangenen Auswertungen.

- **Intensive Aktivitäten sind aus Sicht der Teilnehmenden am charakteristischsten:**

Auf die Frage, welche Arten von Aktivitäten nach ihrer Ansicht von den Teilnehmenden am besten erkannt wurden, nannten der größte Teil *sport activities* (6).¹⁷ Zudem nannten die Teilnehmenden, dass sie Aktivitäten mit großer Intensität (4) oder schnellen Bewegungen (2) gut erkennen konnten. Nur zwei der Teilnehmenden sagten, dass Aktivitäten, die keine durchgehende Bewegung (wie das Bügeln) oder ganz unterschiedliche Richtungen beinhalteten, am besten zu erkennen waren. Eine Person sagte, dass sie nicht darüber nachgedacht habe, wie die Bewegungen der Aktivitäten aussehen könnten, sondern dass sie nur auf das Vorhandensein von erkennbaren Mustern (Patterns) geachtet hätte.

5.2.6 Diskussion der Ergebnisse

Die Studie beinhaltet keine tiefgehende Auseinandersetzung mit den Herausforderungen einer psychologischen Therapie. Das Studiendesign bietet aber die Möglichkeit, sich mit den Möglichkeiten und Grenzen der Vermittlung von Aspekten der Systemmechanismen hinsichtlich eines möglichst realistischen Anwendungsszenarios zu befassen.

Das signifikanteste Ergebnis der Studie war, dass es eine große Differenz gab zwischen den eingezeichneten Schätzergebnissen und dem, was die Teilnehmenden am Ende als nützlich bewerteten. Das konnte man unter anderem bei Visualisierungen sehen, die den Typus Rohdaten beinhalteten, deren zusätzliche Präsenz keine Verbesserung in den Einschätzungen

¹⁷Die Zahl in Klammern zeigt die Anzahl der Personen, die eine solche Aussage getroffen haben.

brachte, wann eine Aktivität stattgefunden hat. Vielmehr war das Gegenteil der Fall und die Ergebnisse wurden sogar schlechter. Gleichzeitig besagten die Ergebnisse des Fragebogens am Ende der Studie, dass die Teilnehmenden die zusätzlichen Informationen nützlich fanden und vor allem mehr Vertrauen in das System hatten.

Im ersten Moment scheint es widersprüchlich zu sein, dass sich die Qualität der Einschätzungen verringert, wenn eine zusätzliche Informationsquelle bereitgestellt wurde. Doch eine mögliche Erklärung liegt auf der Hand: Die Möglichkeit, überhaupt erst Zugang zu einer Quelle zu haben, die eine alternative Lesart zulässt, motiviert dazu, diese Möglichkeit in manchen Fällen auch zu realisieren. Natürlich war es in manchen Fällen wirklich besser, die systemische Klassifikation zu überstimmen und die eigene Einschätzung an den anderen Informationsquellen zu orientieren. Wenn dagegen nur die Klassifikationen des Systems vorlagen und man keinen anderen Anhaltspunkt hatte, musste die Teilnehmende diesen glauben oder völlig zufällig tippen. Da letzteres irrational gewesen wäre, folgten die Personen in solchen Fällen den Systemeinschätzungen, was im Durchschnitt dann die beste Lösung war.

Dies könnte ein Argument dafür sein, gar keine zusätzlichen Informationen zu visualisieren. Doch zeigten die guten generellen Bewertungen der Visualisierungstypen, die Aspekte der Systemmechanismen darstellten, dass die Teilnehmenden einen Wert darin sahen. Dieser Bedarf wurde in der Studie nicht durch weitere Fragen elaboriert, so dass nur Vermutungen angestellt werden können. Ich nehme an, dass der Vertrauensaspekt hierbei wichtig ist, der vor allem mit dem Vorhandensein von Rohdaten zusammenhängt. Wie sich gezeigt hat, waren die Rohdaten alleine keine gute Grundlage für eine Schätzung, wie die Auswertung der eingezeichneten Schätzungen zeigte. Gleichzeitig wurden die Rohdaten in ihrer Gestalt als durchgängiger Datenstrom, der zumindest ein Indikator für mehr oder weniger Aktivität war oder in bestimmten Abschnitten Regelmäßigkeiten aufwies, als etwas aufgefasst, das die Klassifikationen des Systems untermauern könnte.

Hier wäre es ohne Frage interessant zu sehen, wie sich ein längeres Arbeiten mit den Visualisierungen auswirken könnte. Beispielsweise, indem auch Lerneffekte untersucht würden. In dieser Studie wurde versucht, Lerneffekte möglichst zu verhindern, was dadurch realisiert wurde, dass die Visualisierungen in zufälliger Reihenfolge gezeigt wurden, um damit allen Visualisierungstypen gleichermaßen gerecht zu werden. Mit Blick auf das beschriebene Szenario wird angenommen, dass zusätzliche Informationen wie die Rohdaten dabei helfen können, Vertrauen aufzubauen. Die Visualisierung der charakteristischen Muster kann beispielsweise dazu dienen,

zu bewerten, wie sicher das System ist, dass eine bestimmte Aktivität in einem Zeitraum stattgefunden hat. Zudem ist die Intensität der Aktivitäten beispielsweise nur in den Rohdaten zu erkennen - das ist eine weitere qualitative Information, die beispielsweise in der Therapie nützlich sein kann.

5.2.7 Fazit: Erste Studie

Es hat sich gezeigt, dass es zielführend sein kann, Nutzern Visualisierungen anzubieten, die Aspekte der Systemmechanismen vermitteln. Im Interesse der zentralen Fragestellung der Arbeit war es wichtig, eine Vermittlung von Systemmechanismen an dem Kandidaten-Framework zu orientieren und sich dabei auf einen realistischen *State-of-the-Art*-Algorithmus zu beziehen. Teil des realistischen Ansatzes war es auch, dass keine Vereinfachungen für die Darstellung vorgenommen wurden, sondern nur Aspekte visualisiert wurden, die Teil des realen Prozesses der maschinellen Deutung sind.¹⁸

Bezüglich der Anforderungen, die sich aus den Überlegungen zur Hybridisierung ergaben, lässt sich folgendes sagen: Es wurden Bedingungen geschaffen, in denen das potentielle Lesen von Spuren möglich ist. Die Einschätzungen (Klassifikationen) des Systems sind nicht einfach gegeben, sondern es besteht die Möglichkeit, dass sie im Lichte der Rohdaten und der charakteristischen Muster als Zwischenschritt der Erkennung erscheinen. Ich argumentiere, dass die zusätzlichen Informationen dazu führten, dass die Teilnehmenden ein partielles Verständnis über die Möglichkeiten und Grenzen des Systems entwickelt hatten. In diesem Sinne war eine individuelle mediale Erfahrung möglich, die in ihrer Folge einem Prozess der Hybridisierung zuträglich sein kann.

5.3 Studie 2: Self-Tracking-Techniken und die Möglichkeiten und Grenzen einer Selbstreflexion

In der zweiten Studie geht es darum zu untersuchen, inwiefern ein intelligentes Self-Tracking-System dazu geeignet ist, das Personen auf Basis dessen ihr eigenes Verhalten reflektieren können. Die Möglichkeiten und Grenzen einer Selbstreflexion mit Hilfe von Self-Tracking-Techniken ist das *Subziel* dieser Studie. Wie Selbstreflexion in diesem Zusammenhang zu verstehen ist, wird im nächsten Abschnitt diskutiert. Dabei wird die Reflexion hinsichtlich der Rezeption in der HCI-Forschung behandelt und nicht in einem grundlegenden philosophischen Verständnis.

¹⁸Entspricht den Prinzipien *sound* und *complete*, die von Kulueza et al. hinsichtlich einer Vermittlung von Systemmechanismen formuliert wurden und die in Abschnitt 4.2.2 diskutiert wurden.

Die Untersuchung der Möglichkeiten und Grenzen der Selbstreflexion wird in Abhängigkeit von der Darstellung von Systemmechanismen durchgeführt. Dazu werden zwei Arten von Visualisierungen von Aktivitätsdaten hinsichtlich des Aspekts der Selbstreflexion verglichen, wobei es sich bei dem einen Visualisierungstypen um einen handelt, der den Vorgaben des Kandidaten-Frameworks folgt.

5.3.1 Das Ziel ist die Reflexion: Eine Diskussion der Konzepte

Der Zweck, für den die Nutzung von Self-Tracking-Techniken als Mittel dienen soll, ist es, folgt man der gängigen Auffassung in der technischen Forschung, dass eine Selbstreflexion ermöglicht wird. In Abschnitt 2.3.2 wurden Überlegungen vorgestellt, die in den Bereichen *Personal Informatics* sowie *Quantified Self* verortet sind. Unter anderem in einem Artikel zum Thema *Personal Informatics* sprechen die Autoren davon, dass es das Ziel ihres Artikels ist, Entwicklern ein besseres Verständnis zu vermitteln, wie Techniken zur Selbstreflexion genutzt werden können: „developers and designers can better take advantage of Ubicomp technologies to help users self-reflect“.¹⁹

An dieser Stelle werden die Überlegungen aus dem Abschnitt 2.3.2 erweitert und sich intensiver mit den Konzepten von Selbstreflexion hinsichtlich der Technik des Self-Tracking befasst.²⁰ Ausgangspunkt für die Diskussion ist der weiter oben genannte Artikel von Li et al., indem der Reflexion in ihrem „stage based model“ der *Personal Informatics* eine zentrale Rolle zukommt. Dieses Model besteht aus einer Reihe von Phasen, die den Prozess beschreiben, wie Nutzer Techniken des Ubiquitous Computings im Rahmen eines Self-Trackings verwenden. Folgt man ihrem Phasenmodell, beginnt der Prozess beim Auswählen der relevanten Daten, gefolgt vom Sammeln der Daten sowie der Kombination der gesammelten Daten. In der abschließenden „reflection stage“ sind die so aggregierten Daten Gegenstand einer bewussten Auseinandersetzung: „looking at lists of collected personal information or exploring or interacting with information visualizations“.²¹ Das Modell ist Ergebnis einer Reihe von Interviews mit Personen, die unterschiedliche Techniken nutzen, um personenbezogene Daten aufzuzeichnen und auswerten zu können. In dem Artikel ist die Reflexion bzw. Selbstreflexion darüber hinaus nicht weiter konzeptualisiert.

¹⁹Li, Dey und Forlizzi, 2011, S. 405.

²⁰Eine philosophisch-handlungstheoretische Betrachtung von Self-Tracking-Techniken ist nicht das Interesse dieses Kapitels, wurde aber auch hinsichtlich der Technik der Aktivitätserkennung in Dietrich und Laerhoven, 2016b diskutiert.

²¹Li, Dey und Forlizzi, 2010, S. 562.

Eine ausführliche Betrachtung von Reflexion ist von Eric P. S. Baumer in die Diskussion in der Forschung der HCI eingebracht worden.²² Baumer stellt in seinem Framework drei *Dimensionen* bzw. Verständnisse von Reflexion vor, die er als *breakdown*, *inquiry* und *transformation* bezeichnet. Hierfür verweist er auf eine ganze Reihe von Theorien aus unterschiedlichen Disziplinen, unter anderem aus Psychologie, Philosophie und Pädagogik. Baumers Ansatz wird vor allem mit Bezug auf seine Konzeption und nicht auf die theoretischen Grundlagen, auf die er sich bezieht, vorgestellt. Sein Ziel ist es, Methoden und Strategien für ein technisches Design von Self-Tracking-Techniken zu entwickeln, die auf einem tieferen Verständnis von Reflexion beruhen.

Den ersten Typus von Reflexion bezeichnet er als *breakdown*. Der Begriff *breakdown* bezeichnet eine Störung von Abläufen, die zu einem (mentalen) Zusammenbruch führt, der mit Bezug auf die Reflexion nicht nur eine negative Konnotation hat, sondern einen Beginn markiert, den Anfang einer Neuordnung. Diese Art der Reflexion ist dadurch charakterisiert, dass durch einen explizierenden Akt etwas auffällig wird und in deren Folge eine implizite Konzeption der Welt fraglich wird.

Er exemplifiziert *breakdown* mit Bezug auf Forschungskonzepte aus dem Feld der Ethnologie. Ein klassischer Ansatz der Ethnologie ist es, dass Personen einer sozialen Gemeinschaft dazu aufgefordert werden, die Strukturen und Regeln ihres sozialen Zusammenlebens einem Fremden (plausibel) zu erklären („tacit, implicit social structuring of interaction must be explicated by its interactors“²³). Das Explizieren ist ein Akt, bei dem Selbstverständliches fraglich werden kann. Hinsichtlich der Self-Tracking-Techniken, die alltägliche Routinen über einen längeren Zeitraum aufzeichnen und deuten, könnte es dazu kommen, dass implizites Verhalten explizit und damit die bisherigen Bewertungen dieses Verhaltens als beispielsweise gesund/ungesund oder moralisch gut/schlecht fraglich werden. Was Baumer nicht als Teil einer solchen Reflexion sieht, ist die Folge eines solchen Auffällig-werdens bzw. die Bedingungen, unter denen diese Irritation in Richtung eines produktiven Ergebnisses im Sinne einer Erkenntnis oder einer Verhaltensänderung kanalisiert werden können.

Der zweite Typus von Reflexion wird als *inquiry* bezeichnet, was als Nachforschen übersetzt werden kann. Während das Ergebnis des ersten Typus ein unerwartetes Ereignis ist, wenn auch meist Ergebnis eines bewussten Handelns, wird bei der *inquiry* der Prozess des bewussten Suchens

²²Baumer, 2015.

²³Baumer, 2015, S. 590.

nach Zusammenhängen in den persönlichen Daten als Reflexion bezeichnet. Das Prozedere des Nachforschens charakterisiert er als planvolles Vorgehen, das sich analog zu der Methode des wissenschaftlichen Arbeitens durch das Aufstellen von Hypothesen und das Testen dieser und daran anschließend der Anpassung der Hypothesen auszeichnet. Wiederum auf den Anwendungsfall der Aktivitätserkennung bezogen kann es Ziel der Nutzer sein, ihr Verhalten in bestimmter Hinsicht ändern zu wollen, indem sie etwas über die Kontexte von ungewolltem, weil beispielsweise ungesundem Verhalten, erfahren wollen. Ein Beispiel für einen solchen Kontext könnte eine *gesellige Runde* sein, die mit Rauchen und dem Konsum von fettigem Essen einhergeht. Selbstverständlich ist es möglich, dass auch bei einem gezielten Nachforschen das Ergebnis überraschen kann und sich beispielsweise die Hypothese als falsch gestellte Frage herausstellt und völlig andere Faktoren hervortreten.

Eine dritte Dimension von Reflexion ist die *transformation*. Reflexion wird demnach als ein geistiger Prozess betrachtet, der Beobachtungen nicht als Indikatoren für Wirkliches deutet, sondern als Basis für die Vorstellung von Möglichem. Zentral ist demnach nicht das, was ist, sondern das, was sein könnte, eine Variation der aktuellen Verfasstheit. Baumer sagt: „It is not only about examining the current state of the world or one’s self but also about envisioning alternatives“.²⁴

Die drei Dimensionen, wie sie in dem Artikel beschrieben sind, weisen keine scharfen Grenzen auf, so dass *Beobachtungen* von reflexivem Verhalten, teilweise sowohl dem einen, wie dem anderen Typus zugeschrieben werden könnten. Der Fokus von Baumers Framework liegt auch nicht auf einer generellen theoretischen Fundierung einer technisch-vermittelten Reflexion, sondern auf einer konzeptuellen Diskussion, die das Ziel hat, die Perspektive der technischen Community zu erweitern.

Baumers Konzeption ist für meine Arbeit geeignet, weil sie disziplinar in der HCI-Forschung diskutiert wird und dabei vielschichtig genug ist, um den interdisziplinären Ansprüchen der Arbeit gerecht zu werden.

5.3.2 Übersicht: Studiendesign

In der Studie wurden, wie weiter oben erwähnt, zwei Arten von Visualisierungen von Aktivitäten der Teilnehmenden verglichen. Dabei war es Ziel, die beiden Arten von Visualisierungen hinsichtlich der Konzeption von Reflexion aus dem vorherigen Abschnitt zu vergleichen, wobei die eine Darstellungsart Systemmechanismen vermittelt und die andere nicht.

²⁴Baumer, 2015, S. 591.

In der Studie nahmen insgesamt 15 Personen teil, wovon 10 weiblich und 5 männlich waren. Das Alter lag zwischen 26 und 53 Jahren (im Durchschnitt 33 Jahre). Rund die Hälfte der Teilnehmenden (7) wurden in London an der UCL rekrutiert und die übrigen Teilnehmenden fanden sich in Darmstadt an der Technischen Universität (8). Die meisten der Teilnehmenden hatten keinen technischen oder naturwissenschaftlichen Hintergrund und kamen hauptsächlich aus den Fakultäten der Geistes- und Sozialwissenschaften.

Die Aufgabe der Teilnehmenden bestand darin, das *Hedgehog*-Sensorsystem eine Woche lang, nach Möglichkeit Tag und Nacht, zu tragen. Insgesamt ergab das am Ende in Summe aller Teilnehmenden rund 1470 Stunden an aufgezeichneten Aktivitätsdaten. Im Anschluss wurden die Daten scheinbar, so glaubten zumindest sie, *maschinell* gedeutet (dazu gleich mehr) und hinsichtlich der zwei Arten von Visualisierungen dargestellt. Die Teilnehmenden wurden dann gebeten, die Daten anhand der Visualisierungen zu betrachten, wobei die Betrachtung durch eine Reihe von Aufgaben sowie eines *semi-structured* Interview, begleitet wurde. Diese Kombination aus Betrachtung und Interview dauerte ungefähr eine Stunde, was mit Hilfe einer Audioaufzeichnung dokumentiert wurde.

Circa die Hälfte der Teilnehmenden hatten keine Erfahrung mit Self-Tracking-Techniken. Die andere Hälfte hatte schon einmal ein Gerät verwendet, was im weitesten Sinne als Technik zur Aktivitätserkennung bezeichnet werden kann. Dazu gehörten Schrittzähler und *Joggingassistenten* auf dem Smartphone. Die Verwendung eines Systems zur maschinellen Deutung einer umfassenden Menge von Aktivitäten war für alle neu.

