

# Agency und Geschlecht in Mensch/Maschine-Konfigurationen am Beispiel von Virtual Personal Assistants



Diplomarbeit  
zur Erlangung des akademischen Grades Diplominformatiker

eingereicht am Institut für Informatik  
der Humboldt-Universität zu Berlin

**Göde Both**  
goedeboth@gmail.com

Betreuerinnen: Prof'in Dr. Beate Meffert  
Prof'in Dr. Sigrid Schmitz

eingereicht am: 27.6.11

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Grundlagen</b>	<b>4</b>
2.1	Informatische Grundlagen von VPAs . . . . .	4
2.1.1	Multimodale BenutzerInnen-Schnittstellen . . . . .	4
2.1.2	Schnittstellen-AgentInnen . . . . .	5
2.1.3	Software-AgentInnen . . . . .	7
2.1.4	Dialogsysteme . . . . .	9
2.2	Geschlechter- und Technikforschung . . . . .	11
2.2.1	Agency . . . . .	11
2.2.2	Begriffe der Akteur-Netzwerk-Theorie . . . . .	13
2.2.3	Narrationen und Figurationen . . . . .	15
2.2.4	Intra-Aktionen und <i>agential cuts</i> . . . . .	17
2.2.5	Geschlecht und Intersektionalität . . . . .	18
<b>3</b>	<b>Virtual Personal Assistants</b>	<b>22</b>
3.1	Konzeptionen . . . . .	22
3.2	Knowledge Navigator . . . . .	24
3.2.1	Szenario . . . . .	24
3.2.2	Analyse . . . . .	25
3.3	Aktuelle Implementationen . . . . .	29
3.4	Mensch/Maschine-Rekonfigurationen . . . . .	30
<b>4</b>	<b>Siri</b>	<b>33</b>
4.1	Software . . . . .	33
4.1.1	Funktionsumfang . . . . .	33
4.1.2	Vorgeschichte . . . . .	35
4.2	Hardware . . . . .	35
4.2.1	Architektur . . . . .	35
4.2.2	Fertigung . . . . .	37
4.3	Intra-Aktionen . . . . .	38

---

4.3.1	Cyborgisierung von Siri und NutzerIn . . . . .	39
4.3.2	Geschlecht in den Intra-Aktionen . . . . .	42
4.4	Beschreibung eines Szenarios . . . . .	44
4.4.1	Rückverteilung von Performanzen und Kompetenzen . . . . .	44
4.4.2	Präskriptionen . . . . .	47
4.4.3	Konfiguration der Nutzerin bzw. des Nutzers . . . . .	50
4.5	Figuration und Narration . . . . .	51
4.5.1	Arbeitsersparnisgeräte . . . . .	51
4.5.2	Dienstleistungsökonomie . . . . .	53
<b>5</b>	<b>Zusammenfassung und Ausblick</b>	<b>56</b>
	<b>Literaturverzeichnis</b>	<b>59</b>

# 1 Einleitung

You're busy. Between meetings, social events, and hopefully a workout or two, your schedule's packed. Don't you wish you could hand off simple tasks so you could have more time to play?

That's why we built Siri. Because we believe everyone could use an assistant. Because we believe there's a simpler way to get things done [91].

Arbeitsersparnis durch intelligente Maschinen ist ein wiederkehrendes Motiv, welches sowohl InformatikerInnen<sup>1</sup> als auch *science fiction*-AutorInnen immer wieder inspiriert. Die Verheißungen intelligenter Maschinen, welche unser aller Leben leichter machen, manifestieren sich im Interaktionsparadigma der Delegation. Seit Mitte der achtziger Jahre haben prominente InformatikerInnen wie Brenda Laurel, Alan Kay und Nicolas Negroponte Schnittstellen-Agenten zur NutzerInnen-Schnittstelle der Zukunft erklärt (vgl. [65], [51], [77]). Entgegen ihren Ankündigungen hat sich das Interaktionsparadigma bisher jedoch nicht durchsetzen können.

Die wachsende Verbreitung von mobilen Endgeräten wie *smartphones* weckt ein neues Interesse an Schnittstellen-Agenten mit Sprachverarbeitung, welche von einer extensiven militärischen Forschung in diesem Feld flankiert wird. *Virtual personal assistants* (VPAs) heißt die neue Generation von Software-AgentInnen, welche die Nutzung von Internetdiensten erleichtern soll. Bei VPAs handelt es sich um Softwarebasierte (*virtual*), personalisierte (*personal*) AssistentInnen (*assistants*). Die Verheißung der VPAs lautet, dass sie Aufgaben anstelle des Nutzers ausführen, um ihm bestimmte Arbeiten zu ersparen. Fast schon kanonische Anwendungsbeispiele sind Reservierungen von Restaurants, Bestellungen von Veranstaltungskarten und Flugbuchungen.

VPAs sind Cyborgs, da ihnen sowohl traditionell menschliche als auch maschinelle Eigenschaften zugeschrieben werden. Die spezifischen Mensch/Maschine-Konfigurationen von VPAs und ihren NutzerInnen sollen in der vorliegenden Diplomarbeit hinsichtlich der Kategorien *agency* (Handlungsfähigkeit, Wirkmächtigkeit) und Geschlecht analysiert werden. Daraus ergeben sich zwei Forschungsfragen: Erstens, welche Form von *agency*

---

<sup>1</sup> Um dem Ideal einer geschlechtergerechten Orthografie nahe zu kommen, wird im Plural das 'Binnen-I' verwendet und im Singular abwechselnd weibliche und männliche Genera.

wird VPAs in den informatischen Forschungen zur Künstlichen Intelligenz und Mensch-Computer-Interaktion zugeschrieben und wie kann *agency* aus der Perspektive der feministischen Technikforschung alternativ konzipiert werden? Zweitens, wie werden VPAs vergeschlechtlicht, das heißt, wie werden sie durch die Geschlechterordnung strukturiert und wie produzieren die VPAs die Kategorie Geschlecht? Die vorliegende Arbeit untersucht nicht die geschlechtsspezifischen Nutzungs- und Gestaltungsweisen von Männern und Frauen, sondern die Ko-Materialisierung von Geschlecht und Technik.

Neben der Informatik dienen Konzepte aus der Geschlechter- und Technikforschung als theoretische Fundierung der Untersuchung. Dieses interdisziplinäre Diplom-Vorhaben<sup>2</sup> bedurfte einer intensiven und selbstständigen Einarbeitung in, für den Verfasser, zum Teil fremde Wissensbestände. Das Ergebnis des Übersetzungsprozesses soll eine Brücke zwischen unterschiedlichen Wissenschaftsfeldern schlagen und trägt zu einem innovativen Verständnis von VPAs bei. Diese Diplomarbeit knüpft an die Forderung der ehemaligen Arbeitsgruppe *Theorie der Informatik der Gesellschaft für Informatik* an, die Grundlagen der Informatik im breiteren Sinne zu reflektieren. Im Gegensatz zur Theoretischen Informatik geht es bei einer Theorie der Informatik nicht alleine um die mathematische Begründung, weil »viele grundlegende Fragen der Informatik keine formalen Fragen sind [...]« wie Wolfgang Coy aus der o. g. Arbeitsgruppe kommentiert.[23, S. 22]. In diesem Sinne möchte die vorliegende Arbeit einen Beitrag zur Diskussion der philosophischen Vorannahmen und ihrer soziotechnischen Konsequenzen leisten. Die Thematisierung von Ein- und Ausschlüssen strebt so dem Ziel einer inklusiven und damit geschlechtssensiblen Informatik entgegen.

Neben der Auswertung von wissenschaftlichen VPA-Konzepten und der Analyse des visionären Videos *Knowledge Navigator*, wird ein aktueller VPA empirisch untersucht. Die Wahl fiel auf *Siri*, weil sie auf Basis von Forschungsergebnissen aus dem DARPA-Projekt *Cognitive Assistant that learns and organizes* implementiert wurde. *Siri* verkörpert so den aktuellen Stand der Technik und gilt als einflussreich für die weitere Entwicklung von VPAs. Die Diplomarbeit besteht aus fünf Kapiteln. Im Anschluss an die Einleitung werden im zweiten Kapitel die für die Untersuchung relevanten Grundkonzepte aus der Informatik, Technik- und Geschlechterforschung erläutert. Multimodale Schnittstellen, Dialogsysteme sowie Schnittstellen- und Software-AgentInnen bilden die Bezugspunkte

---

<sup>2</sup> Für das entgegengebrachte Vertrauen und die Unterstützung meines interdisziplinären Themas danke ich meinen Gutachterinnen Prof'in Dr. Beate Meffert und Prof'in Dr. Sigrid Schmitz. Mein Dank gilt insbesondere Dr. des. Corinna Bath und der Forschungswerkstatt *Gender in MINT* unter der Leitung von Dr. Petra Lucht für die intensive und hilfreiche Begleitung meiner Diplomarbeit. Für das Korrekturlesen und die konstruktive Kritik zeige ich mich erkenntlich gegenüber (in alphabetischer Reihenfolge) Andrea Knaut, Bärbel Mauss, Claude Draude, Franziska Matthis, Kristin Witte, Micha Plöse, Olaf M. Braun, Philipp Marquardt, Sebastian Winkler und Tobias Florek.

---

für die informatischen Sichtweisen. Die intersektionale Analyse von Geschlecht und die Akteur-Netzwerk-Theorie inklusive ihrer Weiterentwicklung in der feministischen Technikforschung stecken den begrifflichen Rahmen für die Perspektiven der Geschlechter- und Technikforschung ab. Vor diesem Hintergrund werden im dritten Kapitel zwei VPA-Konzepte verglichen und ihre theoretischen Vorannahmen hinterfragt. Die Leitvision *Knowledge Navigator* wird beschrieben und analysiert. Darüber hinaus gibt das Kapitel einen Überblick über aktuelle kommerzielle und militärische Anwendungen von VPAs. Im vierten Kapitel wird der VPA *Siri* hinsichtlich der Hardware, der Software und der Intra-Aktionen von Nutzerin und VPA untersucht. Darauf folgt die Beschreibung eines Szenarios auf Grundlage der erarbeiteten Begrifflichkeiten. Abschließend wird *Siri* im Kontext gesellschaftlicher Erzählungen und symbolisch-struktureller Geschlechterordnung positioniert. Kapitel 5 fasst die Thesen dieser Arbeit zusammen und diskutiert die Übertragbarkeit der Ergebnisse auf andere informatische Artefakte.

## 2 Grundlagen

In diesem Kapitel werden die wichtigsten Grundkonzepte für die Untersuchung von *virtual personal assistants (VPA)* dargelegt. Es steckt gleichzeitig den methodischen Rahmen für die Analyse in den nachfolgenden Kapiteln ab. Diese Arbeit ist interdisziplinär angelegt, so dass in diesem Kapitel sowohl Konzepte aus der Informatik als auch aus der Geschlechter- und Technikforschung diskutiert werden.