5.3.3 Versuchsaufbau

Ein zentraler Aspekt des Versuchsaufbaus der Studie war die Gestaltung der zwei Typen von Visualisierungen, die verglichen werden sollten. Ein wichtiges Kriterium für eine Vergleichbarkeit der Visualisierungen war es, dass beide in ihrer Erscheinung und ihren Nutzungsmöglichkeiten möglichst symmetrisch sind. Das betrifft vor allem die Rohdaten, die wie in der Studie zuvor als eine Art der Vermittlung von Systemmechanismen auch in dieser Studie zur Anwendung kam. Es musste nach einer adäquaten ähnlichen Darstellungsform (Graphen in Form einer Zeitreihe) gesucht werden, die einen Aspekt der Daten darstellt, aber nicht als eine Vermittlung von Systemmechanismen aufgefasst werden kann. Die geeignete Darstellung wurde in der Gestalt des Mittelwertgraphen gefunden. Im Folgenden werden die beiden Typen von Visualisierungen, die als *simple type* und *insights type* bezeichnet werden, getrennt voneinander besprochen.

Die Systemklassifikationen: Ein Wizard of Oz Ansatz

Bei dieser Studie wurde, anders als in der ersten Studie, kein maschinelles Erkennungssystem verwendet, um die Aktivitäten der Teilnehmenden zu deuten. Vielmehr wurde ein *Wizard-of-Oz-Ansatz* gewählt, bei dem die Deutungen der Aktivitäten manuell erstellt wurden. Der Grund dafür lag in der Anforderung, dass trotz des *wild*-Ansatzes ein Maß an Kontrollierbarkeit bewahrt werden sollte, um aussagekräftige Ergebnisse zu erhalten. Anders als bei der ersten Studie machten die Teilnehmenden sehr unterschiedliche Erfahrungen im Umgang mit den Aktivitätsvisualisierungen, da das Dargestellte davon abhing, was die Personen in der Woche erlebt hatten, in der sie den Sensor trugen. Da die maschinelle Deutungsqualität bei nicht individuell gelernten Systemen sehr unterschiedlich ist (eine individuell gelerntes System war vom Aufwand her nicht möglich), wären die Erfahrungen der Teilnehmenden noch einmal unterschiedlicher gewesen. Eine manuelle Deutung, die von mir erzeugt wurde und auf den Erfahrungen mit realen Deutungssystemen basierte, war dahingehend konsistenter und zuverlässiger.

Die manuelle Klassifikation, die einer Systemklassifikation, wie sie in der ersten Studie zur Anwendung kam, nachempfunden ist, wird im Folgenden als Systeminterpretation bezeichnet. Die Bezeichnung Systeminterpretation wurde auch in der Kommunikation mit den Teilnehmenden der Studie verwendet.

Visualisierung A: *simple type*

Bei Visualisierung A handelt es sich um eine einfache Darstellung, welche die Systeminterpretation sowie einen Mittelwertgraphen enthält (Abbildung 5.9, oben). Der Mittelwertgraph ist eine Fusion der Rohdaten, die die Beschleunigungen der drei Achsen zu einer mittleren Bewegung zusammenfasst. Es handelt sich hierbei nicht um eine Vermittlung von Systemmechanismen, sondern um eine zusätzliche Auswertung. Der Mittelwertgraph zeigt das Aktivitätslevel abhängig von der Tageszeit.

Es ist eine Darstellungsweise, wie sie auch ähnlich in gängigen kommerziellen Techniken zur Aktivitätserkennung zur Präsentation von Aktivität genutzt wird. Aus diesem Grund wird dieser Typ von Visualisierung als *simple type* bezeichnet. In der Studie wurde die Bezeichnung Visualisierung A gebraucht und nicht *simple type*, um keine Bewertung nahezu legen.

Visualisierung B: *insights type*

Bei Visualisierung B handelt es sich um den Typus, der zusätzlich zu der Systeminterpretation Aspekte der Systemmechanismen darstellt (Abbildung 5.9, unten). Dieser wird im Folgenden als *insights type* bezeichnet.

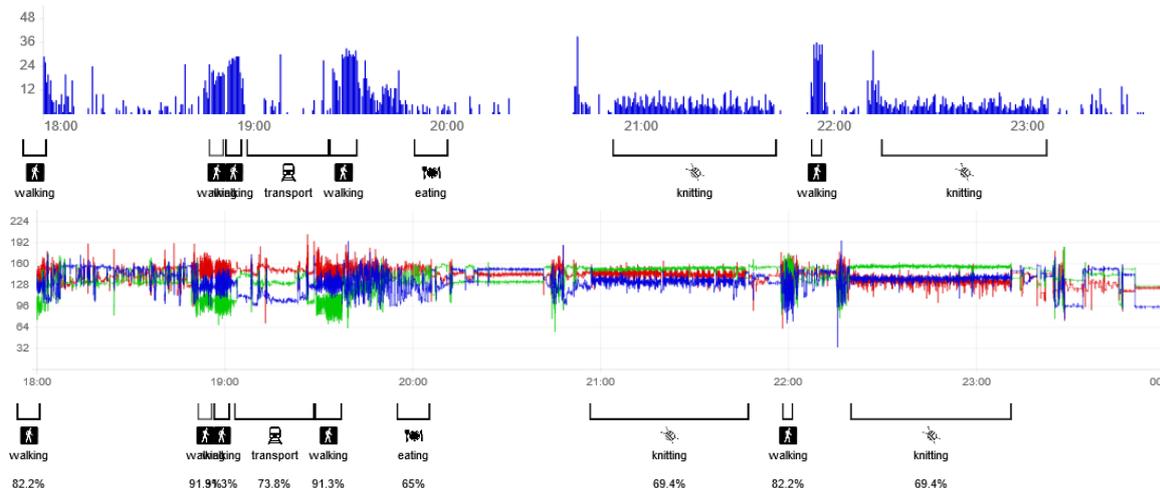


ABBILDUNG 5.9: Ein Beispiel für beide Arten von Visualisierungen. Der obere Graph zeigt ein Beispiel für den *simple type*, im unteren Graphen für den *insights type*. Es handelt sich um eine Sequenz von Aktivitäten in einem Zeitraum von 18.00 Uhr bis 24.00 Uhr. Beide Visualisierungen zeigen die gleiche Sequenz

Die Systemmechanismen werden auf zwei Arten vermittelt: mithilfe der Darstellung der Quelldaten, das heißt, wie zuvor die Rohdaten des dreiaxigen Beschleunigungssensors und anhand einer Darstellung der Qualität der Systeminterpretationen. Die Rohdaten sind in einer dreifarbigen Zeitreihe visualisiert, die mit der Darstellung in der ersten Studie identisch ist. Die Qualität der Deutung ist für jeden Typen einer Aktivität unterschiedlich. Dem Nutzer wurde die Qualität als ein Prozentwert vermittelt. Die Werte entsprechen der durchschnittlichen Genauigkeit, mit der eine Aktivität gedeutet wird, die sich aus den Daten der Evaluierung in der Trainingsphase ergibt. Da es keine wirkliche Interpretation des Systems gab, wurde die Qualität aus den Ergebnissen von Studie 1 und weiteren Erfahrungen mit der maschinellen Erkennung und deren *Erkennungsrate* übernommen (siehe: Aktivitätentabelle Studie 1 in Abschnitt 5.1).

In den folgenden Abschnitten werden die zwei Sektionen, *Interaktion* und *Fragebogen* beschrieben, die die Teilnehmenden im Rahmen der Betrachtung ihrer Aktivitätsdaten durchlaufen mussten. Es lagen 2 bis 4 Tage zwischen dem Ende des Tragens und der Betrachtung der gedeuteten Daten.

Sektion: Interaktion

Nachdem die Teilnehmenden das Sensorsystem für eine Woche getragen hatten, wurden die Daten zur weiteren Bearbeitung heruntergeladen. Basierend auf diesen Daten wurden die Visualisierungen, die wie beschrieben aussehen, erzeugt.

Die Visualisierungen wurden den Teilnehmenden in einem Tool präsentiert, das ein Interagieren zuließ. Die Teilnehmenden hatten die Möglichkeit, frei zwischen den Aktivitätsdarstellungen der Tage hin und her zu wechseln, sowie in dem Mittelwertgraph und dem Rohdatengraph hinein und heraus zu zoomen. Ausgehend von einer 24-Stunden-Übersicht, die bezüglich der Rohdaten identisch mit der aus Studie 1 ist, konnte drei Mal der Ausschnitt vergrößert werden, so dass eine Skalenweite von zwei Minuten möglich war (in der Übersichtsdarstellung lag die Skalenweite bei 4 Stunden).

Damit es den Teilnehmenden möglich war, die zwei Arten von Visualisierungen zu vergleichen, wurden die einzelnen Tage entweder in der einen oder in der anderen Weise dargestellt. Das heißt, wenn beispielsweise am Montag der *insights type* präsentiert wurde, konnten die Teilnehmenden nicht auch noch den *simple type* für den gleichen Tag sehen. Der *simple type* war dann die Darstellungsweise für den Dienstag. Mit welchem Typen von Visualisierung begonnen wurde, war Ergebnis einer Losung.

Vor Beginn dieser Sektion wurde den Teilnehmenden eine kurze Einführung in das Tool und die Visualisierungen gegeben. Das beinhaltete die zuvor beschriebenen Funktionen des Tools, die Benennung der Graphen als Rohdaten beziehungsweise Mittelwert sowie die Qualität der Erkennung in Prozent.

Die Aufgabe der Teilnehmenden bestand darin, ihre Tage zu rekapitulieren und dabei zu bewerten, ob sie mit den Deutungen des Systems einverstanden sind. Zudem wurden sie gebeten, sich mit Hilfe der Daten zu erinnern, was sie am Morgen als erstes auf der Arbeit gemacht haben. Zudem wurden die Teilnehmenden schon bevor sie das Sensorsystem für eine Woche trugen, nach einem speziellen Interesse hinsichtlich ihres Verhaltens gefragt. Hatten sie ein solches Interesse angegeben, wurden sie während des Interaktionsteils der Studie danach gefragt.

Sektion: Fragebogen

Nach Abschluss der Sektion, die eine Interaktion mit dem Tool beinhaltete, gab es einen Fragebogen, in dem generelle Fragen zu den Erfahrungen der Teilnehmenden mit dem System zu beantworten waren. Der Fragebogen bestand aus vier Teilen: Als erstes mussten die Teilnehmenden beantworten, welche der zwei Arten von Visualisierungen sie bevorzugen. Die Bewertung erfolgte anhand einer *7-point Likert-scale*, die das Spektrum von wenig hilfreich bis sehr hilfreich repräsentierte. Zusätzlich mussten sie ihre Entscheidung begründen. Der zweite Teil der Fragen zielte auf einzelne Elemente der Visualisierungen ab. Einige Elemente kamen in beiden Arten

von Visualisierungen vor und manche nur in einer von beiden. Diese wurden anhand der selben siebenstufigen Skala bewertet. Die Elemente waren: Rohdatengraph, Mittelwertgraph, Zoomen, Systeminterpretationen (Deutungen des Systems) und die Genauigkeiten (Qualität der Erkennung). Die dritte Sektion von Fragen bezog sich auf das Empfinden von Vertrauen hinsichtlich der zwei Arten von Visualisierungen. Auch hier wurde die *7-point Likert-scale* verwendet, die von wenig Vertrauen bis viel Vertrauen reichte. Der letzte Teil des Fragebogens beschäftigte sich mit Aspekten der Privatsphäre und der Kontrolle von persönlichen Daten (zusammenfassend als *Privacy* bezeichnet). Die Teilnehmenden wurden gefragt, welche Daten, die in den Visualisierungen zu finden sind, sie im engen Freundeskreis oder mit der Öffentlichkeit teilen würden. Sofern die Teilnehmenden Erfahrungen mit Self-Tracking-Techniken hatten, wurde an dieser Stelle darauf Bezug genommen und nach ihrer bisherigen Praxis zum Teilen bzw. Veröffentlichlichen von personenbezogenen Daten gefragt.

5.3.4 Ergebnisse der Studie

Die meisten der Teilnehmenden haben das Sensorsystem Tag und Nacht getragen, wie es gefordert war. Eine Teilnehmende brach die Studie nach drei Tagen vorzeitig ab, weil sie das Gefühl ständiger Überwachung hatte. Die Teilnehmende erklärte sich bereit, die Studie mit den vorhandenen Daten in der zweiten Phase, das heißt der Sektion *Interaktion* und *Fragebogen*, fortzuführen. Dies geschah Selbstverständlich auf freiwilliger Basis, da die Teilnehmenden auch ohne Angabe von Gründen und ohne weitere Verpflichtungen den Versuch jederzeit abbrechen konnten.²⁵ Eine andere Person machte davon Gebrauch und brach die Studie schon nach einigen Stunden ab, so dass eine Bitte um ein Interview wenig sinnvoll gewesen wäre. Diese Person ist nicht Teil der schlussendlich 15 Teilnehmenden der Studie. Es gab eine Reihe von Gründen, warum einige Personen das Sensorsystem für einen kurzen Zeitraum abgenommen haben. Beispielsweise gaben diese Teilnehmenden an, auf der Geburtstagsfeier der Mutter oder bei einem Treffen mit einem alten Freund, den sie schon lange nicht mehr gesehen hatte, gewesen zu sein. Die Vorstellung, mit dem Sensor gesehen zu werden, wäre ihnen unangenehm gewesen.

Grundlegender Vergleich der Visualisierungen

Für die Darstellung der Ergebnisse der Studie wird mit der quantitativen Auswertung des Fragebogens begonnen. Hier wurden drei Fragen gestellt,

²⁵Sie wurden vor Beginn der Studie in einer Teilnehmererklärung darüber aufgeklärt.

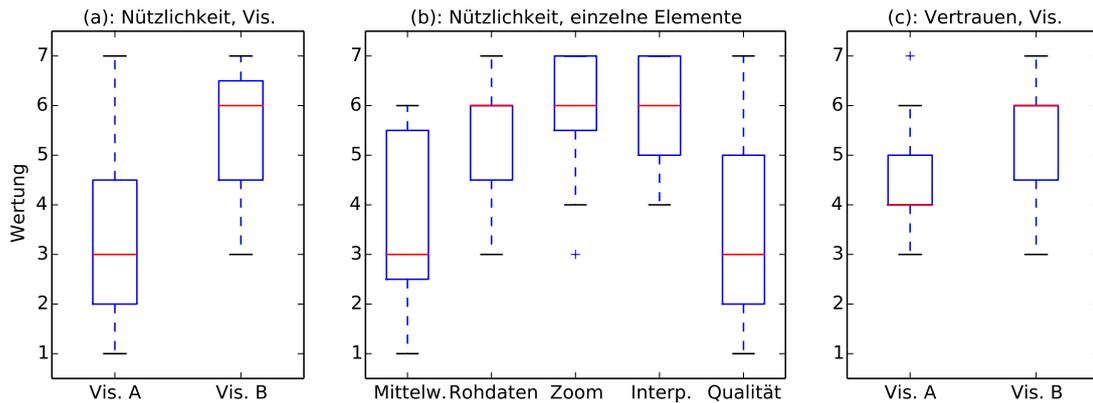


ABBILDUNG 5.10: Evaluierung der Typen von Visualisierungen und der einzelnen Elemente: (a) Nützlichkeitsbewertung, Typen von Visualisierungen; (b) Nützlichkeitsbewertung, einzelne Elemente; (c) Vertrauensbewertung, Typen von Visualisierungen

die anhand einer 7-point Likert-scale bewertet wurden. Die Ergebnisse dieser Bewertung sind in Abbildung 5.10 anhand eines Boxplots dargestellt.

Charakteristika Visualisierung A, der simple type
„More clear to understand“ (P03)
„Pro: Übersichtlichkeit“ (P14)
„Hilfreich für die Tagesanalyse auf einen Blick, sonst zu ungenau“ (P15)
„Für einen schnellen Blick besser“ (P13)
„Simple, less mental work“ (P05)
„More clear, easy to understand“ (P02)
„It is like the Fitbit - its more boring“ (P03)
„The blue one doesn't make so much sense to me“ (P06)
„Ich habe das Gefühl, da fehlt etwas“ (P09)

TABELLE 5.2: Eine Sammlung von Kommentaren, die den simple type betreffen

Die erste Frage, die die Teilnehmenden zu beantworten hatten, betraf die Einschätzung der beiden Typen von Visualisierungen hinsichtlich ihrer Nützlichkeitsbewertung. In Abbildung 5.10a sind die Ergebnisse zu sehen. Die Abbildung zeigt, dass die meisten Teilnehmenden den Visualisierungstypen bevorzugten, der zusätzlich zu den Systeminterpretationen auch Aspekte der Systemmechanismen vermittelt (*insights type*). Zudem wurden die Teilnehmenden explizit nach den Gründen ihrer Präferenz befragt. In den Tabellen 5.3 und 5.2 sind die Charakterisierungen der beiden Typen zusammengefasst, die von den Teilnehmenden als Begründung ihrer Entscheidungen genannt wurden. Die eher positiven Charakterisierungen sind in der jeweiligen Tabelle der Typen weiter oben zu finden und die negativen weiter unten.

Es gibt auch einige Teilnehmende, die argumentierten, dass es gut wäre, beide Typen von Visualisierungen zu haben. Eine solche Meinung findet

sich zum Beispiel in diesem Kommentar einer Teilnehmenden wieder:

„Am Anfang habe ich gedacht, dass ich mit dem blauen [Graphen] mehr anfangen kann; bei dem anderen habe ich anfangs nichts erkannt. Wenn man in beide reinguckt, sieht man dann, dass man mit denen unterschiedliche Sachen sehen kann“ (P08).