### 2.1 Informatische Grundlagen von VPAs

Die US-amerikanischen Informatiker und Psychologen Jun Xiao, Richard Catrambone und John Stasko bezeichnen die Konstruktion von Assistenten als ein komplexes Vorhaben, welches Forschungen auf unterschiedlichen Gebieten erfordert [107, S. 384]. Folglich ist es für diese Arbeit notwendig, zunächst eine Übersicht über diejenigen informatischen Forschungsbereiche zu geben, auf denen VPAs beruhen: Multimodale BenutzerInnen-Schnittstellen (→2.1.1) und Schnittstellen-AgentInnen (→2.1.2) aus der Perspektive der Mensch-Computer-Interaktion (MCI), Software-AgentInnen (→2.1.3) im Bereich der Künstlichen Intelligenz (KI) und Dialogsysteme (→2.1.4) aus dem Blickwinkel der Sprachverarbeitung.

#### 2.1.1 Multimodale BenutzerInnen-Schnittstellen

Mensch-Computer-Interaktion ist ein interdisziplinäres Forschungsfeld, welches sich in Deutschland vorwiegend aus der Informatik und der Psychologie speist. Die britische MCI-Forscherin Christine Faulkner nennt in ihrem Lehrbuch [32, S.1 ff.] weitere Disziplinen: Physiologie, Arbeitswissenschaft, Industriedesign, Ingenieurwissenschaften, Soziologie und Ethnologie. Faulkner definiert MCI als das Studium der Beziehungen zwischen NutzerInnen und Computern, welche beim Erledigen vielfältiger Aufgaben zum Tragen kommen. Ihrer Auffassung nach besteht MCI außerdem darin, die Fähigkeiten der NutzerInnen und Computer im Allgemeinen zu verstehen. Das Ziel sei die Verbesserung der Effizienz der Zusammenarbeit zwischen Mensch und Maschine sowie die Ergonomie der Bedienung [32, S. 2].

Der Ort der Interaktion zwischen Mensch und Computer ist die BenutzerInnen-Schnittstelle. In ihrem populären MCI-Lehrbuch definieren Alan Dix und Janet Finley Interaktion als die Kommunikation zwischen Benutzerin und System [26, S. 104]. Sie schlagen vor, die Interaktion an der Schnittstelle mit einem Interaktionsmodell zu untersuchen. Dix und Finley zufolge machen Interaktionsmodelle deutlich, welche Übersetzungen zwischen dem, was der Benutzer möchte, und dem, was das System tut, erfolgen. Ihr Interaktionsmodell besteht aus vier Komponenten: dem System, der Benutzerin und der Schnittstelle – bestehend aus Output und Input [26, S. 106 ff.]. Der Benutzer teilt dem System seinen Willen über den Input mit. Die Eingabe wird vom Computer verarbeitet und über die Ausgabe präsentiert er ein Resultat. Der Benutzer nimmt das Ergebnis wahr. Auf diese Weise beschreibt Dix einen rudimentären Kreislauf der Interaktion.

Seit der Einführung von Graphischen BenutzerInnen-Schnittstellen (GUI, englisch für *graphical user interface*) in den 1980er Jahren, durch den Xerox Alto und Star hat sich das Paradigma der direkten Manipulation durchgesetzt [26, S. 150]. Anstelle von komplexen Befehlsangaben können Daten, dargestellt durch visuelle Objekte, direkt verändert werden. Die gängigen GUIs bei Microsoft Windows oder Apple Mac OS greifen in der Regel auf Fenster, Symbole, Menüs und Zeiger (WIMP<sup>1</sup>) zurück [26, S. 120 ff.]. Es gibt zahlreiche andere Arten von Schnittstellen, die zusammengenommen fast alle kognitiven Fähigkeiten und Ausdrucksmöglichkeiten von Menschen abdecken: Gestensteuerung, Sprachsteuerung und mechanische, haptische, Geschmacks- und olfaktorische Schnittstellen [55, S. 4-10]. Letztere befinden sich zwar im Experimentierstadium, eine einfache haptische Schnittstelle ist jedoch z. B. in allen Mobiltelefonen mit Vibrationsalarm integriert. Die Mehrheit der Schnittstellen sprechen nicht nur einen Sinn an, sondern kombinieren verschiedene Arten von Schnittstellen. Multimodale Schnittstellen sind nach dem US-amerikanischen Psychologen Philip Kortum keine Seltenheit sondern eher die Regel [55, S. 11]. Wenn ich beispielsweise gerade diesen Text mit meinem Notebook verfasse, bediene ich dabei einer Tastatur und einem *touchpad*, während der Bildschirm vom Notebook meine Eingaben visuell präsentiert. Die gebräuchliche Konfiguration von Maus oder *touchpad*, Tastatur und Bildschirm ist eine **multimodale Schnittstelle**, weil sie unterschiedliche Schnittstellen für einen Zweck kombiniert.

### 2.1.2 Schnittstellen-AgentInnen

VPAAs erweitern die klassische GUI um zwei weitere Schnittstellen: Schnittstellen-AgentInnen<sup>2</sup> und Dialogsysteme (→2.1.4).

---

<sup>1</sup> englisch für *Windows, Icons, Menus, Pointer*

<sup>2</sup> Die Medienpsychologin Nicole Krämer führt aus, dass eine Reihe von englischen und deutschen Synonymen für Schnittstellen-AgentInnen zirkulieren: *embodied conversational agents, virtual assi-*



Alan Kay, dessen Arbeit am Xerox PARC maßgeblich die Entwicklung von heutigen Arbeitsplatzrechnern beeinflusst hat, entwarf 1984 in einem viel zitierten Artikel die Vision von Software-Agenten (»soft robot«) als Mensch-Computer-Schnittstelle [51, S. 8]. In Kays Artikel informiert die Nutzerin den Computer lediglich über das Ziel einer Aufgabe und die Agentin führt selbstständig die notwendigen Arbeitsschritte durch. Bei Bedarf würde die Agentin nachfragen, wenn sie nicht mehr weiter kommt. Kay betont, dass Idee und Konzept von Software-AgentInnen nicht von ihm stammen, sondern bereits Mitte der 1950er am Massachusetts Institute of Technology (MIT) von John McCarty und Oliver G. Selfridge entwickelt wurden. Der KI-Forscher Nicholas Negroponte sieht in Schnittstellen-Agenten ein neues Paradigma, welches die direkte Manipulation zu Gunsten der Delegation ablösen wird [77, S. 7].

Die MCI-Forscherin Brenda Laurel definiert Schnittstellen-AgentInnen als »a character, enacted by the computer who acts on behalf of the user in a virtual (computer-based) environment« [65, S. 356]. Laurel stellt sich vor, dass die Interaktion über einen Avatar erfolgt, der dann anstelle des Nutzers handelt. Die Repräsentation von Schnittstellen-AgentInnen durch menschenähnliche Figuren wird gegenwärtig nicht mehr von allen ForscherInnen geteilt. So kann die Interaktion mit Software-AgentInnen auch über Eingabemasken ähnlich einer Suchmaschine erfolgen, meint der Schweizer Informatiker Rolf Grütter [40, S. 3]. Schnittstellen-AgentInnen müssen folglich äußerlich nicht in der Gestalt von Avataren erscheinen.

Um Schnittstellen-AgentInnen zu kategorisieren schlagen Catrambone et al. eine Reihe von Kategorien vor [18, S. 167 ff.]. Für die vorliegende Arbeit sind *role*, *initiative* und *gender* relevant:

**Rolle** Die Kategorie Rolle beschreibt das Verhalten der Schnittstellen-Agentin in der Interaktion mit dem Nutzer. Assistriert sie immer bei der Erledigung von Aufgaben oder soll sie nur auf Wunsch des Nutzers Unterstützung leisten?

**Initiative** Soll die Agentin unaufgefordert Vorschläge machen und Hilfe anbieten oder soll sie nur dann aktiv werden, wenn sie von der Nutzerin aufgerufen wird?

**Geschlecht** Wird der Agentin ein Geschlecht zugeschrieben?

Ein wichtiges Merkmal von Schnittstellen-AgentInnen ist die Adaptivität in Bezug auf ihre NutzerInnen. Dieser Prozess wird Personalisierung genannt. Der Informatiker André Klahold definiert **Personalisierung** als »[...] die Anpassung von Informationen, Diensten oder Produkten an die definierten Bedürfnisse einer Person« [54, S. 3]. Die Anpassung

---

*stants* und anthropomorphe Schnittstellen [56, S. 13]. Ich bin bei meinen Recherchen zusätzlich auf die Bezeichnungen *intelligent agent* oder *smart assistant* gestoßen.

könne sowohl auf Basis eines NutzerInnenprofils und des Kontextes als auch durch aktive Personalisierung erfolgen. Letzteres erfordere das Anlernen eines Programms durch die Nutzerin. Personalisierung verspricht, die BenutzerInnen-Schnittstelle individuell an die NutzerInnen anzupassen.

### 2.1.3 Software-AgentInnen

Schnittstellen-AgentInnen repräsentieren nur einen Teilbereich von dem, was in der Informatik unter dem Begriff Agent verhandelt wird. Eine klare Abgrenzung von Schnittstellen-AgentInnen gegenüber anderen Software-AgentInnen ist in der Praxis schwierig. Sabine Payr vom *Österreichischen Forschungsinstitut für Artificial Intelligence (OFAI)* unterscheidet zwei Perspektiven zu AgentInnen. Die erste Sichtweise, welche eher typisch für die MCI-Forschung ist, betont, dass AgentInnen Aufgaben im Auftrage von Menschen lösen sollen. Die zweite Perspektive wird tendenziell in der KI-Forschung vertreten, wonach eine Agentin ihre internen Zustände kontrollieren kann [81, S. XV]. In einer BDI-Agenten-Architektur sind es beispielsweise die inneren Zustände der Annahmen (*beliefs*), Wünsche (*desires*) und Absichten (*intentions*). Schnittstellen-AgentInnen können immer durch die MCI-Sichtweise beschrieben werden, sie müssen jedoch nicht der KI-Perspektive gerecht werden. Für die KI-Forschung stellen AgentInnen eher ein spezifisches Paradigma der Software-Entwicklung dar als ein Bedienungskonzept. Eine allgemein akzeptierte Definition, was eine Agentin auszeichnet, gibt es nicht. In diesem Punkt stimmen viele AutorInnen überein [105, S. 15] [90, S. 4] [15, S. 1012] [14, S. 4]. Trotz dieser Übereinkunft gibt der Informatiker Hans-Dieter Burkhard eine Definition, wohl wissend, dass nicht alles, was unter diesen Begriff verhandelt wird, mit dieser Formulierung erfasst werden kann:

Ein Software-Agent ist ein längerfristig arbeitendes Programm, dessen Arbeit als eigenständiges Erledigen von Aufträgen oder Verfolgen von Zielen in Interaktion mit einer Umwelt beschrieben werden kann [15, S. 949].