Charakteristika Visualisierung B, der <i>insights types</i>
„More appealing“ (P03)
„More serious, more scientific, more legit“ (P03)
„Shows characteristics of activity“ (P01)
„Hübscher, differenzierter“ (P11)
„Durchgängiger Datenstrom, mehr Information“ (P10)
„Gut geeignet für die Aktivitätsanalyse in feiner Form“ (P15)
„In the raw data it is easier to see patterns“ (P06)
„Die Rohdaten sind detaillierter“ (P07)
„Eine Ablenkung durch andere Daten kommt bei diesem Typen öfter vor als bei dem anderen“ (P14)
„Looks scientific reasonable but I do not see the value“ (P02)

TABELLE 5.3: Eine Sammlung von Kommentaren, die den *insights type* betreffen

Des Weiteren wurden die Teilnehmenden gebeten, einzelne Elemente der Visualisierungen zu bewerten. Das Ergebnis ist in Abbildung 5.10b dargestellt. Der Vergleich zwischen den Elementen zeigt deutlich, dass es vor allem die Rohdaten sind, die die Teilnehmenden an dem *insights type* schätzen. Ein weiteres exklusives Element des *insights types* ist die Genauigkeit der Deutung, die aber als deutlich weniger hilfreich bewertet wurde. Das spiegelt sich auch in den Kommentaren aus Tabelle 5.3 wider, in der fast ausschließlich der Rohdatengraph in der positiven Charakterisierung des *insights types* Erwähnung findet. Eine wichtige Erkenntnis hinsichtlich einer systemischen Umsetzung ist die Tatsache, dass die Teilnehmenden die Möglichkeit, zoomen zu können, das heißt, zwischen einer detaillierteren und einer allgemeineren Ansicht wechseln zu können, sehr schätzten. Dazu ist anzumerken, dass bestimmte charakteristische Muster von Aktivitäten in den Rohdaten nur in einer bestimmten Granularität des Graphen sichtbar sind. Dieser Aspekt wird in der anschließenden qualitativen Auswertung des Verhaltens der Teilnehmenden näher behandelt.

Im dritten Teil des Fragebogens wurde nach dem Vertrauen gefragt, das die Teilnehmenden mit den unterschiedlichen Visualisierungen verbinden. Abbildung 5.10c zeigt die Ergebnisse. Hier liegt, wie bei der Frage nach der Nützlichkeit, der *insights type* auf dem ersten Platz. Es ist anzunehmen, dass ein Zusammenhang zwischen der Charakterisierung des *insights types* als wissenschaftlich und exakt und dem Gefühl von Vertrauen besteht.

Detaillierte qualitative Auswertungen

Während der Studie konnten eine Vielzahl von Beobachtungen gemacht werden, bei denen die Teilnehmenden die Spezifika der Darstellungsweisen nutzten, um Informationen zu extrahieren, die über die Systeminterpretationen hinausgingen.

- Ausdifferenzieren von Aktivitätsmustern

Im Folgenden werden die Beobachtungen vorgestellt, bei denen die Teilnehmenden die Visualisierungen nutzten, um Aktivitäten, in den Beispielen Schlafen und Laufen auszdifferenzieren.²⁶

Bewegungsdaten als Schlafen zu deuten, gilt als eine Standardaufgabe für Systeme zur Aktivitätserkennung. Es gibt eine Reihe von Beispielen, in denen die Teilnehmenden, die Spezifika der Visualisierungen nutzten, um ihr Schlafverhalten auszdifferenzieren.

In Abbildung 5.11 ist ein Beispiel dargestellt, in dem eine Teilnehmende deutete, wann sie Schwierigkeiten hatte, gut zu schlafen und in welchem Zeitraum die Teilnehmende tief schlief.

"Yes, you can see that, I was a little bit active, and then I fall asleep finally, but only for one hour. It is so interesting that you can say that is deep sleep". (P02)

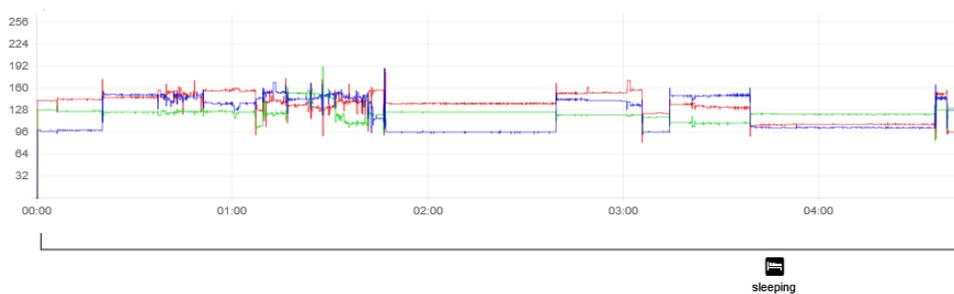


ABBILDUNG 5.11: Die Zeit nach dem ins Bett gehen bis kurz vor 2 Uhr nachts schlief die Teilnehmende nicht gut, danach gab es längere ruhige Phasen

Andere Teilnehmende hatten die Intensität und die Frequenz von Bewegungen in den Graphen als Indikatoren für schlechten oder guten Schlaf gedeutet. Oft war nach einer bestimmten Zeitspanne, beispielsweise jede Stunde, ein deutlicher Peak zu sehen. War das Erscheinen von Peaks regelmäßig und nicht zu oft und zu stark, oder gab es sogar überhaupt keine beobachtbare Bewegung, wurde das als guter Schlaf gewertet. Es ist gut vorstellbar, dass auch bestimmte Schlafphasen,

²⁶In der Studie als *sleeping* und *walking* gelabelt.

wie sie in der Schlafforschung definiert sind, in den Daten teilweise erkennbar wären. Unabhängig davon, ob die Graphen von den Teilnehmenden richtig gedeutet wurden, gab es in jedem Fall eine qualitative Auseinandersetzung mit dem eigenen Schlafverhalten.

Üblicherweise wird die Aktivität Laufen von kommerziellen Systemen zur Aktivitätserkennung als gelaufene Zeit, Distanz oder in Schritten repräsentiert. Selbstverständlich kann sich Laufen qualitativ stark unterscheiden, was sich auch bei den Beobachtungen in der Studie widerspiegelt. Die folgenden Beispiele beziehen sich auf Beobachtungen, bei denen die Teilnehmenden mit dem *insights type* konfrontiert waren. Bei diesem Typen gab es einige Situationen, in denen Personen ihr Verhalten differenzierten, das heißt, Wissen erlangten, das über die Erkenntnis hinausging, dass es sich um Laufen handelt.

Es war einer Teilnehmenden zum Beispiel möglich, das Laufen mit einem Kinderwagen von einem *Solo-Laufen* zu unterscheiden. Dabei handelte es sich um einen Besuch im Park, bei dem sie sich mit ihrem Partner beim Schieben abgewechselt hatte. Das zeigt sich im ersten Teil des unten genannten Zitats. Da sie das Muster einmal identifiziert hatte, fand sie das Muster an einer anderen Stelle wieder, als sie ihr Kind vom Kindergarten abgeholt hatte (siehe Abbildung 5.13). Am Ende der Studie, das heißt im Frageteil, erwähnte die Teilnehmende, dass das Laufen mit dem Kinderwagen etwas sehr Charakteristisches ist, das sich in den Daten zeigt.

„Dazwischen wechseln wir uns immer ab mit dem Kinderwagen. Deshalb gibt es hier solche Unterbrechungen zwischen dem Laufen“. [...] „Und dann noch das Laufen mit Kinderwagen. Das war ja vor allem in der Ansicht mit den drei Farben sehr charakteristisch“. (P09)

Ein anderes Beispiel, dargestellt in Abbildung 5.12, zeigt eine Episode, die vom *System* als Laufen gedeutet wurde, die sich aber optisch in zwei Abschnitte teilt. Die Teilnehmende war erst irritiert über den Bruch in der Darstellung in den Rohdaten, konnte sich dann aber erinnern, dass sie auf dem Hinweg mit ihrem Handy *gespielt* hatte und auf dem Rückweg nicht.

„Doch ich weiß was es sein kann. Es ist meistens so, dass ich beim Hinweg mit dem Handy rumspiele und beim Rückweg laufe ich ganz normal“. (P09)

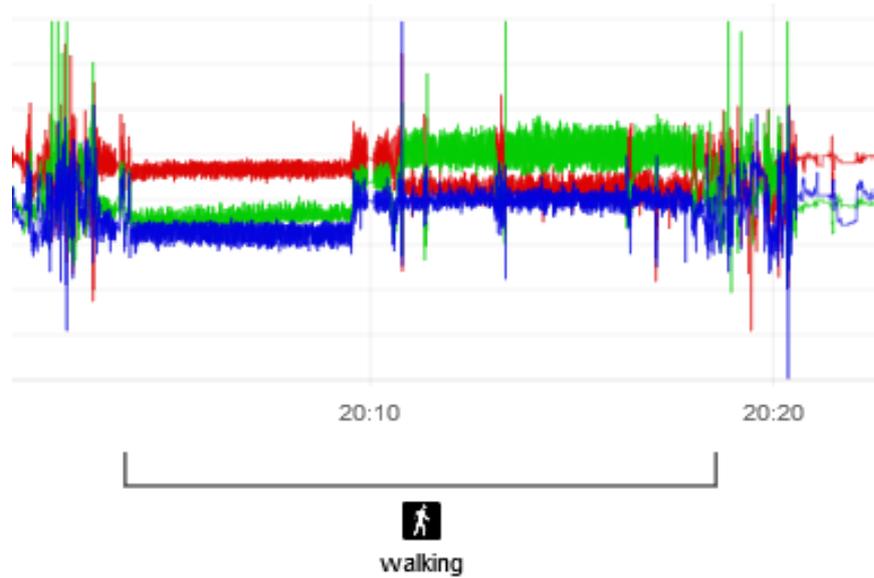


ABBILDUNG 5.12: Laufend mit dem Smartphone in der Hand (links) und laufend ohne Smartphone

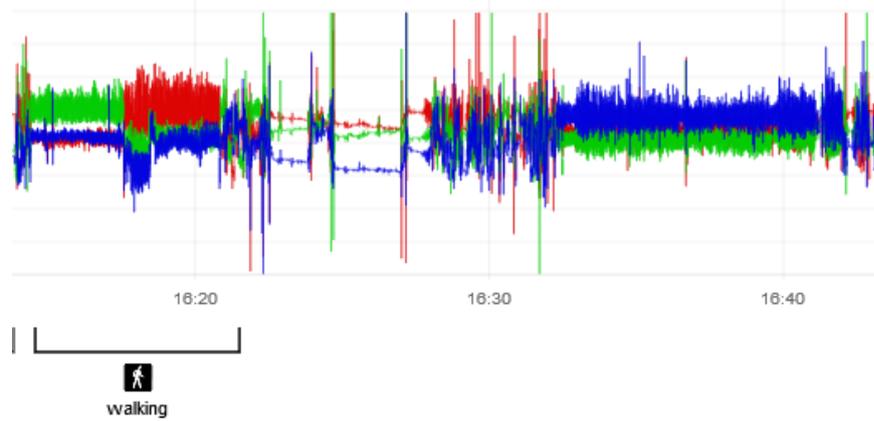


ABBILDUNG 5.13: Links: Auf dem Weg zum Kindergarten, Links: Auf dem Weg zurück mit dem Kinderwagen

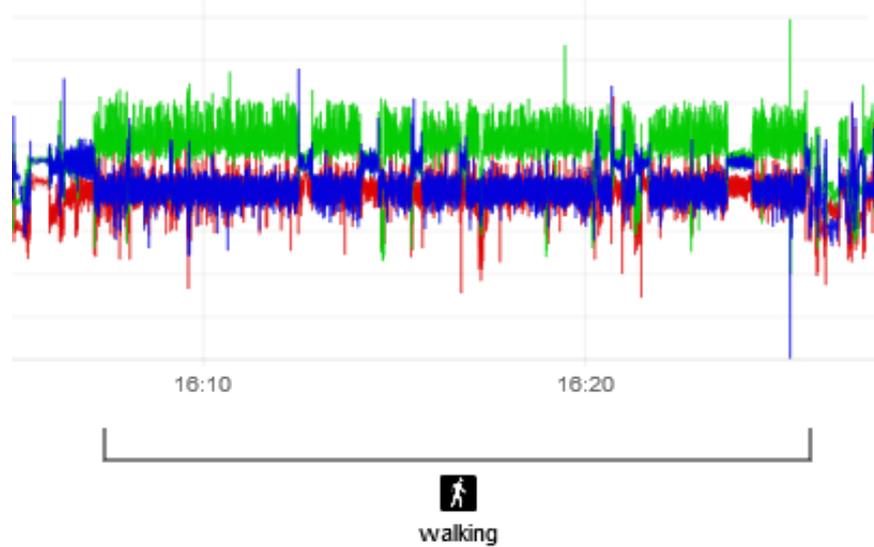


ABBILDUNG 5.14: Laufen auf dem Laufband im Fitnessstudio

Die Erkenntnisse folgten nicht aus der Betrachtung der Daten alleine, jedoch ermöglichte die Betrachtung der Daten eine neue Perspektive, indem beispielsweise durch die Verwunderung über das Dargestellte eine Erinnerung ausgelöst wurde. Es gibt noch andere Beispiele in denen Teilnehmende eine Ausdifferenzierung eines bestimmten Laufverhaltens vornahmen, beispielsweise als eine Teilnehmende Abschnitte in dem sie mit einer Einkaufstüte lief von einem Laufen ohne Tüte unterschied (P05).

Ein weiteres Beispiel ist eine Episode in der eine Teilnehmende ein Laufband im Fitnessstudio benutzte (siehe Abbildung 5.14). Hier war die Teilnehmende irritiert, wo sie denn gelaufen wären, da sie sich zu der Zeit im Fitnessstudio befunden hatte. Die Irritation kam daher, dass die Person nicht dachte, dass das System *Laufen auf dem Laufband* als Laufen erkennen könnte, da man sich doch nicht wirklich fortbewegt. Ein Vergleich mit anderen Mustern *normalen Laufens* und der genauen Erinnerung an die Zeit auf dem Laufband überzeugte die Teilnehmende, dass es dem System grundsätzlich möglich ist auch *Laufen auf dem Laufband* als Laufen zu erkennen.

Die Beispiele zeigen, dass vor allem die Rohdaten zu einem Nachdenken über die Aktivitäten beziehungsweise des Dargestellten anregen.

- Identifizieren von Aktivitätsmustern

Zu Beginn waren die Teilnehmenden insbesondere von den Rohdaten überfordert und fragten sich unter anderem, was jede Farbe repräsentiert. Während der Interaktion mit dem Tool hat eine Mehrzahl der Teilnehmenden zumindest teilweise gelernt, die Graphen in verschiedenen Hinsichten zu deuten. Den Rohdaten als eine Art der Vermittlung von Systemmechanismen wurde in der Studie die meiste Aufmerksamkeit geschenkt.

Zentral war dabei die visuelle Erscheinung von Regelmäßigkeiten, die sich von Segmenten ohne Regelmäßigkeit oder von Segmenten mit anderer Regelmäßigkeit unterschieden. Wie damit umgegangen wurde, zeigten teilweise schon die vorangegangenen Beispiele. Dies wird im Folgenden aber noch einmal vertieft.

In diesem Abschnitt werden die Beispiele zusammengefasst, in denen die Teilnehmenden lernten, einzelne Muster oder generell Muster in den Rohdaten zu verstehen. Das erste Beispiel befasst sich mit einem Abend im Fitnessstudio, in dem die Teilnehmende unterschiedliche Übungen an verschiedenen Geräten durchführte. Das Beispiel, das in Abbildung 5.15 dargestellt ist, wurde von ihr wie folgt kommentiert:

„Die ist die Übung für die geraden Bauchmuskeln. Die mache ich etwas länger. Dann mache ich kurz Pause, und dann links und rechts jeweils einen Satz. Dann mache ich eine Pause. Dann wieder die geraden ein bisschen länger, Pause, links und rechts, Pause, die geraden, links und rechts“.
(P15)

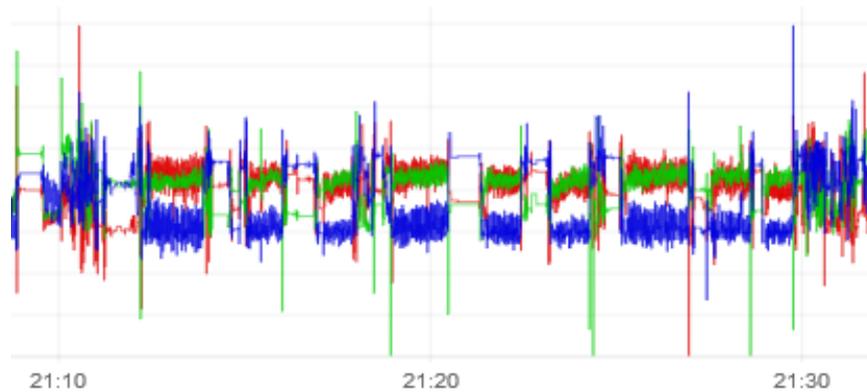


ABBILDUNG 5.15: Übungen im Fitnessstudio

Ein anderes Beispiel befasste sich mit einem Arbeitstag an der Universität. Die Teilnehmende hatte eine Reihe von Meetings im Verlauf des Tages. Mit Blick auf die Rohdaten fand sie Hinweise, welcher Abschnitt ein Treffen mit Studenten zeigte und welcher ein Treffen mit dem eigenen Betreuer (Supervisor). Das folgende Zitat stammt von dem Zeitpunkt, an dem sie erkannte, dass es sich bei einem Zeitabschnitt um ein Treffen mit Studenten handelte. Sie war überrascht, wie aktiv sie in dem Treffen war, so dass sie es fast als demütigend empfand, wie viel sie gesticuliert hatte, wenn sie mit den Studenten sprach bzw. ihnen etwas erklärte (siehe Abbildung 5.16). Die Optik der Rohdaten unterschied sich grundlegend von einem späteren Treffen mit ihrem Betreuer an dem sie weniger gesticuliert hatte. Auf diese Unterscheidung kam sie mehrmals in der Betrachtung der aufgezeichneten Woche zurück.