Im Unterschied zu normativen Definitionen wie etwa bei Payr und Wooldridge (vgl. [81] [105]) zeichnet sich Burkhard's Definition dadurch aus, dass die Eigenschaft ein Agent zu sein, einem Programm zugeschrieben werden muss. Einige grundlegende Merkmale von AgentInnen werden deutlich: Autonomie, Zielorientierung, Interaktion mit Umwelt und Reaktivität. Diese und weitere Eigenschaften sollen nun in Anlehnung an Burkhard und Payr sowie den britischen KI-Forschern Micheal Wooldridge und Nick Jennings erläutert werden [15, S. 1012 ff.] [81, S. XIV ff.] [106, S. 4 ff.]:

**Zielorientierung** Das Handeln der AgentInnen ist auf ein Ziel ausgerichtet.

**Autonomie** Der Agent kontrolliert sein Handeln und seine inneren Zustände. Für das Erreichen von Zielen und Erledigen von Aufträgen notwendige Schritte und Entscheidungen werden vom Agenten selbstständig durchgeführt und der Agent handelt ohne direkte Intervention im Auftrag der BenutzerInnen oder anderer AgentInnen. Burkhard umschreibt die Autonomie-Eigenschaft mit der Analogie, der Agent leiste »Dienst nach Vorschrift«. Payr bezeichnet diese auch als *indirect management*.

**Reaktivität** bezeichnet die unmittelbare Reaktion von AgentInnen auf Stimuli. Was ein Stimulus für eine spezifische Agentin auszeichnet, hängt von der Konstruktion und ihrer Einbettung ab.

**Sozialverhalten** Das Sozialverhalten eines Agenten beschreibt die Interaktion mit anderen AgentInnen oder Menschen. Wenn ein Ziel nicht eigenständig erreicht werden kann, kooperiert der Agent mit anderen AgentInnen. Die Kooperation mit anderen AgentInnen folgt festgelegten Normen und Regeln.

**Rationalität** bezeichnet die Fähigkeit sinnvolle Entscheidungen, auf Basis des der Agentin zur Verfügung stehenden Wissens zu treffen. Burkhard betont, dass diese Entscheidungen nicht in jedem Fall optimal sein müssen, da das Wissen der Agentin über die relevante Welt notwendigerweise beschränkt ist.

**Interaktion mit der Umwelt** meint nur den für die Agentin relevanten Teil der »Welt«.

**Lernfähigkeit** beschreibt die Möglichkeit, aus vergangenen Interaktionen und Entscheidungen zu lernen.

Eine Teilmenge von AgentInnen zeichnet sich zusätzlich durch **Proaktivität** aus. Laut dem Informatiker Jörg Pflüger treffen proaktive AgentInnen für die NutzerInnen kontextabhängige Entscheidungen. Proaktivität geht über Reaktivität insofern hinaus, als dass für die Agentin kein Stimulus notwendig ist, um aktiv zu werden [83, S. 360 ff.]. Wooldridge und Jennings fordern von proaktiven AgentInnen ein, in angemessenen Fällen die Initiative zu ergreifen [106, S. 4]. Laut Pflüger setzt Proaktivität eine **Kontextsensitivität** voraus. Der Agent müsse die Sinnzusammenhänge der Nutzerin interpretieren und das wichtigste Ziel erkennen [83, S. 370]. Im einfachsten Fall lassen sich Ort- und Zeitinformationen zu Nutze machen oder durch Personalisierung das Wissen über die Nutzerin.

Damit die Agentin Rückschlüsse auf die Absicht des Nutzers machen kann, muss das notwendige Wissen in einer formalisierten Form vorliegen. Dies wird in der KI-Forschung Wissensrepräsentation genannt. Im Kontext von VPA sind Domänen- und

Aufgaben-Modelle als Form der Wissensrepräsentation relevant. Die KI-Forscher Bernd Owsnicki-Klewe, Kai von Luck und Bernhard Nebel definieren eine **Domäne** als einen »darzustellenden Weltausschnitt«, und eine **Aufgabe** beschreiben sie als »geforderte Funktionalität eines Problemlösers, der auf der Domäne arbeitet« [80, S. 157]. Das zugehörige **Modell** bezeichnen sie als die »Rekonstruktion eines Weltausschnitts«. Ihre formale Definition eines Modells lautet:

[ $F$  sei eine Formel der Prädikatenlogik erster Stufe. G.B.] Die Formel  $\exists x : F$  wird von der Interpretation erfüllt, falls es eine Belegung von  $x$  mit einem Element der Menge gibt, die  $F$  erfüllt. Die Formel  $\forall x : F$  wird von der Interpretation erfüllt, falls sie für alle Belegungen von  $x$  mit einem Element der  $F$  erfüllt ist. Eine Interpretation, die eine Formel  $F$  erfüllt, heißt auch Modell von  $F$  [80, S. 159].

Einem Modellierungsproblem wird in der Regel mit einer formalen Notation begegnet. Die Gegebenheiten einer Domäne werden durch Axiome formalisiert. Für die Modellierung von Aufgaben müssen dem formalen Notationssystem entsprechende deduktive Fähigkeiten zugeordnet werden [80, S. 158]. In der Regel werden Aufgaben in eine lineare Sequenz von Einzelschritten zerlegt.

### 2.1.4 Dialogsysteme

Ein **Dialogsystem** ist eine Schnittstelle, bei der der Nutzer in gesprochener Sprache mit dem Computer kommuniziert. Es soll nun ein Überblick über die Grundlagen von dialogbasierten Schnittstellen gegeben werden. In ihrem Lehrbuch zur stochastischen Sprachverarbeitung führen Andreas Wendemuth und seine Mit-AutorInnen umfassend in dieses Feld ein. Ein Dialogsystem besteht nach Wendemuth et al. aus fünf Komponenten [100, S. 8]:

1. Spracherkennung
2. Aussprache
3. Verstehen
4. Dialogkontrolle
5. Ausgabe

Unter Spracherkennung verstehen sie die »Ermittlung des syntaktischen Inhalts eines gesprochenen Textes« [100, S. VI]. Wendemuth et al. unterscheiden zwischen zwei Typen von Spracherkennern. Bei Einzelworterkennern muss die Sprecherin Pausen zwischen den

Worten einlegen, wohingegen bei der **kontinuierlichen Erkennung** fließend gesprochen werden kann.

Ein stochastisches Spracherkennungssystem besteht aus zwei Arbeitsschritten: Erstens werden während der Akustischen Analyse im Musterkennungsprozess charakteristische Merkmale aus dem Audio-Signal gewonnen. Der Merkmalsvektor wird zweitens mit Hilfe der Maximum-Likelihood-Schätzung in die wahrscheinlichste Wortfolge übersetzt. Die Suche nach der besten Hypothese erfolgt über die Funktion:

$$\max P(A|W)P(W)$$

Die Wahrscheinlichkeit  $P(W)$  beschreibt das Sprachmodell. Dieses Modell berücksichtigt die statistischen, grammatikalischen und syntaktischen Regeln der zugrunde liegenden Sprache. So ist es möglich, auch falsch oder nicht erkannte Worte zu rekonstruieren. Das Akustische Modell entspricht der bedingten Wahrscheinlichkeit  $P(A|W)$  der aus dem aufgezeichneten Schall gewonnenen Merkmalsvektoren, die einem Wort zugeordnet werden können. Die Maximierung des Produkts aus  $P(W)$  und  $P(A|W)$  führt zu der wahrscheinlichsten Hypothese, in diesem Fall das wahrscheinlichste Wort. In der Regel wird die Erkennung noch durch eine Nachbearbeitung unterstützt. Im n-Best-Verfahren wird beispielsweise die Wortsequenz mit der höchsten Wahrscheinlichkeit ausgewählt [100, S.8ff.]. Die wahrscheinlichste Hypothese entspricht durch die Nachbearbeitung einer Wortfolge. Bei der kontinuierlichen Spracherkennung wird zwischen zwei Typen unterschieden:

**sprecherInnenabhängig** Bei diesem Typ muss die Spracherkennung von der Anwenderin trainiert werden. Die sprecherInnenabhängige Variante hat in der Regel ein kleines Vokabular mit 100-500 Wörtern und verwendet kein Sprachmodell.

**sprecherInnenunabhängig** Durch das Lesen von Referenztexten durch viele Personen, wird SprecherInnenunabhängigkeit angestrebt. Das Vokabular dieses Typs verfügt in der Regel über mehr als 100000 Worte und verwendet ein Sprachmodell.

Die Erkennung einzelner Worte und Sätze, die Spracherkennung, grenzen Wendemuth et al. vom Sprachverstehen ab, der Ermittlung des semantischen Inhalts. Sprachverstehen konzentrierte sich auf die Bedeutung von gesprochenen Texten und diene zur Steuerung der Dialoge zwischen Mensch und Computer. Außerdem ließe die Semantik Rückschlüsse über die mögliche Absicht des Nutzers zu.

## 2.2 Geschlechter- und Technikforschung

Im zweiten Teil des Grundlagenkapitels sollen die Konzepte aus der Technik- und Geschlechterforschung vorgestellt werden, die sich zur Analyse der *virtual personal assistants (VPA)* eignen: *Agency* (→2.2.1), das Vokabular der Akteur-Netzwerk-Theorie (→2.2.2), Figurationen und Narrationen (→2.2.3), Intra-Aktionen (→2.2.4) sowie die unterschiedlichen Analysedimensionen von Geschlecht (→2.2.5).

### 2.2.1 Agency

Ob zufällig oder nicht, die KI-Forschung und die Technikforschung haben in etwa zeitgleich angefangen, sich mit *agency* zu beschäftigen [1, S. 64]. Im Kontext dieser Arbeit ist *agency* als analytische Kategorie gewählt worden, weil es zum einen die beiden wissenschaftlichen Disziplinen verbindet und zum anderen die Differenzen veranschaulicht. In der KI-Forschung und in der Technikforschung wird *agency* unterschiedlich gebraucht. Eine gängige Definition in der KI-Forschung lautet beispielsweise:

Agency degree of autonomy and authority vested in the agent, and can be measured at least qualitatively by the nature of the interaction between the agent and other entities in the system (zitiert nach [14, S. 9]).

Im Kontext der KI-Forschung beschreibt *agency* die Befähigung einer Agentin, Entscheidungen zu treffen. Proaktivität ist der höchste Grad an *agency*. Für die Technikforschung lässt sich gemäß dem Übersetzer von Bruno Latours Monographien, Gustav Roßler, keine adäquate Entsprechung in einem deutschen Wort finden, welches die vielen Bedeutungsnuancen und Facetten abdeckt [63, S.79]. Handeln, Wirken, Kraft, Wirkmächtigkeit, Wirksamkeit, Handlungsfähigkeit, Handlungsmacht, HandlungsträgerInnenschaft sind nur einige der vielen Übersetzungsmöglichkeiten für *agency*.<sup>3</sup>

Die vorliegende Arbeit bezieht sich auf das Konzept von *agency* aus der Akteur-Netzwerk-Theorie und ihrer Weiterentwicklung. Das Konzept bedarf einer Erläuterung seines Entstehungskontextes und seiner theoretischen Grundlage. Unter der Bezeichnung *théorie de l'acteur-réseau* (Akteur-Netzwerk-Theorie, Abkürzung: ANT) entstand sie als Alternative zu den technik- und sozialdeterministischen Auffassungen von Technik. Technikdeterminismus ist eine Theorie, nach der eine Gesellschaft durch ihre technologische Entwicklung bestimmt wird. Der Sozialdeterminismus nimmt spiegelbildlich an, dass Technik das Resultat der sozialen Verhältnisse ist. Beide Theorien arbeiten mit einer groben Vereinfachung auf Basis des Ursache-Wirkung-Prinzips und beruhen auf

---

<sup>3</sup> Nur zur Vollständigkeit sei erwähnt, dass *agency* zusätzlich mit 'Büro', 'Amt', 'Filiale', 'Geschäftsstelle' oder 'Agentur' übersetzt werden kann. Diese Bedeutungen haben für diese Arbeit keine Relevanz.

der strikten Trennung von Gesellschaft und Technik. Die ANT beansprucht für sich, diese Defizite zu überwinden.

Die frühen ProtagonistInnen der ANT arbeiteten zu Beginn der 1980er Jahre im Umfeld der *Ecole Nationale Supérieure des Mines*, einer renommierten technischen Universität in Paris. Zu diesem Kreis zählen zunächst der Philosoph und Technikanthropologe Bruno Latour, die Ingenieurin und Soziologin Madeleine Akrich, der Ingenieur und Soziologe Michel Callon und der britische Soziologe John Law. Die ANT schreibt gegen die Technikvergessenheit der Sozialwissenschaften an. In ihren Studien wenden sie sich den Naturwissenschaften und den Technologien zu, welche sie mit ethnographischen Methoden untersuchen. Die Werkzeuge der Ethnologie werden so auf Wissenschaftskulturen angewendet.

ANT und KI-Forschung haben gemein, dass *agency* nicht nur exklusiv für Menschen reserviert wird. In der KI-Forschung werden RoboterInnen und Software-AgentInnen *agency* zugebilligt. Im Rahmen der ANT wird Handlungsfähigkeit nicht nur Menschen, sondern, wie die ANT-ForscherInnen es ausdrücken, auch **Nicht-Menschen** eingeräumt. Die Spannbreite der Nicht-Menschen reicht von foto-elektrischen Lichtanlagen in der Entwicklungszusammenarbeit [2], über Mikroben bei der »Pasteurisierung Frankreichs« [61] und Fahrbahnschwellen [62], bis hin zum Satz des Pythagoras [63]. Die Beispiele bekräftigen eine entscheidende Differenz. Im Unterschied zur KI-Forschung kann in der ANT *agency* auch weniger 'intelligenten' Entitäten eingeräumt werden. Die Menschen und Nicht-Menschen, zwischen denen Handeln verteilt wird, werden in der ANT-Terminologie **Aktanten** oder Agenten genannt. Um in der ANT den Status eines Aktanten eingeräumt zu bekommen, muss er/sie/es Teil einer Handlungskette sein, in der er/sie/es auf andere Aktanten einwirkt und ihre Ziele verändert.

Als methodische Vorkehrung wird eine Symmetrie zwischen menschlichen und nicht-menschlichen Aktanten angenommen [9, S. 15]. Das bedeutet, dass nicht vor der Analyse anhand der Unterscheidung zwischen Mensch und Nicht-Mensch festgelegt wird, wer am Handeln beteiligt ist und wer nicht. Vielmehr sind Asymmetrien in der Wirkmächtigkeit von unterschiedlichen Aktanten das Ergebnis ihrer Interaktion. Die Unterscheidung Mensch/Nicht-Mensch ist in der westlich-aufklärerisch Denktradition eng verbunden mit der Dichotomie von Subjekt/Objekt. Diese Opposition konstruiert Objekte als passive Entitäten und Subjekte als ihr aktives Gegenüber. Während Subjekte und Objekte in der westlich-aufklärerischen Ontologie jeweils getrennte Bereiche einnehmen, Subjekte gehören zur Politik und Objekte zur Natur, will die ANT die Unterscheidung nicht wiederholen. Sie kritisieren u. a., durch die Subjekt/Objekt-Trennung werde bereits vor der Analyse festgelegt, wer oder was am Handeln beteiligt ist und wer oder was nicht.

## 2.2.2 Begriffe der Akteur-Netzwerk-Theorie

Für die symmetrische Beschreibung von menschlichen und nicht-menschlichen Aktanten entwickelt die ANT ein spezifisches Vokabular. Die für diese Arbeit relevanten Konzepte sollen nun vorgestellt werden. Ein nachfolgendes Beispiel veranschaulicht die Begrifflichkeiten.

**Akteur-Netzwerk** Ein Akteur-Netzwerk deckt sich weder mit der geläufigen Vorstellung eines technischen Netzwerks, wie etwa einem Telefon- oder Eisenbahnnetz, noch mit einem informellen Netzwerk zwischen Personen [63, S. 225]. Es ist ein Oxymoron, welches den hybriden Charakter unterstreichen soll. Akteurs-Netzwerke sind sowohl im landläufigen Sinne technisch als auch sozial. Das heißt, sie sind **soziotechnisch**.

**Substitution** Bei einer Substitution wird ein Aktant in einer gegebenen Handlungskette durch einen anderen ersetzt.

**Übersetzung** Die Übersetzung (*traduction*) ist der zentrale Begriff der ANT und beschreibt bewusst mehrdeutig das Handeln der Aktanten. In der Übersetzung werden die »Identität der Akteure, die Möglichkeit der Interaktion und der Handlungsraum ausgehandelt und abgegrenzt« [17, S. 146]. Mit Übersetzung beschreibt die ANT das gegenseitige Einwirken der Aktanten auf ihr Handeln. Aktanten übersetzen ihre Ziele, das heißt, sie modifizieren sie gegenseitig. Durch eine Übersetzung kann auch die Bedeutung verschoben oder transformiert werden.

**Delegation** Bei einer Delegation steht ein Aktant für einen anderen Aktant ein. Eine Berliner Verkehrsampel ist beispielsweise eine Delegierte der Senatsverwaltung für Stadtentwicklung. Die Verkehrslenkung wird räumlich, zeitlich und aktorial verschoben. Die Delegation erzeugt eine Asymmetrie zwischen den abwesenden HerstellerInnen und den zeitweiligen NutzerInnen [60, S. 230].

**Artikulation** Techniken können die gängige Grenze zwischen Zeichen und Dingen überschreiten. Artikulation ist in der ANT nicht auf die menschliche Rede beschränkt und muss nicht zeichenhaft sein [60, S. 373].

**Präskription** Eine Vorgabe der nicht-menschlichen Delegierten, welches das Verhalten der Aktanten antizipiert, erlaubt, begünstigt, erzwingt oder verbietet [62, S. 401]. Eine Präskription hat in der Regel die Form eines Imperativs: *Benimm dich auf diese und jene Weise!* Präskriptionen stellen die moralische Dimension von Techniken dar, in dem sie ein bestimmtes Verhalten honorieren oder erschweren [59, S. 232].

**Handlungsprogramm** Ein Handlungsprogramm ist eine »Abfolge von Zielen, Schritten und Intentionen« [62, S. 486] und weist den Aktanten ihre Rollen zu.



**Rückverteilung** Anstelle der Unterscheidung zwischen Mensch und nicht-Mensch bzw. Subjekt und Objekt bildet die Rückverteilung von Kompetenzen und Performanzen den Ausgangspunkt für eine ANT-Beschreibung [62, S. 403].

Um diese Begriffe mit Leben zu füllen, sollen sie Mittels eines knappen Beispiels erläutert werden, welches Latour in ähnlicher Weise präsentiert [59]. Betrachten wir folgende fiktive Situation: Die Außentüren eines Universitätsgebäudes sollen möglichst geschlossen bleiben, damit die Wärme im Gebäude bleibt und der Straßelärm nicht nach innen dringt. Wenn die Angehörigen der Universität diszipliniert wären, würden sie nach dem Betreten oder Verlassen des Gebäudes die Tür hinter sich schließen. Da viele Angehörigen der Universität nachlässig sind, hat die Technische Abteilung mehrere Möglichkeiten, um das Ziel von geschlossenen Türen zu erreichen. Sie kann versuchen, die Angehörigen der Universität durch moralische Appelle zu disziplinieren und hoffen, dass die Tür dann häufiger geschlossen wird. Sie kann aber auch die Aufgabe an einen Pförtner delegieren, der extra für diesen Zweck angestellt wird. Dieser müsste jedoch entsprechend diszipliniert sein und die Universität würde für alle Außentüren viel Geld aufwenden, das sie woanders einsparen müsste. Die Technische Abteilung entscheidet sich für eine ökonomische und effektive Lösung: Sie delegiert die Verantwortung an einen stählernen Nicht-Menschen, einen Türschließer, dessen Feder oder Hydraulik dafür sorgt, dass die Tür nach ihrer Benutzung wieder geschlossen wird. Der Türschließer substituiert und übersetzt die vergeblichen moralischen Appelle der Technischen Abteilung in ein technisches Artefakt. »Haltet die Tür geschlossen!«, wird in einen technischen Mechanismus übersetzt. Mit der Änderung der Ausdrucksmaterie des moralischen Appells ändert sich die Präskription: »Wende zusätzliche Kraft auf, um die Tür zu öffnen!« Ein Türschließer präskribiert in diesem Fall eine zusätzliche Menge an Kraft, die notwendig ist, um die Feder zu spannen bzw. die Hydraulik zu betätigen. Das Skript des Türschließers verhindert, dass die Tür offen stehen kann, was für einige NutzerInnen wünschenswert sein könnte. RollstuhlfahrerInnen, Menschen mit Kinderwagen, Kinder, ältere Menschen, Beschäftigte von Paketdiensten können durch die nun schwer gängigen Türen behindert werden. Der Türschließer ist ein Mittler, welcher auf die Performanzen der NutzerInnen einwirkt und bestimmte Kompetenzen von ihnen erwartet.

Wer oder was handelt in dem Beispiel? Die Technische Abteilung oder eher der Türschließer oder die Firma, die diese herstellt oder doch die Angehörigen der Universität, welche sich an den nun schwer gängigen Türen abmühen? Die Antwort der ANT lautet: Alle sind am Handeln beteiligt – wenn auch in unterschiedlichen Maßen. Die Technische Abteilung delegiert eine Handlung an den Türschließer. Dabei verschiebt sich die Artikulation der Handlung sowohl räumlich, zeitlich, als auch aktorial: Räumlich, weil der Tür ein hydraulischer Mechanismus hinzugefügt wurde. Zeitlich, weil die Handlung

nun Tag und Nacht vom Türschließer ausgeführt wird und aktorisch weil der Türschließer für die Technische Abteilung einsteht, wo sie abwesend ist. Trotz ihrer Abwesenheit vollziehen der Türschließer und die undisziplinierten Angehörigen der Universität eine Handlung im Sinne der Technischen Abteilung.