„I had that meeting with a group of students, around 11.30. This is really active, which makes sense, because when I talk, I gesticulate a lot. An that was an hour meeting, pretty much that is me talking. That is almost embarrassing“.
(P03)

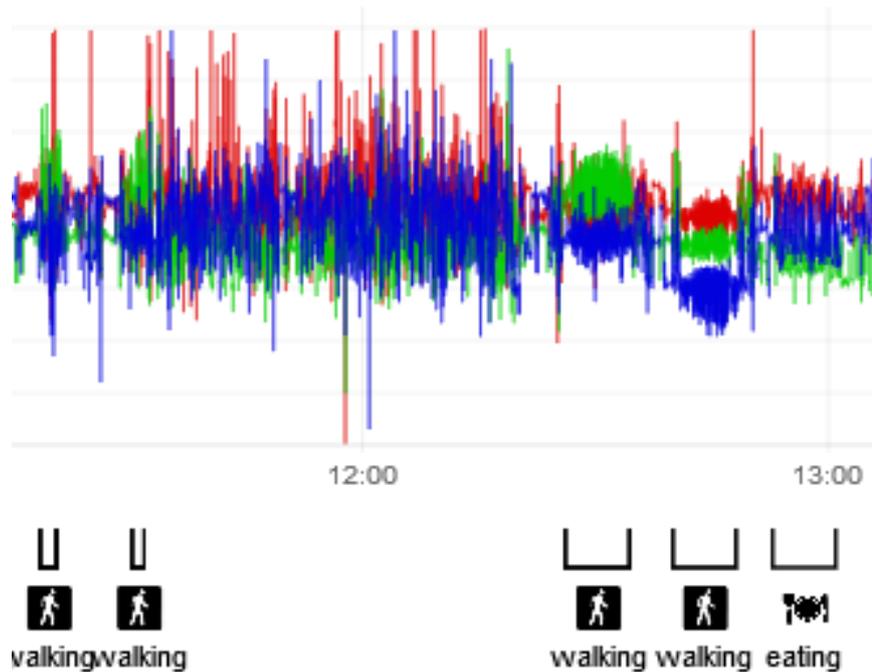


ABBILDUNG 5.16: Um 11.30 traf sich die Teilnehmende mit Studenten, die sie betreute

Privatsphäre und Datenschutz

Der letzte Teil des Fragebogens befasste sich mit Fragen zur Bereitschaft der Teilnehmenden, personenbezogene Daten zu teilen. Die Herausforderungen, die die Privatsphäre sowie den Datenschutz betreffen, sind mit Blick auf die Techniken des Self-Tracking von besonderer Brisanz. Alltägliche Aktivitäten, die heute schon mit kommerziellen Geräten gedeutet werden können, sagen viel über das Leben von Personen aus, wie es zu Beginn mit Blick auf die Anwendung in der psychologischen Therapie angesprochen wurde.

Ein System, wie es hier verwendet wurde, das potentiell ein sehr breites Spektrum an Aktivitäten erkennen könnte, ist besonders kritisch. Mit Bezug auf die Vermittlung von Systemmechanismen liegt der Schluss nahe, dass davon ein positiver Effekt für die Privatsphäre ausgeht, weil sie dem Nutzer Informationen über den Umgang mit den Daten vermittelt.²⁷

In diesem letzten Frageteil wurden die Personen gebeten, anzugeben, welche Elemente in den Visualisierungen sie gewillt wären, zum einen mit der Öffentlichkeit und zum anderen mit Freunden zu teilen. Einige der

²⁷Es gibt eine rege wissenschaftliche Diskussion zum Konzept der Privatsphäre (Privacy), die sehr heterogen ist. Als eine prominente Vertreterin nennt Nelen Nissenbaum die richtigen Vorstellung über den Informationsfluss von persönlichen Daten als eine zentrale Forderung einer Bewahrung von Privatsphäre. Zu diesem Informationsfluss können meiner Meinung nach auch die Systemmechanismen gehören. Siehe dazu Nissenbaum, 2009.

Teilnehmenden erwähnten, dass sie denken, dass sich doch niemand für ihre Aktivitätsdaten interessieren würde. Bei genauerer Nachfrage wurde jedoch eingeräumt, dass solche Daten unter anderem für Versicherungen sowie zukünftige Arbeitgeber relevant sein könnten:

„I think, data tells a lot and the problem is, someone has the data, you do not trust. I think your future employer has a look on your data. This would be really weird“. (P02)

Einige Teilnehmende, die angaben, dass sie nicht gewillt seien, ihre Daten zu teilen, wurden bezüglich ihrer Self-Tracking-Historie und damit zusammenhängend mit ihrer Praxis zum Teilen von Daten befragt. Zwei Zitate zeigen die Reaktionen der Teilnehmenden auf diese Nachfrage:

„I think this is something I would share, because I already share it, and I completely forgot that. Maybe I should just log in and change it“. (P03)

„Naja gut, ich benutze häufig dieses *Runtastic*, das sehen Andere natürlich schon, aber das ist eigentlich kein bewusstes Teilen“. (P13)

Ein anderer interessanter Aspekt ist der Informationsgehalt von Beschleunigungsdaten, die erst durch eine maschinelle Deutung ihr Potential entfalten können. Sensordaten von einem Beschleunigungssensor sind im Gegensatz zu beispielsweise visuellen Sensoren (Kameras) für sich alleine erst einmal nicht sehr aussagekräftig. Eine der Teilnehmenden bezog sich auf diesen Aspekt der Potenzialität.

„Ja, wenn man sich damit auskennt, kann man eine Menge aus den Daten herauslesen. Wenn ich es mit der Öffentlichkeit teile, dann teile ich es ja mit jedem, dann teile ich es auch mit dir und du kannst ja was damit anfangen oder irgendein anderer Experte auf diesem Gebiet. Da hätte ich jetzt nicht so Bock drauf“. (P09)

5.3.5 Diskussion der Ergebnisse

Die Auswertungen der Studie in quantitativer und qualitativer Hinsicht wurden in den vorherigen Abschnitten vorgestellt. Diese Ergebnisse werden im Folgenden hinsichtlich der Konzepte zur Selbstreflexion (5.4) diskutiert, insbesondere mit Bezug auf Baumers Arbeit. Die Möglichkeiten und Grenzen der Selbstreflexion werden dabei in Abhängigkeit von den beiden Visualisierungstypen betrachtet. Eine Diskussion der Ergebnisse mit Bezug auf die Hybridisierungstheorie folgt im abschließenden Abschnitt dieses Kapitels.

Diskussion anhand der Dimensionen von Selbstreflexion

In der folgenden Diskussion dienen die drei Dimensionen von Selbstreflexion, wie sie Baumer vorstellt, als Grundlage. Den Anfang macht ein Verständnis von Reflexion als *breakdown*. Dieser Typ bezeichnet ein auffällig werden von impliziten Strukturen, das heißt, eine Störung des Gewohnten als Ergebnis einer bewussten Explikation. Mit Bezug auf die Technik der Aktivitätserkennung werden Verhalten und Gewohnheiten in der Weise explizit gemacht, indem eine konkrete Deutung durch das System nahegelegt wird.

In der Studie gab es eine Reihe von Beobachtungen, die diese Merkmale aufwiesen, beispielsweise wenn eine der Teilnehmenden sagte: „it is a little depressing; my days are all the same“ (P06). Diese Aussage war mit Blick auf die Übersichtsdarstellungen der Visualisierungen hinsichtlich der alltäglichen Routinen, vor allem der Aktivitäten bei dem alltäglichen Weg zur Arbeit und zurück, gemacht worden. Die Beobachtung der Teilnehmenden bezog sich auf beide Typen von Visualisierungen, zwischen denen sie auf der Übersichtsebene hin- und herwechselte.

Häufiger wurde ein reflexiver *breakdown* beim *insights type*, vor allem hinsichtlich dem Visualisierungselement Rohdaten, beobachtet. Ein Beispiel aus der Studie war das Meeting mit Studenten, bei dem es die Teilnehmende nahezu beschämend („embarrassing“) empfand, als sie herausfand, dass die Ausschläge des Graphen mit den Bewegungen des Armes zusammenhängen müssen, der sich augenscheinlich sehr viel bewegt, wenn sie mit Studenten spricht. Bei beiden Beobachtungen einer Reflexion im Sinne eines *breakdown* ist es nicht klar, ob die Erkenntnisse hinsichtlich der Aktivität dazu führten, dass die Teilnehmenden etwas am eigenen Verhalten ändern wollten.

Der zweite Typus von Reflexion wurde als *inquiry* bezeichnet. Die Reflexion beschreibt demnach die aktive Praxis des Nachforschens in der Form eines Aufstellens und Testens von Hypothesen. Die meisten Anwendungen, die im Bereich des Quantified Self diskutiert werden, können diesem Typus zugeordnet werden. Das heißt, die Beobachtung von bestimmtem alltäglichem Verhalten, das quantitativ erfasst werden soll, um es mit anderen Aspekten zu vergleichen oder Langzeittrends zu erfassen. Die Quantifizierung beinhaltet diesen wissenschaftlichen Aspekt, nachdem das alltägliche Verhalten Gegenstand einer quantitativen Repräsentation sein kann, wobei einzelne Aktivitäten andere bedingen. Ein gezieltes Eingreifen kann dann gegebenenfalls zu einer Optimierung im Sinne einer Vermeidung eines bestimmten Verhaltens oder einer Änderung des Verhaltens führen.

Mit Blick auf die Ergebnisse der Studie kann man sehen, dass die Vermittlung von Systemmechanismen zu einem Wissen über die Bedingungen

der Zuordnung von einzelnen Labels zu bestimmten Daten, unter anderem Laufen und Schlafen, geführt hat. Dieses Wissen konnte dann in andere *Situationen* übertragen werden, wie es im Beispiel mit dem Kinderwagen der Fall war. Ein Ausdifferenzieren hin zu einzelnen Typen von Aktivitäten könnte dann Grundlage für ein weiterführendes Nachforschen werden. Basis für ein Ausdifferenzieren ist es, dass Personen Abschnitte mit verschiedenen Erscheinungen unterscheiden, vor allem die Übergänge zwischen verschiedenen Aktivitätsmustern erkennen. Die Regelmäßigkeit, mit der ein Muster erscheint, ist immer abhängig von der Perspektive. Aus diesem Grund war es sinnvoll, dass Nutzer die Granularität der Visualisierung der Graphen durch Zoomen verändern konnten. In der Studie konnte man beispielsweise beobachten, wie eine Person die Übungen im Fitnessstudio schnell in den Rohdaten wiederfinden konnte, ohne dass sie gelabelt waren. Ein solches Erkennen von Mustern könnte dann Grundlage für ein *händisches* Labeln sein, wobei die gelabelten Daten dann dem System als Trainingsdaten dienen können, was den Raum des Quantifizierbaren erweitern kann.

Bei der *transformation* handelt es sich um eine Reflexion, die eine mediale Vorstellung eines Bereichs als Ergebnis hat, das heißt, nicht nur die Interpretation von Beobachtungen als Ergebnis von tatsächlich Passiertem, sondern als Basis für alternative Vorstellungen. Inwiefern die Deutungen von alltäglichen Aktivitäten eine alternative Vorstellung beispielsweise über das Leben der Personen angeregt haben, lässt sich schwer aus den Ergebnissen der Studie ableiten. Ich argumentiere, dass die Vermittlung von Systemmechanismen die Ergebnisse (Systeminterpretationen) hinsichtlich der Bedingungen der Deutung kontextualisiert und so die Möglichkeit eröffnet, eine alternative Bewertung des eigenen Verhaltens zu erdenken. Zu wissen, wie das System arbeitet, sensibilisiert den Nutzer auch dafür zu begreifen, was nicht quantifizierbar ist und trotzdem Bedeutung hat.

Es hat sich gezeigt, dass alle Dimensionen der Reflexion in der *insights type*-Visualisierung, vor allem auf Basis von Rohdaten, beobachtet werden konnten oder zumindest vorstellbar sind.

Möglichkeiten und Grenzen dieses Ansatzes für ein Systemdesign

Es wurde gezeigt, dass die Vermittlung von Systemmechanismen, wie es anhand des *insights type* realisiert wurde, die Selbstreflexionen wie definiert befördern kann. Man kann sich vorstellen, dass der *insights type* in vielen Aspekten nachhaltiger sein kann, als der *simple type*, da er eine sehr individuelle Anwendung ermöglicht.

Aus praktischer Sicht kann es Gründe geben, die dafür sprechen, dass

bei einer realen Implementierung einer Darstellung von Systemmechanismen die Nutzer überfordert sind, vor allem dann, wenn sie nicht, wie in der Studie, durch die Rolle als Teilnehmende motiviert sind, sich mit den Graphen auseinanderzusetzen. In den Kommentaren, die die Teilnehmenden für die Bewertung der beiden Visualisierungstypen verfasst haben und die in den Tabellen (Tab. 5.2 und Tab. 5.3) zusammengefasst wurden, kann man auch herauslesen, dass nicht alle gewillt sind, sich mit der vielschichtigen Darstellung der Rohdaten auseinanderzusetzen.

Für die Teilnehmenden, die sich auf Muster konzentrierten und dabei nicht jeden Peak überbewerteten, hatte die Vermittlung von Systemmechanismen den größten Nutzen. Die Teilnehmenden waren alle Mitarbeiter oder Studenten an der Universität, so dass davon auszugehen ist, dass ihre Bereitschaft, sich mit komplexen Daten auseinanderzusetzen, höher ist als bei der durchschnittlichen Bevölkerung. Diese Tatsache ist relevant für eine Kontextualisierung der Ergebnisse.

Zudem ist anzunehmen, dass die Grenzen dessen, was Nutzer beispielsweise an Rohdaten verarbeiten können, mit den dreiachsigen Sensordaten, die in drei Farben visualisiert wurden, erreicht ist. Beispielsweise 6- bis 9-achsige Bewegungssensoren, wie sie bei aktuellen IMUs zu finden sind (siehe Abschnitt 5.1.2), übersteigen in ihrer Rohform eine verstehbare Darstellbarkeit. Da die so entstehende große Menge an Daten für die automatische Deutung durch den Klassifikator vorverarbeitet werden müssten, lassen sich auch hier Zwischenschritte finden, die vermittelt werden könnten.

Der Ansatz zur Vermittlung von Systemmechanismen kann als vielversprechend gesehen werden, um eine nachhaltige Nutzung von Self-Tracking-Techniken zu ermöglichen. Doch ist ein solches Systemdesign sicher nicht für jede Anwendung geeignet. Ein großes Potenzial kann man in Anwendungen sehen, in denen Personen gewillt sind, die zu Beginn vielschichtigen Visualisierungen verstehen zu lernen. Eine mögliche Anwendung könnte der Gesundheitsbereich sein, wie es beispielsweise als Szenario in der ersten Studie gezeigt wurde. Neben den schon angesprochenen Anwendungsszenarien, unter anderem in der psychologischen Therapie, ist die Therapie von chronischen Schmerzen ein möglicher Bereich. Eine der Teilnehmenden der Studie hatte einen psychotherapeutischen Hintergrund und erwähnte in der Diskussion ihre Einschätzung zu einem möglichen Nutzen für die Schmerztherapie.

„Oft ist es ja mit den Schmerzen so, dass Leute sich eben überwältigt fühlen und sich vom Schmerz leiten lassen. Aber wenn sie das Problem aktiv managen können, dann haben sie wieder das Heft in der Hand. Das ist ein wesentlicher psychologischer Gesichtspunkt“. (P08)

Man kann die Aussage so verstehen, dass ein einfaches oder sogar spielerisches Feedback, wie es bei gängigen Self-Tracking-Geräten häufig der Fall ist, bei Patienten ein Gefühl produzieren könnte, dass sie in ihrer Erkrankung nicht ernst genommen werden. Eine Beteiligung des Patienten bei der Deutung der eigenen Daten könnte ein geeigneterer Weg sein.

5.4 Die Ergebnisse der Studien aus Sicht der Hybridisierungstheorie

In beiden Studien wurde untersucht, ob die Vermittlung einzelner Aspekte der Systemmechanismen aus Sicht der Nutzer im Rahmen verschiedener Anwendung intelligenter Self-tracking-Systeme als hilfreich empfunden wurde. Verglichen wurden dabei Visualisierungen, die Aspekte der Mechanismen vermittelten, mit solchen, die keine Vermittlung von Aspekten enthielten. Der Vergleich zeigte, dass die Teilnehmenden in der Tat erstere präferierten (5.2.5 und 5.3.4). Die Präferenz konnte anhand der Fragebögen, in denen die Teilnehmenden ihre Erfahrungen mit den Visualisierungstypen resümierten, aufgezeigt werden. Neben dieser Auswertung wurden in der zweiten Studie zusätzlich Audiomitschnitte angefertigt, die den Umgang der Teilnehmenden dokumentierten und als Grundlage dafür dienten, die Möglichkeiten einer Selbstreflexion qualitativ zu untersuchen (5.3.5). Diese qualitativen Ergebnisse werden im Folgenden aus Sicht der Überlegungen zur Hybridisierung von Menschen durch Technik diskutiert.

Die Hybridisierung gilt als ein Prozess, der je nach System, Anwendung und Nutzer von unterschiedlicher Dauer sein kann. Der Beobachtungszeitraum der Studie, in dem der Umgang der Teilnehmenden mit der Technik untersucht wurde, war dagegen stark begrenzt – alles andere hätte den Rahmen der interdisziplinären Arbeit gesprengt. Man kann den in der Studie untersuchten Zeitraum als eine erste, intensive Begegnung mit dem System charakterisieren. Als eine zentrale Bedingung für den möglichen Erfolg einer Hybridisierung wurde die Möglichkeit einer medialen Erfahrung ausgearbeitet, die hinsichtlich von intelligenter Technik als schwierig erachtet wurde, weil eine Wahrnehmung bzw. Erfahrung von Spuren nicht möglich sei. Der Visualisierungstyp mit dem Namen *insights type*, der als einer von zwei Typen in der Studie Verwendung fand, hatte den Anspruch, die Schwierigkeiten einer Spurenwahrnehmung zu kompensieren. Auf Basis dieser Studienkonstellation stellt sich die Frage, ob es Beobachtungen gab, die als Hinweis aufgefasst werden können, dass es Teilnehmenden mit Hilfe dieser Visualisierungstypen möglich war, die Medialität des Systems zu erfahren.

Die Überlegungen von Christoph Hubig, die in Abschnitt 3.1.3 diskutiert wurden und in Abschnitt 3.3 hinsichtlich intelligenter Techniken vertieft wurde, bilden den Rahmen zur Beantwortung dieser Frage. Er nennt „Überraschungs-, Enttäuschungs-, Widerfahrnis- und Gelingenserlebnisse“ im Erproben eines Umgangs mit Technik, die in bestimmter Hinsicht als Spuren des Mediums erfahren werden können. Die mediale Erfahrung ist aber nur möglich, so die Theorie, wenn die genannten Erlebnisse im Kontext ihrer Verursachung betrachtet werden. Um die Ergebnisse der Studie hinsichtlich der Möglichkeit der Hybridisierung zu untersuchen, sind in erster Linie die Fälle von Interesse, in denen davon ausgegangen wird, dass die Teilnehmenden bezüglich ihrer Erwartungen enttäuscht wurden bzw. sie von den Nutzungsergebnissen überrascht waren.