Die frühen Arbeiten der ANT wurden von feministischer Seite als geschlechterblind kritisiert.<sup>4</sup> Die für diese Arbeit relevanten Konzepte, welche an die ANT anschließen und weiterentwickeln, werden jetzt vorgestellt. Es sind im wesentlichen die Arbeiten der feministischen Wissenschaftlerinnen Lucy Suchman, Donna Haraway (→2.2.3) und Karen Barad (→2.2.4). Die US-amerikanische Technikforscherin Lucy Suchman war 20 Jahre im Xerox Palo Alto Research Center (PARC) im Bereich Mensch-Maschine-Interaktion tätig. Computerisierte Arbeitsumgebungen bilden einen der Schwerpunkte ihrer Forschungen. In ihrer Charakterisierung von *agency* knüpft Suchman an die ANT an:

Agency is understood as a material-semiotic<sup>5</sup> attribute not locatable in either humans or nonhumans. Agency on this view is rather an effect or outcome, generated through specific configurations of human and nonhuman entities [96, S. 261].

Suchman weißt damit die humanistische Position vom Menschen als einzigen Ursprung von Handlung zurück. In ihrer Sichtweise 'besitzt' niemand *agency*. Handlungsfähigkeit als Effekt aufzufassen bedeutet, den Blick auf die jeweils spezifischen Anordnungen von Menschen und Nicht-Menschen zu richten. Handlungen sind damit nicht Ausdruck eines individuellen Willens. Suchman zufolge suggeriert unsere Sprache, dass Personen und Artefakte, bzw. Subjekte und Objekte diskrete, getrennte Entitäten sind, die mit eigener *agency* ausgestattet sind [96, S. 263]. Diese Sichtweise auf Personen und Artefakte beruhe jedoch auf systematischer Ausblendung von Verbundenheit und Abhängigkeit.

### 2.2.3 Narrationen und Figurationen

In der Wissenschafts- und Technikforschung wird davon ausgegangen, dass in allen Gesellschaften sinnstiftende Erzählungen (Narrative) existieren, welche die Annahmen über die Welt, wie sie funktioniert und was in ihr Wert geschätzt wird, verbreiten. Narrative definieren stillschweigend den Horizont der möglichen und akzeptierten Handlungen ihrer Aktanten und weisen damit *agency* zu. Sie überlagern sich dynamisch mit den materiellen, institutionellen, ökonomischen, technischen und kulturellen Formen von Gesellschaft [34,

---

<sup>4</sup> Auf die Kritiken kann an dieser Stelle nicht eingegangen werden. Für einen Überblick sei auf folgende Quellen verwiesen: [98, S. 40 ff.], [101, S. 137 ff.], [67].

<sup>5</sup> →2.2.3

S. 73 ff.]. Narrative legitimieren somit Normen und Handlungen. Sie begrenzen zugleich, was im Rahmen der Erzählung vorstellbar ist und was nicht. Somit gestalten Narrative mögliche Zukünfte. Technologien können auf vielfältige Weise Allianzen mit Narrationen eingehen. Das Narrativ des technischen Fortschritts als prominentes Beispiel assoziiert technische Innovation mit gesellschaftlichem Fortschritt (→4.5.1). Wie Latour schreibt, scheint das Fortschrittsnarrativ in unserem Kulturkreis so evident zu sein, dass es jede Diskussion abwürgt [60]. Es gehört zu den Erzählungen, welche unhinterfragt als wahr gehandelt werden. Narrative stützen sich nicht nur auf Technologien, sondern werden auch durch Technologien verbreitet.

Der Prozess der Figuration (*figuration*) ist eng mit Narrationen verwandt. Gemäß dem Langenscheidt Englisch-Deutsch Wörterbuch bedeutet das Adjektiv *figurative*, dass etwas bildlich, übertragen, metaphorisch oder symbolisch gemeint ist [74]. Die Biologin und Wissenschaftsforscherin Donna Haraway hat dieses Konzept vom Germanisten Erich Auerbach entliehen und für die Naturwissenschafts- und Technikforschung fruchtbar gemacht. Im Gegensatz zu Auerbachs Vorstellung einer mimetischen Praxis müssen Haraways Figurationen nicht notwendigerweise etwas repräsentieren. Sie müssen jedoch im übertragenden Sinne (*tropic*) gemeint sein. Nach Haraway können Figuren (*figures*) nicht wörtlich (*literal*) und selbst-identisch sein. Sie enthielten vielmehr eine Verschiebung, welche Gewissheiten in Frage stelle [43, S. 10 f.].

Figurationen basieren auf Erzählungen [44, S. 1]. Trotzdem funktionieren sie nicht nur als Bestandteil der Rhetorik. Die Germanistin Karin Harrasser schreibt in ihrem Kommentar, Haraway nutze Figurationen, um die »Fragestellung, Konzepte, Methoden, politischen Standpunkt und konkrete Materialitäten« ihrer Untersuchungsgegenstände zu verdichten [45, S. 451]. Beispielsweise analysiert Haraway die transgene Maus *OncoMouse*, die zur Krebsforschung hergestellt wurde, als Figur in weltlichen techno-wissenschaftlichen Erlösungserzählungen (*secular technoscientific salvation stories*) [43, S. 8]. Haraway weitet so Figurationen auf Technologien aus. Suchman erläutert das Konzept der materialisierten Figuration:

Technologies, Haraway argues, are forms of *materialized figuration*; that is, they bring together assemblages of stuff and meaning into more and less stable arrangements [96, S. 227].

Technologien, verstanden als materialisierte Figurationen, vereinen Stoffliches mit Bedeutung. Figurationen sind **materiell-semiotische** (*material-semiotic*) Prozesse. Analog zur ANT arbeitet Haraway mit einer Symmetrie zwischen Materie und Sprache. Das heißt, sie privilegiert in ihren Analysen weder das eine noch das andere als Erklärendes. Dieser Ansatz ermöglicht es, die symbolische und soziale Dimension von technischen

Artefakten zu erfassen, ohne den nur vermeintlich ausschließenden Gegensatz von Technik und Gesellschaft zu wiederholen.

Suchman knüpft an Haraways Analyse von Figurationen an. In ihrem Buch *Human-Machine Reconfigurations* untersucht sie u. a. die Figuration des Menschen in der KI-Forschung und Robotik und wie dabei Mensch und Maschine zusammen konfiguriert werden (vgl. [96]). Laut *Duden* wird der Begriff Konfiguration in der Informatik dazu gebraucht, um die Zusammenstellung von Hardware oder verschiedener Softwareprodukte zu einem System zu bezeichnen [21]. Suchmans Auffassung von Konfiguration geht über die Zusammenstellung von technischen Artefakten hinaus. Sie begreift Konfigurationen als materiell-semiotische, bi-direktionale Prozesse, welche definieren, inszenieren, beschränken und kontrollieren, was Menschen und Nicht-Menschen auszeichnet und voneinander unterscheidet. Die spezifischen Konfigurationen können machtvoll, soziale und materielle Effekte hervorrufen, zum Beispiel durch bestimmte Ein- und Ausschlüsse [97, S. 149-153]. In der vorliegenden Arbeit orientiert sich der Gebrauch von Konfiguration und konfigurieren an Suchman. Die Effekte von Konfigurationen sollen im Zuge dieser Arbeit insbesondere anhand der Kategorie Geschlecht nachvollzogen werden (→2.2.5).

#### 2.2.4 Intra-Aktionen und *agential cuts*

Die Physikerin und Philosophin Karen Barad radikalisiert die Ansätze der ANT. Ihr theoretischer Ansatz einer *posthumanist performativity* schließt an Judith Butlers Konzept der Performativität an [5]. Unter **Performativität** versteht Butler eine Reihe von erzwungenen, fortlaufenden und nicht notwendigerweise bewussten Wiederholungen, welche *das produzieren, was sie sprachlich benennen*. Butler kann mit Performativität erfassen, wie Sprache materielle Effekte haben kann. Barad weitet Butlers Ansatz auf Nicht-Menschen aus. Hierfür führt sie eine eigene Terminologie ein. Zwei dieser Begriffe, Intra-Aktionen (*intra-actions*) und *agential cuts*, sollen nun erläutert werden.

Der geläufigen Verwendung des Wortes Interaktion liegt zu Grunde, dass sich zwei oder mehr bereits formierte Entitäten in einen Austausch begeben. Barad dreht diese Vorstellung um: Abgetrennte und von einander unabhängige Entitäten sind das Ergebnis materiell-semiotischer Prozesse, welche sie **Intra-Aktionen** nennt. Ihr Ausgangspunkt ist die Quantenphysik. In Barads Lesart bilden die Untersuchungsinstrumente, das Untersuchungsobjekt und die Observierenden ein untrennbares Ganzes. Durch materiell-semiotische Grenzziehungen, die sie **agential cuts** nennt, werden eigenständige Entitäten erzeugt. Die Grenzen sind keineswegs 'naturegegeben', sondern das Ergebnis der gemeinsamen Anstrengungen von Sprache und Materie. Abgetrennte Objekte und Subjekte sind nicht die Voraussetzung sondern das Ergebnis der Intra-Aktionen [5, S. 816]. Nur durch die ständige Wiederholung der Grenzen erscheinen sie selbstverständlich und natürlich.

Zum Beispiel ist die Auffassung, dass der menschliche Körper mit seiner Haut endet, ein *agential cut*. Ende des 17. Jahrhunderts kommt es in Europa zu einer gründlichen Neuorientierung der Körpervorstellungen. Zuvor wurden Körper als fließend verstanden – ohne strikte Abgrenzung von Außen und Innen [49, S. 106 ff.].

Suchman greift Barads Ansatz auf, um die Analyse von Mensch-Maschine-Schnittstellen neu zu konzipieren. Sie schlägt in Anschluss an Barad vor, die Schnittstelle als Ort von Intra-Aktionen aufzufassen, an dem Personen und Maschinen als diskrete Entitäten erzeugt werden [96, S. 268]. Die Schnittstelle (*interface*) solle nicht als gegebene Grenze zwischen Mensch und Maschine untersucht werden, sondern als eine inszenierte (*enacted*) Grenze, die ihre Wirkung in den jeweiligen Konfigurationen entfaltet [96, S. 263].