Die transkribierten Audioaufnahmen bildeten die Datengrundlage für die Auswertung der Studie zu den Möglichkeiten und Grenzen der Selbstreflexion mit Self-Tracking-Techniken. Es war Aufgabe der Teilnehmenden, Auskunft über das eigene Verhalten zu geben und ihre Erfahrungen zu erklären. Da das Verhalten von den Teilnehmenden verbalisiert wurde, handelte es sich dabei um eine bewusste bzw. reflektierte Aktivität. Im Rahmen der Diskussion der phänomenologischen Ansätze wurde argumentiert, dass die Hybridisierung vor allem im Modus leiblichen Lernens erfolgt und gerade keine intellektuelle Leistung ist. In dieser Diskussion wurde aber auch hervorgehoben, dass sich das Erlernen bzw. Erproben eines Umgangs durch einen gewissen kognitiven Akt auszeichnet. Damit ist das Studiendesign prinzipiell geeignet, die Hybridisierungstheorie in dieser Hinsicht zu reflektieren.

In der Diskussion der Ergebnisse zur Selbstreflexion zeigte sich, dass die Teilnehmenden in vielen Fällen auf Basis der *insights typ*-Visualisierung Erfahrungen machten, die als eine mediale Erfahrung aufgefasst werden kann. Einige Situationen, die in dieser Hinsicht aufgefasst werden können, werden im Folgenden noch einmal dargestellt.

Zum Beispiel gab es Situationen, in denen Teilnehmende irritiert waren, dass die Ergebnisse, die das System nach Auswertung ihrer Aktivitäten visualisierte, nicht mit ihren Erinnerungen übereinstimmten. Da die Teilnehmenden eine adäquate Deutung ihres Verhaltens erwarteten, charakterisiere ich diese Fälle als enttäuschte Erwartungen. In einem Teilabschnitt in 5.3.4 wurden Nutzungssituationen beschrieben, in denen Teilnehmende Aktivitäten, die als Laufen gedeutet wurden (in der Studie *walking*), aus verschiedenen Gesichtspunkten weiter ausdifferenzierten. Es handelte sich dabei in einigen Fällen um enttäuschte Erwartungen, weil die Personen die Systemdeutung nicht teilten.

Das zeigte sich am Beispiel der Nutzungssituation, in der eine Teilnehmende sich wunderte, warum das System behauptet, dass sie über einen

längeren Zeitraum gelaufen sei, was sich aber nicht mit ihren Erinnerungen deckte. Vielmehr erinnerte sie sich, dass sie in der Zeit im Fitnessstudio gewesen sei. Als sie sich mit den Mustern in den Rohdaten auseinandersetzte, erkannte sie deutlich das Muster für Laufen, das sie an vielen anderen Stellen zuvor schon gesehen hatte und sich dort mit ihren Erinnerungen gedeckt hatte. Dabei wurde ihr klar, so die Deutung ihrer Aussagen, dass es sich um ein Laufen nicht nur handeln könnte, sondern geradezu handeln *musste*. Es musste demnach das Laufen auf dem Laufband sein, was hier erkannt wurde. Etwas, das ihr deshalb nicht in den Sinn gekommen war, weil sie Laufen mit einem zurückgelegten Weg verknüpft hatte, was ja auf dem Laufband nicht der Fall war. Als Folge revidierte bzw. passte sie ihrer Vorstellung über die Systemleistungen an. Es waren die Rohdaten als ein Element der *insights typ*-Visualisierung, die dazu dienten, die Störung in der Erwartungshaltung schnell aufzulösen.

Zudem gab es in den Studien Situationen, in denen Personen selbst Muster in den Rohdaten herauslesen konnten und einmal identifizierte Muster zu einem anderen Zeitpunkt in den Daten wiedererkannten (*Buggy-Beispiel* in 5.3.4). Charakteristische Muster (im engl. Patterns) im zeitlichen Verlauf von Sensordaten bilden die Basis dessen, wie physische Aktivitäten von einem maschinellen Lernsystem gedeutet werden. Die entwickelten Fähigkeiten, Muster zu erkennen und zu vergleichen, zeigt, dass die Personen ein grundlegendes Prinzip des lernenden Systems praktisch verstanden hatten.

Es war zu beobachten, dass die Vermittlung einzelner Aspekte der Systemmechanismen vor allem in Form der Rohdaten eine partielle Rekonstruktion der Medialität des Systems zuließ. Das heißt, die Möglichkeiten und Grenzen technischer Wirksamkeit wurde in bestimmter Hinsicht besser erfahren als es dies ohne den Zugang zu diesen Elementen der Systemmechanismen der Fall gewesen wäre. Eine weitreichende Erfahrung der Fülle möglichen Wirkens wird als Basis eines Vertraut-werdens mit der Technik mit dem Ziel einer gelingenden Hybridisierung verstanden.

Kapitel 6

Schluss und Ausblick

In diesem abschließenden Kapitel werden die Ergebnisse der Arbeit zusammengefasst und in Form eines Ausblicks Überlegungen angestellt, wie die Hybridisierung von Menschen durch intelligente Techniken aus normativer Perspektive zu bewerten ist.

6.1 Zusammenfassung

Ziel dieser Arbeit war es, das Verhältnis von Mensch und intelligenter Technik als eine Form der Hybridisierung begrifflich zu bestimmen und daraus Konsequenzen für eine praktische Gestaltung zu erörtern.

Die Hybridisierung des Menschen durch Technik wurde in der Arbeit hinsichtlich des Verständnisses eines Prozesses, der die Genese der funktionalen Beziehung beider Entitäten betrifft, behandelt. Mensch-Technik-Hybride sind funktionale Verbindungen von Menschen und Technik in alltäglichem Handeln und Wahrnehmen. Ausgehend von einem allgemeinen Verständnis der Hybridisierung des Menschen durch Technik wurde dieses Verständnis auf die Rolle und Bedeutung der intelligenten Techniken hin erweitert bzw. spezifiziert. Ermöglicht wurde dies in erster Linie durch den interdisziplinär motivierten Bezug auf konkrete intelligente technische Systeme. Dieser bot sich an, da so die philosophischen Überlegungen am Gegenstand geschärft und die philosophischen Konzepte gleichzeitig für die technische Forschung *greifbar* gemacht werden konnten.

Intelligente Techniken sind in dieser Arbeit als Systeme charakterisiert worden, die Muster in Daten, beispielsweise Sensordaten, im Rahmen eines festgelegten Kategoriensystems deuten und darauf aufbauend bestimmte Aktionen vollziehen können. Eine zentrale Bedingung dabei ist, dass die Deutungskompetenz der intelligenten Techniken nicht auf expliziten Festlegungen der Gestalter basiert, sondern das Resultat eines Lernprozesses ist, in dem die Techniken Trainingsdaten sowie ein bestimmtes Bereichswissen akkumulierten und auszuwerten lernten (siehe Abschnitt 2.1.1).

Das philosophische Verständnis der Hybridisierung wurde ausgehend von der Medienphilosophie der Technik und der Leibphänomenologie entwickelt.

Die Rolle der individuellen Erfahrung von Medialität für die Hybridisierung

Aus Sicht der Medientheorie sind Medien keine neutralen Vermittler, Mittel oder Instrumente, die vollständig unter menschlicher Kontrolle stehen, vielmehr werden Medien – je nach theoretischem Ansatz – eine autonome oder teilautonome Wirksamkeit zugeschrieben. Sybille Krämer (3.1.2) argumentiert, dass Medien, zu denen sie auch eine Vielzahl an Techniken zählt, eine Eigenständigkeit zugesprochen werden kann, die sich aber nur im Zusammenspiel mit dem menschlichen Handeln entfaltet. Sie beschreibt das gemeinsame Wirken von Menschen und Medien als eine „operative Einheit“, wobei es sich stets um eine Aktualisierung der Potentialität beider Entitäten handelt.¹ Krämer spricht davon, dass sich die Einheit aus Mensch und Medium, die ich als eine hybride Konstellation verstehe, dadurch auszeichnet, dass die Technik als Medium in der Nutzung verschwindet und eine Neutralisierung des Mediums im „störungsfreien Vollzug“ stattfindet.² Dabei bezieht sie sich auf andere populäre Autoren der Medientheorie wie Heider, Luhmann oder Mersch, die hinsichtlich des Verschwindens der Technik ein ähnliches Bild entwerfen.³

Der Fokus meiner Arbeit liegt auf der Entstehung dieser *operativen Einheit*, verstanden als eine in individuellen und lokalen Situationen entstehende hybride Konstellation von Mensch und Technik. Bei Krämer spielt die Entstehung, das heißt der Prozess der Hybridisierung, selbst keine Rolle, vielmehr betrachtet sie Konstellationen, die sich, so kann man ihre Ausführungen deuten, in ihrer Hybridität verfestigt haben. Dem gegenüber wurde in der Arbeit argumentiert, dass das Verschwinden des Mediums im Vollzug nicht einfach gegeben, sondern bereits das Ergebnis einer Hybridisierung von Mensch und Medium darstellt.

Diese Hybridisierung als Prozess einer Verbindung von Mensch und Technik wurde mit Bezug auf Überlegungen von Christoph Hubig (3.1.3) und Bruno Latour (3.1.4) untersucht. Latours soziologisches Interesse richtet sich auf hybride Konstellationen aus menschlichen und nicht-menschlichen Akteuren, wobei es sich bei Nicht-Menschen vor allem um Techniken handelt (präziser: materielle Dinge, die als Techniken verstanden werden bzw. im Umgang mit dem Menschen zu Techniken werden). Er spricht in diesem Zusammenhang von *Kollektiven*, in denen Techniken als (materielle) Dinge mit Menschen ein hybrides Verhältnis eingehen und so eine soziale Rolle einnehmen.⁴ Hybride aus Menschen und Technik sind

¹Siehe u. a. Krämer, 2008, Karte als technisches Medium, S. 334.

²Krämer, 2008, S. 335, S. 27.

³Krämer, 2008, S. 273f verweist auf: Heider, 1927, Luhmann, 2001, Mersch, 2003.

⁴Latour, 2006, S. 504.

nach seiner Ansicht nicht nur funktionale Verbindungen, obwohl er einige Beispiele in dieser Hinsicht diskutiert (u. a. Mensch-Waffe, Pipette-Pasteur), sondern sind grundlegender als Teil eines sozialen Netzwerks zu verstehen, in dem Eigenschaften des einen im anderen wirksam werden.⁵ In Latours Überlegungen findet sich das Konzept des „Blackboxing“, welches ein Verbinden von Menschen und Technik als einen (reversiblen) Prozess beschreibt, der die „vereinte Produktion [Agieren] von Akteuren und Artefakten völlig undurchsichtig macht“.⁶ Das „Blackboxing“ ist ein erster Aspekt bei der Konkretisierung der Hybridisierung von Menschen durch Technik.

Christoph Hubigs Ansatz ist geeignet, die im alltäglichen Gebrauch vonstattengehende Verbindung von Menschen und Technik als einen Prozess in einer bilateralen funktionalen Hinsicht zu vertiefen. Für seine Konzeption spielt die Medialität als *Gegenstand* einer individuellen Erfahrung eine entscheidende Rolle. Eine mediale Vorstellung, die vom Nutzer entwickelt wird, besteht in dem praktischen Wissen über die Möglichkeiten und Grenzen des Wahrnehmens und Handelns im Medium der Technik. Die Medialität einer Technik kann dabei partiell, in der Form eines „negative[n] Sich-Zeigen[s]“ in der Nutzung Schritt für Schritt erfahren werden.⁷ Die Medialität als Möglichkeitsraum kann aber laut Hubig nie Gegenstand einer abschließenden Vorstellung werden. Als Resultat dieses Prozesses, der als eine mediale Erfahrung bzw. Rekonstruktion verstanden wird, zeigt sich, wie die Technik im alltäglichen Umgang als Vermittelndes, als Genutztes, als Mittel zu einem Zweck für den Nutzer zu verschwinden scheint. Die Rekonstruktion des Mediums erfolgt laut Hubig infolge bestimmter Erfahrungen wie z. B. Erwartungsenttäuschungen der Nutzer.

Der Ausgangspunkt für eine Hybridisierung ist die erste Begegnung von Mensch und Technik. Lernt eine Person eine Technik *neu* kennen⁸, fasst die Person, je nach Vorwissen, das sich aus einem vorherigem Umgang mit ähnlichen Dingen speist, den Gegenstand in einer bestimmten Weise auf – sie deutet ein Element des Gegenstands etwa als Griff oder als Taste. Bestimmte, mehr oder weniger explizite Zwecke und Vorstellungen über die Umgangsmöglichkeiten mit der Technik führen bei den Techniknutzern zu der Ausbildung bestimmter Erwartungen und Handlungsweisen. Entsprechen die Ergebnisse der Nutzung nicht den Erwartungen, passt der Nutzer als Reaktion darauf seine Aktionen an oder korrigiert seine Erwartungen. Schritt für Schritt erfährt eine Person auf diese Weise die Medialität der

⁵Siehe dazu die Überlegungen von Latour, 2006, S. 504.

⁶Latour, 2006, S. 491.

⁷Hubig, 2006, S. 160.

⁸Bei Latour ein erstes *Interesse* oder bei Callon ein *interressement*, Callon, 1986.

Technik, d. h. alles das, was mit und durch eine bestimmte Technik möglich und ermöglicht wird. Unter Medialität der Technik lassen sich insofern die Möglichkeiten und Grenzen einer Technik verstehen, die im Rahmen konkreter Handlungskontexte erfahrbar werden. Im gleichen Zuge verfeinern sich eventuell auch die Fähigkeiten der Nutzer im Umgang mit der Technik. Die Erfahrung ist dabei nicht als eine explizite Thematisierung zu verstehen, sondern als ein implizites Abgleichen von Erwartung und Resultat.

Die Hybridisierung als eine Weise der Verleiblichung

Die Überlegungen zur Hybridisierung aus medientheoretischer Sicht ließen sich in der Arbeit unter Rückbezug auf die Phänomenologie des Leibes erweitern. Unter Bezug auf Merleau-Ponty kann die Hybridisierung als eine Weise der Verleiblichung charakterisiert werden. Merleau-Ponty beschreibt die Verleiblichung mithilfe des Konzepts des Körperschemas. Das Körperschema ist eine Konzeption menschlich-körperlicher Existenz, die die Tatsache beschreibt, dass der *eigene Körper* als Leib beim Handeln und Wahrnehmen nicht in seinen einzelnen Segmenten, beispielsweise in Form bestimmter Muskelgruppen angesprochen wird, sondern in seiner Gesamtheit bekannt ist und erfahren wird. Dabei gilt das Körperschema nicht nur für die Glieder des Körpers, die sich als Leibliches versammeln, sondern Merleau-Ponty versteht auch Dinge, die als Technik gebraucht werden, als Teil dieses Schemas. Ein Beispiel für eine Technik, die in das Körperschema integriert und als Erweiterung des Gliedersystems dient, ist der Stock des Blinden, der maßgeblich den Zugang zu dessen Welt mitbestimmt (3.2.3). Eine so entstandene funktionale Verbindung von Mensch und Technik als Ergebnis einer Verleiblichung kann als ein Hybrid beschrieben werden. Im Zuge einer Verleiblichung mündet der Umgang mit Technik in einer Selbstverständlichkeit, was Merleau-Ponty neben dem Blindenstock auch anhand der Schreibmaschine oder des Automobils veranschaulicht. Demzufolge ist die Verleiblichung eine Konzeption, die dazu geeignet ist, die Hybridisierung von Mensch und Technik zu charakterisieren.

In Abschnitt 3.2.4 wurde Heideggers phänomenologische Technikphilosophie hinsichtlich der Hybridisierung von Menschen durch Technik diskutiert. Bei Heidegger ist das technische Handeln ein basales Element menschlichen Seins in der Welt. Der Terminus *In-der-Welt-sein* wird in Bezug auf das *Wohnen* in der Welt verstanden, welches maßgeblich durch die unhintergehbare körperliche Präsenz (als Leib des Menschen) und dessen

vorreflexives, gewohntes Agieren in der Welt bestimmt wird.⁹ Das leibliche Sein des Menschen, wenn auch nicht explizit von Heidegger als Leib benannt, schließt das Verwenden von Dingen im Sinne eines technischen Handelns nicht nur mit ein, sondern ihm kommt eine grundlegende Bedeutung zu. Dinge, die in diesem Sinne hinsichtlich ihrer Verwendung als Technik erscheinen, bezeichnet er als *Zeug* und das vorreflexive Verhältnis der Technik zum Menschen als *Zuhandenheit*.

Die Überlegungen Heideggers zeigen eine Verwandtschaft zu dem *Ansatz* der Medienphilosophie, demnach die Technik im störungsfreien Gebrauch verschwindet bzw. darin verborgen ist. Die Frage nach der Genese des Verhältnisses von Mensch und Technik, in der die Technik als zuhandene gebraucht und nur so überhaupt verstanden wird bzw. im engeren Sinne Technik ist, spielt bei Heidegger keine Rolle, da er dieses Verhältnis als das ursprüngliche tituliert. Vielmehr ist für ihn der umgekehrte Weg von Interesse, wenn er eine Erklärung dafür sucht, wie sich die Technik als „nächstzuhandene[s] Seiende [...] als unverwendbar, als nicht zugerichtet für seine bestimmte Verwendung“, herausstellt. Dann erscheint sie als Vorhandene und ein gebrauchender Umgang ist nicht mehr möglich.

Die Möglichkeiten und Grenzen einer Mensch-AI-Hybridisierung

Die zentrale Aufgabe dieser Arbeit bestand darin, ein Verständnis für die Hybridisierung von Menschen durch intelligente Technik zu erarbeiten. Es schloss sich die Frage an, welche Konsequenzen sich daraus für die technische Gestaltungspraxis ergeben.