### 2.2.5 Geschlecht und Intersektionalität

In diesem Abschnitt werden die Grundbegriffe für die Analyse von Geschlecht im Rahmen dieser Arbeit dargelegt. Im *Handbuch der Frauen- und Geschlechterforschung* zählen die Informatikerinnen Corinna Bath und Heidi Schelhowe sowie die Soziologin Heike Wiesner vier Forschungsperspektiven der Frauen- und Geschlechterforschung auf:

1. Frauen in der Informatik
2. Vergeschlechtlichung des öffentlichen Bildes, der Fachkultur und Curricula der Informatik
3. Vergeschlechtlichung von Software und Informationstechnologien
4. Internet, Cyberfeminismus und feministische Technikforschung [7]

Die vorliegende Arbeit macht hauptsächlich von der dritten Forschungsperspektive Gebrauch. Bevor im Detail erläutert werden soll, welche Forschungsfragen in dieser Arbeit zum Tragen kommen, soll zunächst die Intersektionalität von Geschlecht erläutert werden. Laut den Techniksoziologinnen Gabriele Winker und Nina Degele geht das Konzept auf die Kritik schwarzer Frauen am weißen Feminismus in den USA zurück [103, S. 11]. Intersektionalität versucht der Tatsache Rechnung zu tragen, dass Menschen nicht nur einem Geschlecht zugeordnet werden, sondern u. a. durch *race*/Ethnizität, Klasse und Körper<sup>6</sup> markiert werden. Menschen gehören immer mehreren Gruppen an und folglich besteht Identität aus mehreren Ebenen. Eine intersektionale Analyse erfolgt daher nicht in den Begriffen Mann und Frau, da die Unterschiede innerhalb der Gruppen zu groß sind.

---

<sup>6</sup> Winker und Degele verstehen unter der Kategorie Körper Herrschaftsverhältnisse im Zusammenhang mit körperlichen Merkmalen wie z. B. Alter, Attraktivität, Generativität und körperliche Verfasstheit [103, S. 51].

Intersektionalität unterstreicht hingegen die Verschränkung von multiplen Herrschafts- und Machtverhältnissen. Geschlecht, *race*/Ethnizität, Klasse und Körper sind immer relational und stehen in einer wechselseitigen Beziehung. Vor dem Hintergrund von Intersektionalität kann Geschlecht nicht getrennt von den anderen Differenzkategorien betrachtet werden.

Geschlecht und so auch *race*/Ethnizität, Klasse und Körper werden in dieser Arbeit als soziokulturelle Konstruktionen aufgefasst. Zahlreiche Untersuchungen legen nahe, dass Geschlecht weder zeitlich noch kulturell unabhängig ist. Eine Pionierin der Geschlechterforschung in Deutschland, die Historikerin Karin Hausen, konnte zeigen, dass die heute verbreitete Annahme von zwei komplementären Geschlechtern erst mit Trennung von Erwerbsarbeit und Familienleben in den westlichen kapitalistischen Gesellschaften aufgekommen ist [46]. Londa Schiebinger und Thomas Laqueur kommen in ihren Untersuchungen der westlichen Medizin zu dem Schluss, dass sich erst Ende des 17. Jahrhunderts die binäre Geschlechterordnung durchsetzen konnte. Bis dahin basierte die europäische Medizin auf einem Ein-Geschlechter-Modell indem Frau und Mann lediglich graduell unterschieden wurden. Frauen galten als 'mindere' Männer, jedoch wurden die Geschlechter nicht in fundamentaler Opposition zueinander gesehen [86] [57]. Ein Blick in nicht-westliche Kulturkreise bestätigt die Zweifel an der verbreitete Vorstellung von der Natürlichkeit der Zweigeschlechtlichkeit. Die feministische Biologin Sigrid Schmitz verweist auf die Sambia in Papua Neuguinea, die Berdache und die Navajos in Nordamerika, welche mehr als zwei Geschlechter kennen und leben [87, S. 33]. Insbesondere erscheint die Zweigeschlechtlichkeit vor dem Hintergrund der Intersexualität als soziokulturelle Konstruktion (vgl. [33]). Von dem Phänomen Intersexualität wird gesprochen, wenn Körper genetisch, anatomisch und hormonell nicht eindeutig entsprechend den herrschenden Normen einem Geschlecht zugeordnet werden können.

Geschlecht als performativ zu begreifen bedeutet, den Geschlechtskörper nicht als Ursache oder Basis für Vergeschlechtlichung zu sehen sondern als Materialisierung von Geschlechter-Normen. Der Geschlechtskörper ist nicht Ausdruck von zu Grunde liegenden Differenzen, seien sie nun biologisch, psychologisch oder soziologisch begründet. Butler argumentiert, dass die Normen und Attribute vielmehr konstitutiv für die Geschlechter-Identitäten sind, welche sie nur scheinbar ausdrücken [16, S. 192]. Geschlecht ist somit ein Effekt von Institutionen, Sprache und Praktiken. Bath schlägt vor, die Vergeschlechtlichung informatischer Artefakte theoretisch in Anschluss an Suchman und Barad zu begreifen. Technologien und Geschlecht werden in einem Prozess gleichzeitig durch zitathafte Wiederholungen hervorgebracht. Informatische Artefakte können so als »Materialisierungen im Sinne einer mehr oder weniger umstrittenen normativen Technikgestaltung« verstanden werden [6, S. 239]. Technikgestaltung wird in diesem Zusammenhang nicht als individueller Akt von EntwicklerInnen aufgefasst, sondern als

Effekt von soziotechnischen Anordnungen. Die Techniksoziologin Judy Wajcman fasst das Verständnis des Geschlechter-Technik-Verhältnisses in der gegenwärtigen feministischen Technikforschung wie folgt zusammen:

gender relations can be thought of as materialized in technology, and masculinity and femininity in turn acquire their meaning and character through their enrolment and embeddedness in working machines [98, S. 107].

Informatische Artefakte sind sowohl Ursache als auch Effekt von Geschlecht. Umgekehrt erhält Geschlecht seine Bedeutung erst durch die Verschränkung mit Technologien.

In dieser Arbeit werden die Ko-Materialisierungen von VPAs, als informatische Artefakte, und Geschlecht unter Berücksichtigung der Wechselwirkungen mit den o. g. Differenzkategorien untersucht. Daraus ergeben sich die drei Schwerpunkte:

- (1) Geschlechter-Performanzen in den Intra-Aktionen
- (2) Einschreibungen der geschlechtsspezifischen Arbeitsteilung
- (3) Materiell-semiotische Ein- und Ausschlüsse

Zu (1): Auf dieser Ebene werden die Geschlechter-Performanzen der VPA in der Intra-Aktionen mit der Nutzerin analysiert. Gemäß der Kulturwissenschaftlerin Alexandra Chasin geht die Anthropomorphisierung eines informatischen Artefakts zwangsläufig mit einer Vergeschlechtlichung einher, da die Performanz des Menschlichen immer durch Geschlecht, aber auch *race*/Ethnizität und Klasse markiert ist [19, S. 75]. Anthropomorphe AgentInnen wiederholen wie ihre menschlichen Gegenüber zitathaft Geschlechter-Normen.

Zu (2): Der Begriff **geschlechtsspezifische Arbeitsteilung** beschreibt den Zustand, in dem bestimmte Tätigkeiten und Berufe von Angehörigen eines bestimmten Geschlechts ausgeübt werden. Die Einschreibung der geschlechtsspezifischen Arbeitsteilung erfolgt auf zwei Ebenen: Erstens auf der narrativen Ebene der Figurationen der VPAs und zweitens auf der Ebene der Tätigkeiten, die durch die VPAs automatisiert werden sollen. Letzteres betrifft auch die Frage, welche menschlichen Arbeiten für das Funktionieren der VPAs nötig sind. Die übergeordnete Frage lautet, inwiefern die VPAs dazu beitragen, die geschlechtsspezifische Arbeitsteilung zu (de-)stabilisieren?

Zu (3): Ein- und Ausschlüsse werden in der vorliegenden Arbeit entlang der Präskriptionen (→2.2.2) und der Konzepte analysiert. Der Informationswissenschaftlerin Susan Leigh Star zufolge passen die Bedürfnisse von Individuen nie ganz zu einem technologischen System [93, S. 36]. Indem VPAs bestimmte Bedürfnisse befriedigen können und andere nicht, verkörpern sie Normen und Werte. Die »behind-the-scenes decisions« [94, S. 110], bei denen die EntwicklerInnen und DesignerInnen bewusst oder unbewusst Ein- und Ausschlüsse vornehmen, werden mit der Einschreibung in die Artefakte

in der Regel unsichtbar. Als Ausgangspunkt für eine solche Untersuchung schlägt Star die Perspektive der *non-user* vor. Der nicht-Nutzer ist die Perspektive, die im Sinne der ANT nicht übersetzt werden kann. Umgekehrt können durch Einschluss vergeschlechtlichte NutzerInnenbilder eingeschrieben werden.

Die begrifflichen Konzepte mit denen VPAs in der KI- und MCI-Forschung beschrieben werden, nehmen auf symbolischer Ebene Ein- und Ausschlüsse vor. Begriffe und Handlungen, die Mitbedeutungen in Bezug auf Geschlecht haben, heißen geschlechtlich konnotiert. Konnotationen sind in der Linguistik Sekundärbedeutungen, Mitbedeutungen und Begleitvorstellungen von Worten und Praktiken [38]. Ausschlüsse auf symbolischer Ebene können soziale Ausschlüsse zur Folge haben.



## 3 Virtual Personal Assistants

In diesem Abschnitt wird der Kontext für die Analyse im Kapitel 4 erläutert. Zunächst erfolgt eine Eingrenzung von dem, was in dieser Arbeit unter *virtual personal assistants* verstanden wird, und ein Vergleich von zwei Konzeptionen aus der Informatik (→3.1). Anschließend wird mit Apples *Knowledge Navigator* jenes visionäre Video beschrieben und analysiert, welches die Entwicklung von VPAs maßgeblich beeinflusst hat (→3.2). Danach erfolgt ein kurzer Überblick über aktuelle Implementationen von VPAs (→3.3). Zum Abschluss des Kapitels werden die Grundannahmen des Interaktionsparadigma untersucht und mit Ansätzen aus der Geschlechter- und Technikforschung gegengelesen (→3.4).

### 3.1 Konzeptionen

Die Informatik verwendet häufig Berufe als Bezeichnungen für Soft- und Hardware. Der Literaturwissenschaftlerin Katherine Hayles zufolge waren Computer in den 1930er und 1940er die gängige englische Bezeichnung für weibliche Rechenkräfte [48, S. 1]. Die neue Bedeutung hat in diesem Fall die ursprüngliche Bedeutung verdrängt. Das Substantiv AssistentIn ist abgeleitet aus vom lateinischen Verb *assistere* und kann mit »beistehen« übersetzt werden. Ein Assistent ist dazu bestimmt, jemanden zu unterstützen. Das englische Wort *assistant* hat laut dem *Oxford English Dictionary* auch die Bedeutung »the second« (der/die Zweite). Ein Assistent bzw. eine Assistentin befindet sich in einer Hierarchie zum 'ersten'. Die Figur des Assistenten steht für jemanden, der einer übergeordneten Person zuarbeitet oder sie unterstützt. Die tatsächlichen Tätigkeiten können sehr unterschiedlich sein, je nachdem ob es sich um eine Assistenz in der Wissenschaft oder Wirtschaft handelt oder um die Assistenz von Menschen mit Behinderungen.

Als *virtual personal assistants* werden sowohl eine bestimmte Klasse von Software-AgentInnen als auch LeiharbeiterInnen bezeichnet, die virtuell – folglich 'nicht vor Ort' – tätig sind. Letzteres ist eine aufstrebende Branche, bei der LeiharbeiterInnen, oft in Ländern mit niedrigerem Lohnniveau, Menschen und Firmen in reicheren Ländern über das Internet oder Telefon, virtuell unterstützen. Unternehmen wie *getFriday* und *askSunday* bieten Privatpersonen und Firmen die Dienste ihrer VPAs an. Bei *getFriday* lauten typische Aufgaben: HandwerkerInnen organisieren, Datenbanken pflegen, Bücher

und CDs kaufen, Termine absprechen und Reisen organisieren [37]. Softwaregestützte VPAs werden häufig auch unter anderen Bezeichnungen verhandelt.<sup>1</sup> Im Unterschied zu VPAs handelt es sich bei *virtual assistants* um eine interaktive Hilfe, die beispielsweise auf Unternehmenswebsites anzutreffen sind. Dieser Typ von virtuellen AssistentInnen ist in der Regel nicht personalisiert.

In der Informatik gibt es keine einheitliche Definition von VPAs. Jede Charakterisierung setzt unterschiedliche Schwerpunkte in Bezug auf Aufgabe, Funktionsweise und Integration der VPAs. Aus diesem Grund sollen im folgenden zwei Konzeptionen exemplarisch mit einander verglichen werden. Die Informatikerin Patti Maes vom MIT Media Lab verwendet in ihrem 1997 veröffentlichten Artikel den Begriff »personal assistant« (PA), um Programme zu bezeichnen, die mit der Nutzerin in der gleichen Arbeitsumgebung kollaborieren [70, S. 145f.]. Zweck eines PAs ist es, die Arbeitsbelastung des Nutzers zu reduzieren und ihn vor dem »information overload« zu bewahren. Aus diesem Grund soll der Agent die Vorlieben, Interessen und Gewohnheiten der Nutzerin erlernen. Schwierige Aufgaben könnten in ihrer Komplexität vor dem Nutzer verborgen werden oder anstelle der Nutzerin vom Agenten erledigt werden. Darüber hinaus sei der PA in der Lage, die Nutzerin in der Anwendung zu schulen und die Kollaboration zwischen verschiedenen NutzerInnen zu unterstützen. Ein PA überwache Ereignisse und Prozesse im Hintergrund. Als Beispiel für einen PA nennt Maes einen E-mail- und News-Client, welcher die Informationen für die Nutzerin filtert.

Tom Gruber (2009) und Dag Kittlaus (2010) konzipieren VPAs als Schnittstelle für mobile Endgeräte, welche die Internet-Nutzung verbessern sollen. Sie vergleichen ihr Konzept mit dem einer herkömmlichen Internet-Suchmaschine. In ihrer Lesart präsentieren Suchmaschinen nur Ergebnisse und können jedoch nicht Aufgaben für die Nutzerin erledigen (»get things done«). Mit *things* meinen sie Reservierungen, Buchungen und Bestellungen über das Internet. VPAs seien dagegen eher eine »Do Engine« als eine »search engine« (Suchmaschine) [53]. »Task delegation« oder »Task completion« sollen die Suchmaschinen als Schnittstelle zum Internet ablösen. Auch sie führen Arbeitserleichterung und Zeitersparnis als wesentliche Motivation für VPAs an. Unterschiedliche Internetdienste sollen zentral durch einen VPA genutzt werden können, ohne dass die Nutzerin viele Portale durchsuchen muss. Stattdessen werden die vorkonfigurierten Dienste durch den VPA genutzt. Die BenutzerInnen-Schnittstelle eines VPAs unterscheidet sich fundamental von einem gewöhnlichen Webbrowser. Eine dialogbasierte Schnittstelle mit kontinuierlicher Spracherkennung und -verstehen ersetzt die klassischen

---

<sup>1</sup> Schnittstellen-AgentInnen, *embodied cognitive assistants*, *virtual personal assistants*, *indirect management*, *personal assistant*, *personal internet assistant*, *intelligent software assistant*, *software personal assistant* oder schlicht Software-Agent.

*point-and-click*-Operationen in einem Webbrowser. Analog zum Konzept von Maes sollen persönliche Informationen dazu dienen, dass die Angebote und Suchergebnisse für die Anwenderin maßgeschneidert sind.

Beide Konzeptionen von VPAs versprechen Arbeitserleichterungen für den Nutzer. Bei Maes liegt der Schwerpunkt in der Imitation der Anwenderin mit dem Ziel, ihre Interessen zu erlernen. Eine Bibliothekarin oder ein Buchhändler dienen hier als Vorbilder. Mit der Zeit werde der VPA wie seine menschlichen Vorbilder die Vorlieben der Nutzerin einschätzen können. Gruber/Kittlaus richten den Fokus auf die Reduzierung des Rechercheaufwands beim Konsum von Internet-Dienstleistungen. Die Personalisierung erfolgt entweder aktiv mit der expliziten Angabe von Präferenzen durch den Nutzer oder passiv indem der VPA persönliche Informationen aus Kalendern und Mails berücksichtigt. Während der Agent von Maes visuell im Hintergrund tätig ist und auch umgangen werden kann, ist der Agent von Gruber/Kittlaus als Mittler zwischen Internetdiensten und Nutzerin im Vordergrund präsent.

## 3.2 Knowledge Navigator

Als maßgebliches Vorbild für heutige VPAs soll nun der *Knowledge Navigator*, im folgenden als KN abgekürzt, diskutiert werden. Jahre vor dem *world wide web* schildert 1987 der damalige Vorstandsvorsitzende von Apple, John Sculley, seine Vision vom »Macintosh der Zukunft, den es Anfang des 21. Jahrhunderts geben wird« [89, S.397]. Der KN wird als ein mobiler Computer beschrieben, der in in der Lage sein sollte, über vernetzte Computer Informationen aus »Bibliotheken, Museen, Datenbanken und Regierungsdokumenten« zu durchsuchen und zu interpretieren, um so für die NutzerInnen personalisiertes Wissen zu präsentieren.

Basierend auf Sculleys Vision ließ Apple einen aufwendigen Videofilm [28] produzieren. Der Informatiker Jared Spool geht heutzutage davon aus, Form, Struktur und Inhalt seien kein bloßer Zufall sondern Ergebnis eines intensiven Arbeitsprozesses der *Creative Services* von Apple und der Filme-MacherInnen. Das Video habe einen nachhaltigen Effekt für Apples Produktentwicklung gehabt [92]. Wie die zahlreichen Referenzen auf den KN beweisen, beeinflusste er die Forschung der KI und MCI. Insbesondere jedoch verweisen die EntwicklerInnen von Siri (→4) in Artikeln, Interviews und Vorträgen immer wieder auf den KN. Aus diesem Grund bedarf es einer genaueren Analyse des Videos.

### 3.2.1 Szenario

Die Handlung, in die das Szenario einbettet ist, soll nun zusammengefasst werden. Anschließend werden die Struktur, die Form und die ProtagonistInnen des Videos

analysiert. Die ProtagonistInnen des Szenarios sind ein Geographie-Professor, dessen Mutter, eine Kollegin, ebenfalls Geographie-Professorin, und »Phil«, der Avatar und Schnittstellen-Agent des KN. Das Video beginnt mit einem Kamera-Schwenk über einen rustikalen Schreibtisch, auf dem gerahmte Schwarz-Weiß-Photos, verschiedene Dokumente und ein futuristisch anmutendes Notebook, der *Knowledge Navigator*, liegen. Dazu ertönt eine klassische Symphonie. In der folgenden Einstellung wird das Büro in der Totale gezeigt. Es ist ein geräumiges, historisch anmutendes Studierzimmer mit Bücherregalen und einem Globus. Ein Professor betritt das Büro und schlägt das Notebook auf. Die Musik wird ausgeblendet, statt dessen hört die Zuschauerin die Stimme von Phil, dem Schnittstellen-Agenten des KN. Phil informiert den Professor über neue persönliche Nachrichten und seinen Tagesplan. Der Professor setzt sich an seinem Schreibtisch. Er fordert Phil auf, ihm seine Notizen für die am Nachmittag bevorstehende Geographie-Vorlesung zu zeigen. Der Professor möchte die Notizen aktualisieren und lässt Phil neuere wissenschaftliche Artikel suchen und sie mündlich zusammen fassen. Anschließend hält er eine Video-Konferenz mit einer Kollegin von einer entfernten Universität ab, um gemeinsam die Abholzung des Amazonas-Regenwaldes zu diskutieren. Der KN berechnet eine visuelle Simulation über die zukünftige Entwicklung des Baumbestandes. Nach dem Abschluss seiner Vorlesungsvorbereitung geht der Professor zum Mittagessen. Phil wünscht ihm einen guten Appetit (»Enjoy your lunch!«). Nachdem der Professor sein Arbeitszimmer verlassen hat, hört die Zuschauerin ein Telefonklingeln. Phil meldet sich und entschuldigt die Abwesenheit des Professors. Es ist die Mutter des Professors, die ihn an eine wichtige Aufgabe erinnern möchte. Ihre Anruf wird ausgeblendet und das Video endet mit der Einblendung des Textes »The Power to be your best«.

### 3.2.2 Analyse

Die Analyse der Struktur des Videos widmet sich den Fragen, wie die Handlung aufgebaut ist und wie sie erzählt wird. Der Zuschauerin wird zunächst ein klischeehafter Arbeitsplatz eines Wissenschaftlers präsentiert. Das Video bettet auf diesem Weg ein innovatives Narrativ, den *Knowledge Navigator*, in ein konservatives Narrativ ein. Die aufgezählten Gegenstände sollen dem wissenschaftlich ausgebildeten Zuschauer das Gefühl von Vertrautheit vermitteln. Obwohl das Video 20 Jahre in der Zukunft spielt und an das *science fiction*-Genre<sup>2</sup> anknüpft, werden bewusst die dystopischen Elemente dieser Erzählform vermieden.

Die im Video präsentierten Technologien waren zum Zeitpunkt der Produktion weit

---

<sup>2</sup> Jahre vor der Produktion des Videos wurden ähnliche, anthropomorphe Schnittstellen-AgentInnen im *science fiction*-Genre inszeniert, beispielsweise in Gestalt von HAL 9000 (1968) aus dem Film und Roman *2001: A Space Odyssey* [20] und von KITT (1982) aus der Fernsehserie *Knight Rider* [58].

vom technisch Möglichen entfernt. Die menschlichen ProtagonistInnen äußern jedoch an keiner Stelle Erstaunen oder Verwunderung. Der KN wird als ein völlig alltäglicher Gebrauchsgegenstand verhandelt. Anders als es von einer IT-Firma wie Apple erwartet werden könnte, werden die Hintergründe und Komponenten der Technik nicht erläutert. Das Video ist im Gegenteil so angelegt, dass es den Zuschauer kontinuierlich von der Technik ablenkt, indem es auf den Professor und seinen Software-Assistenten Phil, fokussiert. Hugh Dubberly, der zusammen mit Doris Mitsch das Drehbuch für das Video verfasst hat, begründet, warum auf explizite Details über die verwendete Technik verzichtet wurde:

Lastly, I must note what Alan Kay once told me – that the technology shown is not the issue in these videos. The issue is the social use of the technology – what it does to and for people [11, S. 1585].