Dabei wurde gezeigt, dass eine mediale Erfahrung gar nicht oder nur schwer möglich ist, wenn das technische System intelligentes Vermögen besitzt (3.3). Demnach erscheint aus Sicht eines Nutzers ein intelligentes System auch nach fortgesetztem Lernen eines Umgangs unklar hinsichtlich der Möglichkeiten und Grenzen des Wirkens. Die Hybridisierung ist, wie weiter oben dargestellt, durch die Möglichkeiten einer partiellen Erfahrung der Medialität festgelegt und damit hinsichtlich intelligenter Technik problematisch. Wie in Abschnitt 4.2 dargestellt, zeigten empirische Untersuchungen zum Nutzerverhalten die symptomatische Problemlage im Umgang von Menschen und intelligenter Technik auf.

Um die Hybridisierung hinsichtlich des Verhältnisses von Mensch und

⁹Das Wohnen steht nach Heidegger in terminologischer Verwandtschaft mit Gewohnheiten.

intelligenter Technik präzise zu bestimmen und so die Probleme zu analysieren, wurden die Überlegungen zum Konzept der Spur verwendet. Sybille Krämer verwendet diesen Begriff in ihrem medientheoretischen Ansatz (3.1.2) wie auch Christoph Hubig, auf dessen Arbeit zur Medialität der Technik in Abschnitt 3.1.3 Bezug genommen wurde. Die Spur ist ein Konzept, das dabei helfen kann, den Übergang von der Erfahrung der Nutzer während des Vollzugs hin zu einer medialen Vorstellung zu verstehen. Die Idee ist, dass jede Nutzung eines technischen Systems Spuren des Mediums *hinterlässt*, das heißt Effekte der Nutzung, die auf die mediale Verfasstheit der Technik verweisen. Die Medialität der Technik wird in der Nutzung Schritt für Schritt, wie oben beschrieben, durch den Abgleich von Erwartungen und den Beobachtungen der Ergebnisse der Nutzung erschlossen. Eine mediale Vorstellung besteht in dem praktischen, mithin leiblichem Wissen über die Möglichkeiten und Grenzen des Wahrnehmens und Handelns im Medium der Technik. Im Handeln mit Technik sind die Spuren die *Hinterlassenschaften* realer Nutzung (Spuren von) und verweisen damit auf Mögliches, das heißt auf die Bedingungen möglicher zukünftiger Nutzung (Spuren für).

Wie in Abschnitt 3.1.3 dargestellt, werden Spuren üblicherweise als Hinterlassenschaften eines realen Ereignisses verstanden, sie verweisen auf Wirkliches, das zeitlich zurückliegt. Die Spuren, wie sie hinsichtlich von Medien verstanden werden (siehe Abschnitt 3.1.3), sind zwar anteilig auch Hinterlassenschaften als Effekte wirklicher Nutzung, verweisen aber nicht nur auf Wirkliches, sondern auch auf Mögliches, das heißt, auf die Medialität der Technik als einem Raum möglichen Wirkens. Demnach sind die Spuren nicht die Effekte einer Nutzung alleine, beispielsweise die Schramme in der Wand als Ergebnis eines fehlgeschlagenen Hammergebrauchs oder die Wärme des Computers als Nebeneffekt der Nutzung. Vielmehr verweisen die Effekte nur dann auf die Medialität und sind damit Spuren, wenn sie vom Nutzer hinsichtlich seiner Erwartung und im Kontext der Entstehung der Effekte, das heißt z. B. hinsichtlich des Wissens über bzw. der Beobachtung von zugrundeliegenden (u. a. physischen oder elektronischen) Wirkzusammenhängen, interpretiert werden können.

Mit Blick auf intelligente Techniken wurde argumentiert, dass ein Spurenlernen problematisch wird. Die Spuren als Voraussetzung für eine Rekonstruktion von Medialität werden nicht mehr in der Nutzung *hinterlassen*. Dieser *Verlust von Spuren* entsteht dadurch, dass die Differenzenerfahrung zwischen den erwarteten Effekten (Ergebnissen, Zielen) und den realen Effekten nicht mehr im Kontext ihrer Entstehung erscheint. Das heißt, der Nutzer erfährt eine Abweichung von dem, was er zu erreichen hoffte oder beobachtet zusätzliche, unerwartete Effekte, wobei er keine Anhaltspunkte dazu findet, warum dies der Fall ist. Der Nutzer kann nicht erkennen, was

die Differenz verursacht hat.

Mit Bezug auf nicht-intelligente Techniken aus dem Bereich der Elektronik oder der Computertechnik stellte sich die Frage, ob nicht in gleicher Weise gilt, dass Spuren verlustig gehen, da die Hauselektronik hinter der Wand oder die Platinen im Gehäuse nicht für den Nutzer zugänglich sind (3.1.5). Bei diesen Techniken gilt: Wenn der Umgang damit Effekte zeitigt, die direkt auf das eigene Verhalten zurückzuführen sind, wie z. B. das Drücken einer Taste auf der Computertastatur zur Darstellung des Buchstabens auf dem Bildschirm führt oder das Drücken des Lichtschalters zum An- oder Ausschalten des Lichts, so lassen sich die Zusammenhänge im iterativen Prozess *schnell* erschließen. Das Licht und der Bildschirminhalt eignen sich als Effekte, um sie im Kontext ihrer Entstehung (Bedienung der Taste) als Spuren zu lesen und um damit auf die Medialität als Möglichkeiten und Grenzen der Nutzbarkeit zu verweisen.

Kandidaten zur Vermittlung von Systemmechanismen

Die bisherigen Überlegungen zeigten, dass die Möglichkeit, die Medialität auf Basis von Spuren zu rekonstruieren, zentral für den Prozess der Hybridisierung ist. Hinsichtlich des Designs von intelligenten Systemen wurde untersucht, wie auf der gestalterischen Ebene Bedingungen geschaffen werden können, bei denen ein Lesen von Spuren möglich ist, ohne auf intelligente Funktionalitäten zu verzichten.

Es wurde angenommen, dass dazu die Mechanismen, denen gemäß ein intelligentes System agiert bzw. prozessiert, in der Nutzung vermittelt werden müssen. Mechanismen sind in einem allgemeinen Verständnis als die Funktionsweisen zu verstehen, nach denen der maschinelle Algorithmus, meist ein Klassifikator, als Grundlage des intelligenten Systems eine Situation deutet und darauf basierend bestimmte Effekte erzeugt. Je nach Lernalgorithmus, der zum Einsatz kommt, können die Systemmechanismen sehr unterschiedlich aussehen und eignen sich auf bestimmte Weise dazu, sie in ihrem Wirken darstellbar zu machen.

Das Feld maschinellen Lernens als ein Teilbereich der Informatik befasst sich mit der Veranschaulichung algorithmischer Funktionalität im Rahmen einer Diskussion zur Verstehbarkeit von Lernalgorithmen. Dabei wird unter der Verstehbarkeit meist eine Nachvollziehbarkeit verstanden, das heißt die Möglichkeit, zu verstehen, wie ein Ergebnis des Algorithmus (Klassifikationsergebnis) auf Basis der Eingangsgrößen zu Stande kommt.

In der vorliegenden Arbeit wurden die Überlegungen zur Verstehbarkeit als Grundlage genommen, maschinelle Lernansätze hinsichtlich der Möglichkeit der Vermittlung von Systemmechanismen zu vergleichen. Die Möglichkeiten und Grenzen der Verstehbarkeit wurden anhand von zwei

sich grundlegend unterscheidenden Lernansätzen, den Entscheidungsbäumen und den neuronalen Netzen, dargestellt (siehe die Abschnitte 4.1.2 und 4.1.3). Entscheidungsbäume gelten als einfach verstehbar, da die Ergebnisse auf Basis einer hierarchischen Struktur entstehen, was ein Nachvollziehen leicht möglich macht. Einzelne Instanzen einer Klassifikation, beispielsweise im Rahmen einer Situationsdeutung oder Optionsentscheidung der Technik, können anhand der Merkmale in u. a. Sensordaten nachvollzogen werden. Bei neuronalen Netzen ergeben sich die Ergebnisse auf Basis einer komplexen Struktur von Neuronen aus den Eingangswerten. Die Struktur ist komplex, weil die Neuronen in mehreren Schichten angeordnet sind, wobei jedes Neuron das Gesamtergebnis beeinflussen kann. Einzelne, das Ergebnis beeinflussende Schwellenwerte und relevante Merkmale in den Sensorwerten können nicht oder nur schwer herausgelesen werden.

Mit Blick auf Systeme maschinellen Lernens, die in intelligenten Systemen zum Einsatz kommen, zeigt sich, dass aktuelle leistungsstarke Algorithmen keine Lernansätze verwenden, die im besprochenen Sinne leicht verständlich sind. Beispielsweise sind *Random-Forest*-Ansätze Erweiterungen der Entscheidungsbäume, wobei bei diesem Ansatz eine ganze Reihe von Entscheidungsbäumen mit unterschiedlichen Konfigurationen zusammenwirken und somit Ergebnisse deutlich schwerer nachvollziehbar sind.

Auf Basis der Diskussion zur Verstehbarkeit von lernenden Algorithmen wurde nach Typen von Elementen aktueller Algorithmen gesucht, die für den Nutzer vermittelt werden können, um das interne Prozessieren intelligenter Technik zugänglich zu machen. Diese Elemente bezeichne ich als Kandidaten, die am Ende von Kapitel 4 in einem Kandidaten-Framework zusammengefasst wurden (4.3).

Die Kandidaten wurden nicht dahingehend ausgewählt, den Anspruch zu erfüllen, den Algorithmus vollständig zu repräsentieren, was schlicht nicht möglich wäre, sondern sie wurden als Indikatoren verstanden, die zunächst nur auf einzelne Instanzen des Wirkens der Mechanismen verweisen. Die Indikatoren lassen die Systemmechanismen nur im Kontext der konkreten Anwendung, der Historie vorheriger Anwendungen und erst nach mehreren Nutzungsinstanzen erscheinen. Im Rahmen dieses Prozesses des Schritt-für-Schritt-Erschließens werden die Kandidaten als Spuren im Zuge bestimmter Differenzerfahrungen wirksam und somit eine teilweise Rekonstruktion der Medialität möglich.

In dieser Arbeit wurden vier Kandidatentypen ausgearbeitet. Die vier Typen bilden ein Suchraster, das, je nach Algorithmus und Anwendung, unterschiedlich genutzt werden kann. Der erste Typ von Kandidaten sammelt Elemente des Algorithmus, die die Chronologie der maschinellen Klassifikation betreffen, das heißt die Zwischenschritte, die letztendlich zu einem Klassifikationsergebnis führen. Das sind beispielsweise die *Votings*

parallel laufender Klassifikatoren, wie beim *Random-Forest*-Ansatz und die Schwellenwerte, die die Anzahl von Stimmen (Votes) als ausreichend für eine Klassifikation auszeichnen. Der zweite Typ betrifft die Quellen der maschinellen Deutung, das heißt die Messdaten bzw. bestimmte Aspekte der Messdaten sowie sonstige Grundlagen, beispielsweise von Wissensdatenbanken, die die Grundlage für eine Klassifikation bilden. Die Quellen repräsentieren den Wahrnehmungs- und Erkenntnisraum der Technik, welcher dem Nutzer durch Abgleich mit dessen eigener Erfassung den Kontext der Entstehung von Ergebnissen erfahrbar macht. Der dritte Typ versammelt Kandidaten, die als Metadaten gelten, das heißt, er umfasst alle Information über die Beschaffenheit des Systems, unter anderem Daten zur allgemeinen Leistungsfähigkeit des Systems, beispielsweise die durchschnittliche Genauigkeit, mit der bestimmte Situationen klassifiziert werden können. Der vierte Typ von Kandidaten bezieht sich auf das allgemeine Bereichswissen, das als Grundlage für eine Modellierung des Algorithmus dienen kann, beispielsweise das Wissen über die Physiologie von Menschen als Grundlage zur maschinellen Deutung von menschlicher Aktivität.

Abschluss des interdisziplinären Bogens: Nutzerstudien und deren philosophische Reflexion

Die im praktischen Teil der Arbeit dargestellten Untersuchungen bezogen sich auf Anwendungen im Feld der Self-Tracking-Techniken. Die Technik der Aktivitätserkennung bildet die Grundlage für die computertechnisch vermittelte Selbstbeobachtung. Die Technik zeichnet sich dadurch aus, dass Sensorsysteme zum Einsatz kommen, die der Mensch am Körper trägt bzw. die sich in seiner Umgebung befinden, mit deren Hilfe menschliche Aktivitäten maschinell gedeutet werden können. Die Nutzung lernender Algorithmen als Grundlage der Deutung qualifiziert die Systeme als intelligente Techniken.

In zwei Studien wurde in erster Linie untersucht, inwieweit die Vermittlung von Systemmechanismen, die sich an dem Kandidaten-Framework orientieren, für einen Umgang mit intelligenten Self-Tracking-Systemen für die Anwendung in verschiedenen Bereichen eignet. Des Weiteren wurden die Ergebnisse aus Sicht der theoretischen Überlegungen zu den Möglichkeiten einer Hybridisierung reflektiert.

Das technische Setup der Studien basierte auf einem System mit dem Namen *Hedgehog*, welches maßgeblich in der *Embedded Sensing Systems Group* an der Technischen Universität Darmstadt entwickelt wurde. Dabei handelte es sich um ein Sensorsystem, das am Handgelenk getragen werden kann und intern hauptsächlich aus einem dreiachsigen Beschleunigungssensor besteht (5.1.1). Der *Hedgehog* wurde vor allem für ein *long-term*

activity tracking entwickelt, das heißt, er kann eine ununterbrochene Messung der Bewegungsdaten in einem Zeitraum von über zwei Wochen realisieren. Der Beschleunigungssensor bildet die Grundlage für die Deutung der Aktivität. Die Bewegungen des Arms, die in den Sensordaten repräsentiert sind, dienen als Indikatoren für eine Vielzahl von täglichen Aktivitäten, u. a. Fortbewegungsarten wie Laufen, Treppensteigen und Radfahren sowie eine Reihe von Sport- und Freizeitaktivitäten.

In der ersten Studie (5.2) wurden verschiedene Visualisierungstypen verglichen, die sich hinsichtlich der Darstellung von Aspekten von Systemmechanismen unterschieden, wobei es bei den Aspekten um Zwischenschritte im Prozess maschineller Klassifikation mit dem Namen *dense motif discovery* handelte. Das Studiendesign orientierte sich an einem Anwendungsszenario, das die Möglichkeit eines Einsatzes von Techniken der Aktivitätserkennung in der psychologischen Therapie betraf. Das Szenario beeinflusste das Studiendesign u. a. dahingehend, dass das System der Aktivitätserkennung nur einzelne Schlüsselaktivitäten erlernte, da eine verlässliche maschinelle Deutung einzelner Aktivitäten in einer Therapiesituation für Psychologen/Psychiater von Interesse sein könnte. Die Studie zeigte vor allem zwei Dinge: Zum einen, wie das Kandidaten-Framework in einem realistischen Anwendungsszenario systemisch umgesetzt werden kann und zum anderen, dass Visualisierungen, die bestimmte Aspekte des Systemmechanismus darstellten, von den Teilnehmenden als nützlicher eingeschätzt wurden und dem Dargestellten mehr *vertraut* wurde (Im Detail: 5.2.5).

Für die zweite Studie (5.3) wurde ein Studiendesign gewählt, bei dem die Aktivitäten der Teilnehmenden in einer einwöchigen Phase mit Hilfe des Sensorsystems *Hedgehog* aufgezeichnet wurden. Die Sensordaten wurden in der Folge hinsichtlich eines definierten Satzes von Aktivitäten wie Freizeit- und Sportaktivitäten sowie hinsichtlich alltäglicher Beweglichkeitsaktivitäten gedeutet (Bei dem System handelte es sich um einen *Wizard-of-Oz*-Ansatz, der in Abschnitt 5.3.3 ausführlich beschrieben wurde) und den Teilnehmenden anhand von zwei Visualisierungstypen präsentiert wurde. Bei den Visualisierungstypen handelte es sich zum einen um eine *einfache* Darstellungsweise, die als *simple type* bezeichnet wurde, zum anderen um eine, die auf dem Kandidaten-Framework basierte und als *insights type* bezeichnet wurde. In der Studie wurden die Möglichkeiten einer technik-vermittelten Reflexion alltäglichen Verhaltens untersucht. Dabei handelte es sich um eine Forschungsfrage, die im Ubiquitous Computing aktuelle Relevanz besitzt und auch im Rahmen der disziplinären Konzeptionen von Selbstreflexion besprochen wurde (Es wurde definiert, dass es sich u. a. um eine *gelungene Selbstreflexion* handelte, wenn durch

den Technikgebrauch ein systematisches *Nachforschen* hinsichtlich des eigenen Verhaltens möglich war bzw. wenn ein *Perspektivenwechsel* angeregt wurde).

Die Aufgabe der Teilnehmenden bestand darin, die maschinell gedeutete Aktivitätshistorie der vergangenen Woche in einer einstündigen Sitzung anhand der Visualisierungen auf einem Computerbildschirm anzusehen und auf Basis dessen versuchen, verschiedene Begebenheiten bzw. Ereignisse der Woche, private und berufliche, ins Gedächtnis zu rufen. Die Sitzung beinhaltete zudem eine anschließende Beantwortung eines Fragebogens. Die Teilnehmenden hatten die Vorgabe, die Erfahrungen im Umgang mit den Visualisierungen zu kommentieren, was in einer Audioaufnahme aufgezeichnet wurde. Die transkribierten Aufnahmen bildeten die Grundlage für die qualitative Auswertung der Studie.

Die Auswertung der Fragebögen zeigte, dass die Teilnehmenden den *insights type* besser bewerteten als den *simple type*, sowohl hinsichtlich der allgemeinen Nützlichkeit als auch des *Vertrauens* in die Systemleistungen. Die Elemente des *insights type*, die sie als besonders nützlich bzw. vertrauensfördernd erachteten, waren die sogenannten Rohdaten. Die Rohdaten sind die der Deutung zugrundeliegenden Sensordaten, die bezüglich des Kandidaten-Frameworks zu dem Typus *Quellen* gehören. In der Visualisierung des *insights type* wurden die Rohdaten in einer mehrfarbigen Darstellung vermittelt (Im Detail: 5.3.4).

Die qualitative Auswertung der Audioaufnahmen bildete die Grundlage, um die Möglichkeiten einer Selbstreflexion in Abhängigkeit von den Visualisierungstypen zu bewerten. Die qualitativen Ergebnisse zeigten, dass eine Selbstreflexion vor allem in der Form eines systematischen Nachforschens und der Möglichkeit eines Perspektivenwechsels mit dem *insights type* besser möglich war. Dabei spielten vor allem die Rohdatenvisualisierung des *insights types* eine Rolle: Diese bot den Teilnehmenden die Möglichkeit, die *Deutungen* des Systems zu hinterfragen und diese gegebenenfalls in ihrem Sinne weiter auszudifferenzieren.

Eine ergänzende Auswertung der Studienergebnisse wurde in Abschnitt 5.4 vorgenommen, indem die Ergebnisse hinsichtlich der Möglichkeit einer Hybridisierung bewertet wurden. Dabei wurde die Frage gestellt, ob es Beobachtungen gab, die den Schluss zulassen, dass der Verlust von Spuren erfolgreich kompensiert wurde und durch die Vermittlung von Aspekten der Systemmechanismen in Form der Kandidaten eine mediale Rekonstruktion und somit die Bedingungen für die Möglichkeit einer Hybridisierung erfüllt waren. Die Beobachtungen, die im Rahmen einer Bewertung der Möglichkeiten der Selbstreflexion gemacht wurden, zeigten eine ganze Reihe von Fällen, in denen im Umgang Erwartungen an die

Leistungen des Systems enttäuscht wurden oder es zu Überraschungserlebnissen kam. In vielen Fällen konnten Nutzer beispielsweise die Ursachen nicht erfüllter Erwartungen identifizieren, indem sie sich auf die Quelldaten (Rohdaten) bzw. Informationen zu der Qualität der Systemdeutung, stützten.

In der beschriebenen Studie wurde gezeigt, dass es Personen aufgrund der Vermittlung von Aspekten der Systemmechanismen in einer ersten Begegnung möglich war, die Technik in medialer Hinsicht zu erfahren. Die Erklärungen und Beschreibungen der Teilnehmenden bildeten die Grundlage der qualitativen Bewertung. Das schränkt die Aussagekraft hinsichtlich der medialen Erfahrung insofern ein, als dass hier nur bewusst Verbalisiertes und somit in gewisser Hinsicht reflektiertes Verhalten ausgewertet werden konnte. Wie im theoretischen Teil der Arbeit herausgearbeitet, kommt dem impliziten Sich-Verhalten beim Lernen eines Umgangs mit Technik eine zentrale Rolle zu, doch ist gerade das erste *Kennenlernen* durch bewusstes Agieren und Reflektieren geprägt. Ob eine erste Begegnung zu einem hybriden Verhältnis führt, hängt von vielen weiteren Faktoren ab, die in den Studien nicht abgebildet werden konnten. Doch war der Umfang der Studie ausreichend, um zu zeigen, dass die Medialität des Systems von den Nutzern teilweise erschlossen wurde und somit die Annahme getroffen werden kann, dass die Grundlage für eine gelingende Hybridisierung gelegt wurde (3.3).

6.2 Ausblick: Die Hybridisierung aus einer normativen Perspektive

In diesem Ausblick wird die Frage behandelt, inwiefern die Hybridisierung eine Verbindung von Menschen und intelligenter Technik darstellt, die aus moralischer Sicht erstrebenswert ist.¹⁰

Eine Richtung, die Hybridisierung aus einer moralischen Perspektive zu bewerten, kann anhand des Konzepts der *Transparenz* erfolgen. Das normative Argument lautet, dass die Vermittlung von Systemmechanismen eine gewisse *Transparenz* herstellt, was moralisch erstrebenswert ist und somit in der Gestaltung berücksichtigt werden sollte. *Transparenz* wird hinsichtlich vieler Aspekte sowohl in wissenschaftlichen Arbeiten als auch in

¹⁰In der Forschung zur HCI gibt es eine ganze Reihe an Ansätzen, in denen Methoden entwickelt werden, die die Technikgestaltung nach moralischen Werten organisiert. Beispiele sind das „value-sensitive design“ sowie das „critical design“, siehe u. a. Bardzell, 2011, Friedman, Kahn Jr. und Borning, 2013.

der politischen Diskussion als ein normatives Ziel genannt, da sie in einer ersten Charakterisierung als eine Offenlegung von Prozesses und Inhalten für alle ist, einen gerechten Zugang zu Wissen gewährleistet und eine Grundlage für autonome Entscheidungen bietet. Ein Mangel an *Transparenz* von technischen Entscheidungssystemen ist eines der zentralen Bedenken, die in Hinblick auf die Entwicklung intelligenter Systeme und ihrer zunehmenden Wirksamkeit in vielen Lebensbereichen hervorgebracht wird.¹¹

Der Begriff der *Transparenz* wird metaphorisch – ursprünglich aus dem Bereich der Optik – gebraucht und wurde in keiner der im Absatz zuvor genannten Arbeiten klar definiert. Die Herausforderungen einer Konzeption von *Transparenz* liegt darin, dass es zu klären ist, was *durchsichtig* gemacht werden soll und was das andere ist, was dadurch erscheinen soll. Im Folgenden wird die *Transparenz* hinsichtlich möglicher Bedenken, die sich aus einem Mangel an *Transparenz* ergeben, anhand von Beispielen verschiedener Mensch-Technik-Konstellationen weiter präzisiert.

In einer Nutzer-System-Konstellation, wie sie für diese Arbeit zentral ist, stehen Personen in einem Verhältnis zu intelligenten Systemen, wenn sie mit ihrem intelligenten Auto ein Fahrziel erreichen wollen, sie am Computer mit einer Suchmaschine bestimmte Informationen suchen oder mit personalisierter Werbung konfrontiert sind. Neben diesen Konstellationen gibt es auch Beispiele, in denen (intelligente) Algorithmen Aufgaben in sozialen Institutionen übernehmen. Beispielsweise werden Algorithmen eingesetzt, um in den USA die Vergabe von Greencards zu regeln oder über die Gewährung von Bankkrediten und deren Höhe zu entscheiden. Einige Publikationen haben sich in diesem Zusammenhang mit der Fairness solcher autonomen Entscheidungssysteme befasst und fordern in diesem Zuge eine *Transparenz* dieser Systeme.¹² Auch der weltgrößte Verband der Ingenieure (IEEE) hat sich mit den ethischen Herausforderungen intelligenter und autonomer Systeme befasst und *Transparenz* und Nachvollziehbarkeit als eine wichtige Forderung hervorgehoben.¹³

Hinsichtlich des Einsatzes algorithmischer Entscheidungen in sozialen Institutionen sind es Fragen sozialer Gerechtigkeit, die die Forderung nach *Transparenz* rechtfertigen.¹⁴ Aus Sicht eines individuellen Verhältnisses von Menschen und intelligenten Systemen kann die Forderung nach

¹¹Siehe unter anderem hier: Dwork et al., 2012, IEEE, 2016, Charisi et al., 2017.

¹²Dwork et al., 2012.

¹³Eine Veröffentlichung des „Institute of Electrical and Electronics Engineers“, besser bekannt als IEEE. Siehe dazu: IEEE, 2016, erster Entwurf. Weitere Arbeiten, die sich unter dem Begriff „accountable algorithms“ versammeln, befassen sich mit diesem Thema. Siehe dazu: Goodman, 2016, Kroll et al., 2016.

¹⁴Dwork et al., 2012.

mehr *Transparenz* in Hinblick auf den ethischen Anspruch einer Gewährleistung menschlicher Autonomie gerechtfertigt werden. Das Autonomie-Argument besagt, dass Menschen jederzeit autonom im Sinne einer Selbstgesetzgebung, das heißt auf Basis von sich selbst gegebenen Normen handeln können sollten.¹⁵ Demnach sollten aus Sicht von System-Designern Bedingungen geschaffen werden, die ein autonomes Handeln im Umgang mit der Technik gewähren. Hinsichtlich intelligenter Systeme kann es sein, dass sich Personen nicht in ein Verhältnis dazu setzen können, in welcher Beziehung sie zu dem System stehen, weil die undurchsichtigen Mechanismen beispielsweise Handlungsanteile *verschleiern*. Zentral für ein autonomes Agieren mit Technik ist es, sich ablehnend und anerkennend zu einer Technik bzw. einem technischen System, das einen umgibt, verhalten zu können. Hubig spricht hinsichtlich dieses bewussten Eingehens einer Beziehung von der Kontrolle, die potentiell erhalten bleiben sollte. Es geht hier nicht um eine ständige Kontrolle aller Prozesse, sondern um ein Potential oder ein Vermögen als „Kontrollkompetenz“.¹⁶ Wiegerling argumentiert in einer ähnlich gelagerten Einschätzung mit Bezug auf intelligente Systeme, dass es Möglichkeiten geben sollte, „aus der Systemunterstützung bzw. Systembegleitung aussteigen zu können“.¹⁷ Er weist hinsichtlich einer normativen Gestaltungspraxis auf eine *Transparenz* im Sinne einer Aufklärung über Funktionsweisen hin:

„Es wird in intelligenten Handlungsumgebungen darauf ankommen, dem Menschen Handlungsoptionen zu vermitteln, ihm Eingriffsmöglichkeiten in ein autonom agierendes System sichtbar zu machen und ihn soweit wie möglich über die Funktionsweisen des Systems aufzuklären.“¹⁸

Wenn man sich der *Transparenz* aus Sicht des Autonomiearguments nähert, wird das Bild klarer und es kann begründet werden, welche Aspekte einer Technik sichtbar gemacht werden sollen. Die Forderung nach *Transparenz* ist in diesem Zusammenhang erfüllt, wenn die Eingriffsmöglichkeiten disponibel bleiben. Im Fall einer hybriden Verbindung von Mensch und Technik sind Eingriffe möglich, da die Möglichkeiten eines damit Umgehens, im Prozess der Entstehung der Verbindung zum Gegenstand einer praktischen Vorstellung wurden. Im Prozess einer Hybridisierung wird nutzerseitig die Funktionsweise als Möglichkeit eines zweckgebundenen

¹⁵Hubig, 2015, S. 131f, unter Bezugnahme auf Kants Autonomiebegriff in der „Grundlegung der Metaphysik der Sitten“ Kant, 2012.

¹⁶Hubig, 2015, S. 134.

¹⁷Wiegerling, 2011, S. 203.

¹⁸Wiegerling, 2011, S. 72.

Umgangs praktisch bzw. leiblich *verstanden*. Eine so entstandene Verbindung ist im Fall einer *Krise*, das heißt einer Störung der gewöhnlichen Nutzung, im Prinzip reversibel.¹⁹ Eine Vorstellung über die für den Nutzer relevanten Funktionsweisen hinsichtlich der unterschiedlichen Handlungsoptionen, die auch Ausstiegsoptionen sein können, wie sie Wiegerling fordert, haben sich in der Erfahrung der Medialität der Technik während des Prozesses der Hybridisierung entwickelt.

Ein alternatives Mensch-Technik-Verhältnis beschreibt beispielsweise ein Agieren in einer technischen Infrastruktur, auf die man angewiesen ist, die aber in der Gesamtheit ihres potentiellen Wirkens unklar bleibt. Man kann zwar einen Umgang damit finden, der auch zu einer Gewohnheit werden kann, der aber nicht selbstverständlich oder vertraut ist. Don Ihde, dessen Arbeit zu Mensch-Technik-Verhältnissen in Abschnitt 3.2.5 diskutiert wurde, beschreibt *Extremfälle* solcher Typen von Verhältnissen, u. a. *background relations* oder die Verschmelzung als *total embodiment*, bei denen nicht mehr von einem Verhältnis zur Technik gesprochen werden kann, weil die Grenze zwischen den Entitäten nicht mehr als solche erfahrbar ist.

Eine Designstrategie, die eine Hybridisierung mithilfe der Vermittlung von Systemmechanismen gewährleistet, ist ethisch erstrebenswert, da sie die Handlungsautonomie von Nutzern berücksichtigt. Eine weitere Ausarbeitung des ethischen Arguments ist Aufgabe einer zukünftigen Arbeit.

6.3 Fazit

In der vorliegenden Arbeit wurde das Verhältnis von Mensch und intelligenter Technik aus einer interdisziplinären philosophisch-informatischen Perspektive betrachtet. Die Arbeit beinhaltete zum einen eine Auseinandersetzung mit philosophischen Konzepten, die hinsichtlich neuer Beziehungsgeflechte zwischen Mensch und Technik erweitert und vertieft wurden. Zum anderen wurden neue Ideen für eine technische Gestaltungspraxis entwickelt, die die aktuelle Landschaft erweitern und Herausforderungen begegnen. Die interdisziplinäre Ausrichtung bedurfte einer starken Fokussierung, sowohl auf der Ebene der Theorie als auch hinsichtlich der praktischen Umsetzung, so dass bestimmte philosophische und informatische Diskussionen und Forschungsfragen, die eng mit dem Thema in Verbindung stehen, nicht behandelt werden konnten. Zudem beschränkten sich die praktischen Konsequenzen in ihrer Umsetzung auf nur einen Anwendungsbereich intelligenter Techniken. Eine Erweiterung dessen könnte wiederum Aufgabe einer zukünftigen Arbeit sein.

¹⁹Hubig, 2007, S. 180.

Literatur

- Abowd, Gregory D. und Elizabeth D. Mynatt (2000). „Charting Past, Present, and Future Research in Ubiquitous Computing“. In: *ACM Transactions on Computer-Human Interaction* 7.1, S. 29–58.
- Abowd, Gregory D. et al. (1999). „Towards a Better Understanding of Context and Context-Awareness“. In: *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing*. Heidelberg: Springer Verlag, S. 304–307.
- Alpaydin, Ethem (2008). *Maschinelles Lernen*. München: Oldenbourg Verlag.
- Amft, Oliver, Holger Junker und Gerhard Troster (2005). „Detection of Eating and Drinking Arm Gestures Using Inertial Body-worn Sensors“. In: *Proceedings of the 2005 International Symposium on Wearable Computers*. Los Alamitos: IEEE, S. 160–163.
- Aristoteles (1983). *Über die Seele*. 6. Auflage, Übersetzt von Willy Theiler. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Bardzell, Jeffrey (2011). „Interaction Criticism: An Introduction to the Practice“. In: *Interacting with Computers* 23.6, S. 604–621.
- Baumer, Eric P.S. (2015). „Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection“. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. New York: ACM, S. 585–594.
- Berlin, Eugen (2014). „Early Abstraction of Inertial Sensor Data for Long-term Deployments“. Diss. Darmstadt University of Technology, Computer Science.
- Berlin, Eugen und Kristof Van Laerhoven (2012). „Trainspotting: Combining Fast Features to Enable Detection on Resource-constrained Sensing Devices“. In: *Ninth International Conference on Networked Sensing*. Los Alamitos: IEEE, S. 1–8.
- Berlin, Eugen und Kristof Van Laerhoven (2012). „Detecting Leisure Activities with Dense Motif Discovery“. In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. New York: ACM, S. 250–259.
- Bonfiglioli, Rudi und Federico Nanni (2016). „From Close to Distant and Back: How to Read with the Help of Machines“. In: *History and Philosophy of Computing: Third International Conference*. Hrsg. von Fabio Gadducci und Mirko Tamosanis. Cham: Springer International Publishing, S. 87–100.

- Brey, Philip (2001). „Hubert Dreyfus: Humans Versus Computers“. In: *American Philosophy of Technology: The Empirical Turn*. Bloomington: Indiana University Press, S. 37–63.
- Callon, Michel (1986). „Some Elements of a Sociology of Translation: Do Mestication of the Scallops and the Fishermen of St Brieuc Bay“. In: *Power, Action, and Belief: A New Sociology of Knowledge?* Hrsg. von John Law. London: Routledge, S. 196–223.
- Chalmers, Matthew und Ian Maccoll (2003). „Seamful and Seamless Design in Ubiquitous Computing“. In: *Proceedings of Workshop At the Crossroads: The Interaction of HCI and Systems Issues in UbiComp 2003*.
- Chang, Keng-hao, Mike Y. Chen und John Canny (2007). „Tracking Free-weight Exercises“. In: *Proceedings of the 9th International Conference on Ubiquitous Computing*. Hrsg. von John Krumm et al. Berlin: Springer Verlag, S. 19–37.
- Charisi, Vicky et al. (2017). *Towards Moral Autonomous Systems*. Website. Online erhältlich unter <http://arxiv.org/abs/1703.04741>; abgerufen am 11. Juni 2018.
- Chiauzzi, Emil, Carlos Rodarte und Pronabesh Dasmahapatra (2015). „Patient-centered Activity Monitoring in the Self-management of Chronic Health Conditions“. In: *BMC Medicine* 13.1, S. 77.
- Choe, Eun Kyoung, Bongshin Lee und M.C. Schraefel (2015). „Characterizing Visualization Insights from Quantified Selfers' Personal Data Presentations“. In: *Computer Graphics and Applications, IEEE* 35.4, S. 28–37.
- Choe, Eun Kyoung et al. (2014). „Understanding Quantified-selfers' Practices in Collecting and Exploring Personal Data“. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM, S. 1143–1152.
- Cruz, Luis et al. (2015). „A Wearable and Mobile Intervention Delivery System for Individuals with Panic Disorder“. In: *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. New York: ACM, S. 175–182.
- Cuttone, Andrea, Michael Kai Petersen und Jakob Eg Larsen (2014). „Four Data Visualization Heuristics to Facilitate Reflection in Personal Informatics“. English. In: *Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*. Bd. 8516. Springer International Publishing, S. 541–552.
- Dennett, Daniel C. (1991). „Real Patterns“. In: *Journal of Philosophy* 88.1, S. 27–51.
- (2004). „Can Machines Think?“ In: *Alan Turing: Life and Legacy of a Great Thinker*. Hrsg. von Christof Teuscher. Berlin: Springer Verlag, S. 295–316.
- Dietrich, Manuel, Eugen Berlin und Kristof van Laerhoven (2015). „Assessing Activity Recognition Feedback in Long-term Psychology Trials“.

- In: *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. New York: ACM, S. 121–130.
- Dietrich, Manuel und Kristof van Laerhoven (2015). „A Typology of Wearable Activity Recognition and Interaction“. In: *Proceedings of the 2Nd International Workshop on Sensor-based Activity Recognition and Interaction*. New York: ACM, 1:1–1:8.
- (2016a). „An Interdisciplinary Approach on the Mediating Character of Technologies for Recognizing Human Activity“. In: *Philosophies* 1.1, S. 55–67.
- (2016b). „Reflect Yourself!: Opportunities and Limits of Wearable Activity Recognition for Self-Tracking“. In: *Lifelogging: Digital self-tracking and Lifelogging - between disruptive technology and cultural transformation*. Hrsg. von Stefan Selke. Wiesbaden: Springer Fachmedien Wiesbaden, S. 213–233.
- Donald J. Patterson Dieter Fox, Henry Kautz und M. Philipose (2005). „Fine-Grained Activity Recognition by Aggregating Abstract Object Usage“. In: *Proceedings of the 2005 International Symposium on Wearable Computers*. Los Alamitos: IEEE, S. 44–51.
- Dourish, Paul und Genevieve Bell (2011). *Divining a Digital Future: Mess and Mythology in Ubiquitous Computing*. Cambridge: MIT Press.
- Dreyfus, Hubert L. (1972). *What Computers Can't Do: The Limits of Artificial Intelligence*. New York: Harper und Row.
- (2002). „Intelligence Without Representation: Merleau-Ponty's Critique of Mental Representation“. In: *Phenomenology and the Cognitive Sciences* 1, S. 367–383.
- (2007). „Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian“. In: *Artificial Intelligence* 171.18, S. 1137–1160.
- Dwork, Cynthia et al. (2012). „Fairness Through Awareness“. In: *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*. New York: ACM, S. 214–226.
- Fällman, Daniel (2003). „In Romance with the Materials of Mobile Interaction : A Phenomenological Approach to the Design of Mobile Information Technology“. Diss. Umeå University, Informatics.
- Fisher, Ronald A. (1936). „The Use of Multiple Measurements in Taxonomic Problems“. In: *Annals of Eugenics* 7.2, S. 179–188.
- Friedman, Batya, Peter H. Kahn Jr. und Alan Borning (2013). „Value Sensitive Design and Information Systems“. In: *Human-Computer Interaction in Management Information Systems: Foundations*. Hrsg. von Ping Zhang und Dennis Galletta. New York: M.E. Sharpe, S. 1–27.
- Gaitsch, Peter (2012). „Das „Schwierigste“ des Leibes. Methodische Überlegungen in Auseinandersetzung mit Heidegger“. In: *Den Menschen im*

- Blick: phänomenologische Zugänge : Festschrift für Günther Pöltner zum 70. Geburtstag.* Hrsg. von Reinhold Esterbauer, Günther Pöltner und Martin Ross. Würzburg: Königshausen & Neumann, S. 159–178.
- Goodman, Bryce W. (2016). „A Step Towards Accountable Algorithms?: Algorithmic Discrimination and the European Union. General Data Protection“. In: *29th Conference on Neural Information Processing Systems*.
- Grün, Felix et al. (2016). *A Taxonomy and Library for Visualizing Learned Features in Convolutional Neural Networks*. Website. Online erhältlich unter <https://arxiv.org/abs/1606.07757>; abgerufen am 11. Juni 2018.
- Grunwald, Armin und Yannick Julliard (2005). „Technik Als Reflexionsbegriff: Überlegungen Zur Semantischen Struktur des Redens Über Technik“. In: *Philosophia Naturalis* 42.1, S. 127–157.
- Hård, Mikael und Andrew Jamison (2013). *Hubris and Hybrids: A Cultural History of Technology and Science*. New York: Routledge.
- Heidegger, Martin (1962). *Die Technik und die Kehre*. Opuscula aus Wissenschaft und Dichtung. Pfullingen: Neske.
- (1967). *Sein und Zeit*. Tübingen: Max Niemeyer Verlag.
- (1992). *Der Ursprung des Kunstwerkes*. Stuttgart: Reclam.
- Heider, Fritz (1927). *Ding und Medium*. Berlin: Weltkreis Verlag.
- Helal, Abdelsalam, Diane J. Cook und Mark Schmalz (2009). „Smart Home-Based Health Platform for Behavioral Monitoring and Alteration of Diabetes Patients“. In: *Journal of Diabetes Science and Technology* 3.1, S. 141–148.
- Helaoui, Rim, Daniele Riboni und Heiner Stuckenschmidt (2013). „A Probabilistic Ontological Framework for the Recognition of Multilevel Human Activities“. In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. New York: ACM, S. 345–354.
- Holleczek, Thomas et al. (2010). „Recognizing Turns and Other Snowboarding Activities with a Gyroscope“. In: *Proceedings of the 2010 International Symposium on Wearable Computers*. Los Alamitos: IEEE, S. 1–8.
- Houben, Steven et al. (2016). „Physikit: Data Engagement Through Physical Ambient Visualizations in the Home“. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York: ACM, S. 1608–1619.
- Hubig, Christoph (2006). *Die Kunst des Möglichen : 1. Technikphilosophie als Reflexion der Medialität*. Bielefeld: transcript Verlag.
- (2007). *Die Kunst des Möglichen : 2. Ethik der Technik als provisorische Moral*. Bielefeld: transcript Verlag.
- (2011). *Technik als Medium und „Technik“ als Reflexionsbegriff*. Website. Online erhältlich unter <https://www.philosophie.tu->

- darmstadt . de / media / institut _ fuer _ philosophie / diesunddas / hubig / downloadshubig / technik_als_medium_und_reflexionsbegriff.pdf; abgerufen am 15. Juli 2017.
- (2015). *Die Kunst des Möglichen : 3. Macht der Technik*. Bielefeld: transcript Verlag.
- Husserl, Edmund (2012). *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie : eine Einleitung in die phänomenologische Philosophie*. Hamburg: Meiner.
- Huynh, Tam und Bernt Schiele (2005). „Analyzing Features for Activity Recognition“. In: *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services*. New York: ACM, S. 159–163.
- IEEE (2016). *The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems*. Website. Online erhältlich unter http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html; abgerufen am 11. Mai 2017.
- Ihde, Don (1990). *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press.
- (1993). *Philosophy of Technology: An Introduction*. New York: Paragon House.
- IrisDataset (1937). *Iris Flower Data Set*. Website. Online erhältlich unter <https://archive.ics.uci.edu/ml/datasets/iris>; abgerufen am 11. Mai 2017.
- Jänicke, Stefan et al. (2015). „On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges“. In: *Eurographics Conference on Visualization*. Hrsg. von R. Borgo, F. Ganovelli und I. Viola. Geneva: The Eurographics Association, S. 1–21.
- Kaminski, Andreas (2010). *Technik als Erwartung : Grundzüge einer allgemeinen Technikphilosophie*. Zugl.: Darmstadt, Techn. Univ., veränd. Diss. Bielefeld: transcript Verlag.
- Kant, Immanuel (2012). *Grundlegung zur Metaphysik der Sitten*. [Nachdr. von 1985]. Stuttgart: Reclam Verlag.
- Kirsh, David (2013). „Embodied Cognition and the Magical Future of Interaction Design“. In: *ACM Transactions on Computer-Human Interaction* 20.1, 3:1–3:30.
- Kotsiantis, Sotiris B. (2007). „Supervised Machine Learning: A Review of Classification Techniques“. In: *Informatica* 31, 249–268.
- Krämer, Sybille (1998a). „Das Medium als Spur und als Apparat“. In: *Medien, Computer, Realität. Wirklichkeitsvorstellungen und Neue Medien*. Hrsg. von Sybille Krämer. Frankfurt am Main: Suhrkamp Verlag, S. 73–94.

- Krämer, Sybille (1998b). „Medien, Boten, Spuren. Wenig mehr als ein Literaturbericht“. In: *Was ist ein Medium?* Hrsg. von Stefan Münker. Frankfurt am Main: Suhrkamp Verlag.
- (2008). *Medium, Bote, Übertragung: Kleine Metaphysik der Medialität*. Frankfurt am Main: Suhrkamp Verlag.
- Kroll, Joshua A et al. (2016). „Accountable Algorithms“. In: *University of Pennsylvania Law Review* 165, S. 633–705.
- Kulesza, Todd et al. (2013). „Too Much, Too Little, or Just Right? Ways Explanations Impact end Users' Mental Models“. In: *2013 IEEE Symposium on Visual Languages and Human Centric Computing*. Los Alamitos: IEEE, S. 3–10.
- Kulesza, Todd et al. (2015). „Principles of Explanatory Debugging to Personalize Interactive Machine Learning“. In: *Proceedings of the 20th International Conference on Intelligent User Interfaces*. New York: ACM, S. 126–137.
- Lake, Brenden M. et al. (2016). „Building Machines That Learn and Think Like People“. In: *Behavioral and Brain Sciences*, S. 1–58.
- Latour, Bruno (2006). „Über technische Vermittlung: Philosophie, Soziologie und Genealogie“. In: *ANThology*. Hrsg. von Andréa Belliger und David J. Krieger. transcript Verlag, S. 483–528.
- Le Masurier, Guy C. und Catrine Tudor-Locke (2003). „Comparison of Pedometer and Accelerometer Accuracy under Controlled Conditions“. In: *Medicine and Science in Sports and Exercise* 35.5, S. 867–871.
- Lester, Jonathan et al. (2005). „A Hybrid Discriminative/Generative Approach for Modeling Human Activities“. In: *Proceedings of the 19th International Joint Conference on Artificial Intelligence*. San Francisco: Morgan Kaufmann Publishers Inc., S. 766–772.
- Li, Ian, Anind Dey und Jodi Forlizzi (2010). „A Stage-based Model of Personal Informatics Systems“. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM, S. 557–566.
- Li, Ian, Anind K. Dey und Jodi Forlizzi (2011). „Understanding My Data, Myself: Supporting Self-reflection with Ubicomp Technologies“. In: *Proceedings of the 13th International Conference on Ubiquitous Computing*. New York: ACM, S. 405–414.
- Lim, Brian Y., Anind K. Dey und Daniel Avrahami (2009). „Why and Why Not Explanations Improve the Intelligibility of Context-aware Intelligent Systems“. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM, S. 2119–2128.
- Luhmann, Niklas (2001). „Das Medium der Kunst“. In: *Niklas Luhmann: Aufsätze und Reden*. Hrsg. von Oliver Jahraus. Stuttgart: Reclam, 198–217.

- Lupton, Ellen (1993). *Mechanical Brides: Women and Machines from Home to Office*. New York: Princeton Architectural Press.
- Mahdaviani, Maryam und Tanzeem Choudhury (2008). „Fast and Scalable Training of Semi-Supervised CRFs with Application to Activity Recognition“. In: *Advances in Neural Information Processing Systems 20*. Hrsg. von J. C. Platt et al. Red Hook: Curran Associates, Inc., S. 977–984.
- McLuhan, Marshall, Quentin Fiore und Jerome Agel (1999). *The Medium Is the Massage: An Inventory of Effects*. San Francisco: HardWired.
- Merleau-Ponty, Maurice (1966). *Phänomenologie der Wahrnehmung*. Berlin: de Gruyter.
- (1976). *Die Struktur des Verhaltens*. Bd. 13. Phänomenologisch-psychologische Forschungen. Berlin: de Gruyter.
- Mersch, Dieter, Hrsg. (2003). *Die Medien der Künste : Beiträge zur Theorie des Darstellens*. München: Wilhelm Fink Verlag.
- Miyamori, Hisashi und Shun ichi Iisaku (2000). „Video Annotation for Content-based Retrieval Using Human Behavior Analysis and Domain Knowledge“. In: *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*. Los Alamitos: IEEE, S. 320–325.
- Nissenbaum, Helen (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford: Stanford University Press.
- Nordmann, Alfred (2008). *Technikphilosophie zur Einführung*. Hamburg: Junius Verlag.
- Oliver, Nuria und Eric Horvitz (2005). „A Comparison of HMMs and Dynamic Bayesian Networks for Recognizing Office Activities“. In: *Proceedings of the 10th International Conference on User Modeling*. Hrsg. von Liliana Ardissono, Paul Brna und Antonija Mitrovic. Berlin: Springer Verlag, S. 199–209.
- Partridge, Kurt und Bo Begole (2011). „Activity-Based Advertising“. In: *Pervasive Advertising*. Hrsg. von Jörg Müller, Florian Alt und Daniel Michelis. London: Springer London, S. 83–101.
- Peirce, Charles Sanders (1867). „On the Natural Classification of Arguments“. In: *Proceedings of the American Academy of Arts and Sciences 7*, S. 261–287.
- Perera, Charith et al. (2014). „Context Aware Computing for The Internet of Things: A Survey“. In: *IEEE Communications Surveys Tutorials 16.1*, S. 414–454.
- Philipose, Matthai et al. (2004). „Inferring Activities from Interactions with Objects“. In: *IEEE Pervasive Computing 3.4*, S. 50–57.
- Rammert, Werner und Ingo Schulz-Schaeffer (2002). „Technik und Handeln-Wenn soziales Handeln sich auf menschliches Verhalten und

- technische Artefakte verteilt“. In: *Können Maschinen handeln? : Soziologische Beiträge zum Verhältnis von Mensch und Technik*. Hrsg. von Werner Rammert. Frankfurt am Main: Campus-Verlag, S. 11–64.
- Ravi, Nishkam et al. (2005). „Activity Recognition from Accelerometer Data“. In: *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence*. Cambridge: AAAI Press, S. 1541–1546.
- Ritter, Joachim, Karlfried Gründer und Gottfried Gabriel, Hrsg. (1980). *Historisches Wörterbuch der Philosophie: Band 5*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Rogers, Yvonne (2011). „Interaction Design Gone Wild: Striving for Wild Theory“. In: *Interactions* 18.4, S. 58–62.
- Rooksby, John et al. (2014). „Personal Tracking As Lived Informatics“. In: *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*. New York: ACM, S. 1163–1172.
- Russell, Stuart J. und Peter Norvig (2010). *Artificial Intelligence: A Modern Approach*. 3. ed. Boston: Pearson.
- Scholl, Philipp M und Kristof Van Laerhoven (2012). „A Feasibility Study of Wrist-worn Accelerometer Based Detection of Smoking Habits“. In: *Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. Los Alamitos: IEEE, S. 886–891.
- ScikitLearn (2017). *Decision Trees on Scikit Learn*. Website. Online erhältlich unter <http://scikit-learn.org/stable/modules/tree.html>; abgerufen am 11. Mai 2017.
- Searle, John (1984). *Minds, Brains and Science*. Cambridge: Harvard University Press.
- Singh, Aneesha et al. (2014). „Motivating People with Chronic Pain to Do Physical Activity: Opportunities for Technology Design“. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM, S. 2803–2812.
- Sokolova, Marina und Guy Lapalme (2009). „A Systematic Analysis of Performance Measures for Classification Tasks“. In: *Information Processing & Management* 45.4, S. 427–437.
- Stiefmeier, Thomas et al. (2008). „Wearable Activity Tracking in Car Manufacturing“. In: *IEEE Pervasive Computing* 7.2, S. 42–50.
- Suchman, Lucy A. (1987). *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge: Cambridge University Press.
- Svanæs, Dag (2013). „Interaction Design for and with the Lived Body: Some Implications of Merleau-pony’s Phenomenology“. In: *ACM Transactions on Computer-Human Interaction* 20.1, 8:1–8:30.
- Turing, Alan (1950). „Computing Machinery and Intelligence“. In: *Mind* 59.236, 433–460.

- Van Laerhoven, Kristof, Eugen Berlin und Bernd Schiele (2009). „Enabling Efficient Time Series Analysis for Wearable Activity Data“. In: *Proceedings of the 8th International Conference on Machine Learning and Applications*. Los Alamitos: IEEE, S. 392–397.
- Verbeek, Peter-Paul (2005). *What Things Do: Philosophical Reflections on Technology, Agency, and Design*. University Park: Pennsylvania State University Press.
- Von Foerster, Heinz (2002). *Wissen und Gewissen: Versuch einer Brücke*. Frankfurt am Main: Suhrkamp Verlag.
- Waldenfels, Bernhard (2000). *Das leibliche Selbst : Vorlesungen zur Phänomenologie des Leibes*. Frankfurt am Main: Suhrkamp Verlag.
- Weiser, Mark (1991). „The Computer for the 21st Century“. In: *Scientific American* 265.3, S. 94–104.
- Weizenbaum, Joseph (1990). *Die Macht der Computer und die Ohnmacht der Vernunft*. Frankfurt am Main: Suhrkamp Verlag.
- Wiegerling Klaus und Küchenhoff, Joachim (2008a). *Leib und Körper*. Philosophie und Psychologie im Dialog. Göttingen: Vandenhoeck und Ruprecht.
- Wiegerling, Klaus (2008b). „Leib als symbolische Form und Ursprung von Medialität“. In: *Lebendige Form : zur Metaphysik des Symbolischen in Ernst Cassirers "Nachgelassenen Manuskripten und Texten"*. Hrsg. von Reto Luzius Fetz. Hamburg: Cassirer-Forschungen, S. 77–92.
- (2011). *Philosophie intelligenter Welten*. Paderborn: Wilhelm Fink Verlag.
- (2013). „Von Leibern und Körpern zur sekundären Leiblichkeit“. In: *Heimatschichten: Anthropologische Grundlegung eines Weltverhältnisses*. Hrsg. von Joachim Klose. Wiesbaden: Springer Fachmedien Wiesbaden, S. 211–237.
- Winner, Langdon (1977). *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*. Cambridge: MIT Press.
- (1980). „Do artifacts have politics?“ In: *Daedalus* 109.1, S. 121–136.
- Woolgar, Steve und Geoff Cooper (1999). „Do Artefacts Have Ambivalence? Moses' Bridges, Winner's Bridges and Other Urban Legends in STS“. In: *Social Studies of Science* 29.3, S. 433–449.
- Yang, Rayoung und Mark W. Newman (2013). „Learning from a Learning Thermostat: Lessons for Intelligent Systems for the Home“. In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. New York: ACM, S. 93–102.
- Ye, Juan, Simon Dobson und Susan McKeever (2012). „Review: Situation Identification Techniques in Pervasive Computing: A Review“. In: *Pervasive and Mobile Computing* 8.1, S. 36–66.
- Zeiler, Matthew D. und Rob Fergus (2014). „Visualizing and Understanding Convolutional Networks“. In: *Proceedings of the 13th European Conference*

on Computer Vision. Hrsg. von David Fleet et al. Cham: Springer International Publishing, S. 818–833.