Für Dubberly sind Videos dieser Art ein Instrument des *prototyping*. Streng genommen handelt es sich bei dem KN um ein *mock-up* (Attrappe), weil die Funktionalitäten nur mit Hilfe der Schnitttechnik simuliert wird. Nichtsdestotrotz lassen sich einige technische Aspekte des KN identifizieren, da sie mittlerweile zur Standardausstattung von Laptops gehören. Das Gehäuse des KN, ähnelt einer Mischung aus Notebook und *tablet* (Abbildung 3.1). Es kann ausgeklappt werden, sodass sich ein farbiger Flachbildschirm unterbrechungsfrei auf beide Seiten erstreckt. Wie bei vielen gegenwärtigen *smart phones* und *tablets* ist der Bildschirm des KN berührungssensitiv. Der Professor kann mit seinen Fingern Objekte auf dem Bildschirm verschieben und durch Menüs navigieren. Wie bei heutigen Laptops üblich, sind Kamera und Lautsprecher in dem Rahmen des Gehäuses integriert. Die MacherInnen des Videos gestalteten den KN ohne Tastatur, so wie es bei *tablets* verbreitet ist. Obwohl es zum Zeitpunkt der Produktion noch kein WLAN gab, benötigt der KN kein Kabel, um mit anderen Computern zu kommunizieren. Der KN bietet Möglichkeiten zur Kollaboration zwischen entfernten NutzerInnen, wie zum Beispiel das *desktop sharing* und die Video-Telefonie. Während der Video-Konferenz kann die Kollegin eine animierte Simulation direkt auf seinem Bildschirm vorführen, ohne dass der Professor dafür selbst aktiv werden muss.

Das Video legt dem Zuschauer nahe, dass bestimmte Technologien in Zukunft möglich sind. Gleichzeitig transportiert es damit bestimmte Grundannahmen und Werte auf deren Basis die gezeigten Technologien vorstellbar werden. Vor allen Dingen suggeriert das Video, dass eine menschenähnliche Bedienungsschnittstelle möglich ist, die ohne erkennbare Probleme versteht, was eine Nutzerin möchte. Ferner legt das Video nahe, dass Wissen in einer stabilen und Maschinen-lesbaren Form existiert und dass aus wissenschaftliche Studien in Echtzeit animierte Simulationen erstellt werden können, die eine objektive Wirklichkeit repräsentieren. Die Sorge um den Amazonas-Regenwald gibt

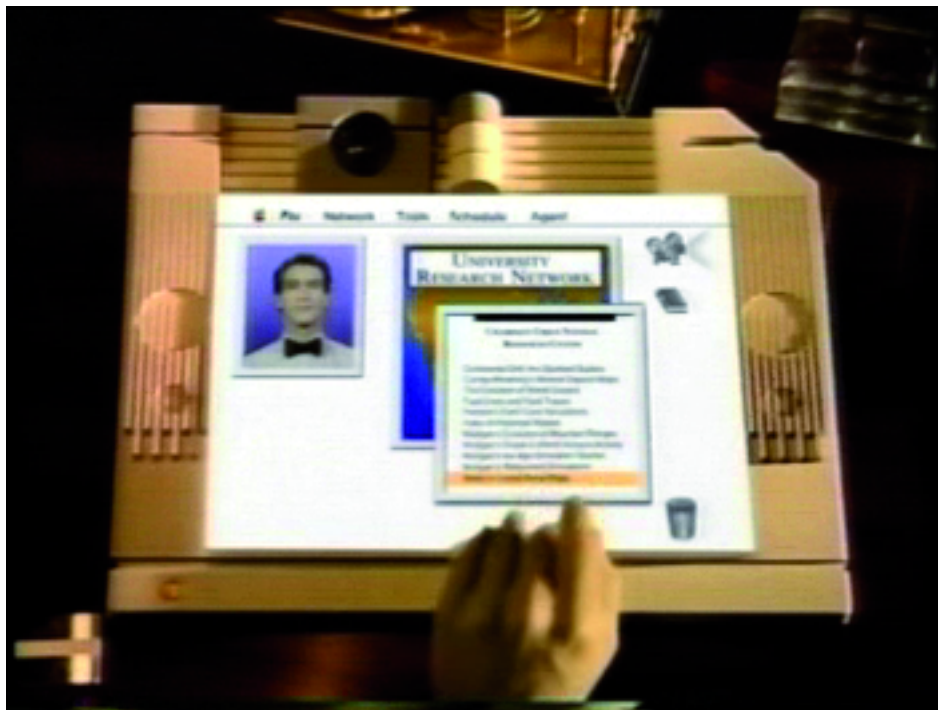


Abbildung 3.1: Standbild von dem *mock-up* des *knowledge navigator* aus [28]

einen verantwortungsvollen Umgang mit der Natur vor und wertet sowohl die Technologie als auch ihre Nutzer auf. Mit dem KN kann der Amazonas-Regenwald beherrscht werden, ohne den lokalen Gegebenheiten und Menschen Rechnung tragen zu müssen. Das Video knüpft damit an koloniale Phantasien und Motive an. Der Name »Navigator«, die zentrale Position des Globus und die Art und Weise wie die Sorge um den Regenwald inszeniert wird, legitimieren die Herrschaft des Westens über die ehemaligen Kolonien.

Die Handlung wird ausschließlich aus der männlichen Perspektive des Professors erzählt. Die weiblichen Protagonistinnen kommen nur als Randfiguren vor und sind im Falle der Mutter unsichtbar. Die Technikforscherinnen Nelly Oudshoorn, Els Rommes und Marcelle Stienstra weisen auf die männliche Konnotation von Technik hin:

Technology is one of the most stable and powerful symbols of masculinity. Technological competency has become an important aspect of hegemonic masculinity [79, S. 54].

Die Männlichkeit des Professors zusammen mit der exklusiven Technik wiederholen so die traditionelle Geschlechterordnung. Die Kollegin vom Professor nutzt zwar vermutlich auch einen KN. In der Inszenierung erscheint sie jedoch passiv. Sie ist diejenige die angerufen wird und welche nur in Form eines Live-Bildes gezeigt wird. Der Professor wird hingegen dabei gezeigt, wie er aktiv mit dem KN arbeitet und spontan während ihrer Videokonferenz eine animierte Simulation berechnen lässt. Der Nexus zwischen

Männlichkeit und Technikkompetenz steht im Vordergrund des Videos.<sup>3</sup> Der Anruf der Mutter bewegt sich im Rahmen einer traditionellen Geschlechterordnung. Sie verwendet eine veraltete Technologie, das Telefon, um mit ihrem Sohn in Kontakt zu treten. Sie sorgt sich um familiäre Beziehungen, einem weiblich-konnotierten Bereich. Der Professor empfindet seine Mutter als lästig, da er den direkten Kontakt mit ihr vermeidet und die Kommunikation an den VPA delegiert.

Der Star des Videos ist die dialogbasierte Schnittstelle des KN: der Software-Agent Phil. Die kontinuierliche Spracherkennung, das Sprachverstehen und das Gespür für die Gesprächs- und Arbeitssituation lassen ihn menschlich wirken. Der Software-Agent Phil versteht alles, was der Professor sagt, weiß wann er angesprochen ist, wartet geduldig auf Gesprächspausen, bevor er etwas sagt, und spricht flüssiges Englisch. Phil kann Daten aus wissenschaftlichen Artikeln in Echtzeit interpretieren und daraus animierte Simulationen erstellen. Phil ist nicht nur reaktiv, er ist ein proaktiver Agent, der entscheiden kann, wann und aus welchen Gründen er seinen Nutzer unterbrechen kann. Mit Phil wird angedeutet, wie weit Personalisierung gehen könnte. Er weiß um die persönlichen Beziehungen des Professors und dessen Werte. Beispielsweise erfährt jener erst nach der Videokonferenz, dass seine Mutter wieder angerufen hat.

Der Avatar Phil ist anthropomorph. Mit seiner großen schwarzen Fliege und dem weißen Hemd dargestellt, weckt er Assoziationen an einen englischen Butler oder einen Kellner. Dies steht im Gegensatz zu seinen Tätigkeiten, die ihn eher als wissenschaftlichen Assistenten, Sekretär und Bibliothekar positionieren. Zusammen mit dem geräumigen Büro und dem formal gekleideten Phil wird der KN als eine exklusive Technologie für weiße, männliche, gebildete Angehörige der gehobenen Mittelschicht inszeniert. Der Wissenschaftler im Video steht stellvertretend für alle wissensintensiven Branchen, in denen lebenslanges Lernen Voraussetzung für einen Arbeitsplatz ist. Nach Sculley ist es die Aufgabe des KN, »die ökonomische Bedeutung der wohlhabenden Mittelschicht Amerikas zu schützen« [89, S. 403]. In diesem Kontext steht die abschließende Botschaft des Videos. »The Power to be your best« unterstellt der Zuschauerin den Willen zur Arbeit an sich selbst, was in diesem Fall die selbstständige Fortbildung und Qualifizierung für den Beruf meint. Die Botschaft basiert so auf einem Narrativ der Aktivierung und Effektivierung des Menschen.

---

<sup>3</sup> Das Demonstrationsvideo *Starfire* von Sun Microsystems bricht hingegen mit der Verknüpfung von Männlichkeit und Technik, indem weibliche Protagonistinnen das innovative Computersystem ausreizen. Bedauerlicherweise zitiert es im Gegenzug viele weibliche Stereotype.

### 3.3 Aktuelle Implementationen

Wie in 2.1.2 erläutert, ist das Konzept der indirekten Steuerung von Computern durch Schnittstellen-AgentInnen bereits in den 1950er Jahren entwickelt worden. Zu einer massenhaften Anwendung ist es bisher nicht gekommen. Ein berühmtes Beispiel stellt der ehemalige Hilfe-Assistent von Microsoft Office dar. »Clippy« – in der deutschen Version hieß er »Karl Klammer« – wurde von vielen AnwenderInnen eher als unangenehm, denn als hilfreich empfunden [107, S. 384]. Während im folgenden Kapitel der VPA Siri im Detail untersucht wird, sollen in diesem Abschnitt zwei VPAs vorgestellt, in dessen Kontext Siri steht: Zum einem Siris direkter Konkurrent der Rearden Personal Assistant und zum anderen Siris Vorläufer CALO.

Einer der wenigen kommerziell erfolgreichen VPAs ist der Rearden Personal Assistant. Der VPA wird von 7300 Unternehmen (Stand Mai 2011) eingesetzt, unter ihnen namhafte Großkonzerne wie Coca-Cola, AT&T und Siemens. Er ist spezialisiert auf die Abwicklung von Geschäftsreisen. Der VPA integriert die Buchung von Flügen, Hotels und Mietwagen. Darüber hinaus kann er den Zeitplan organisieren und einen Teil des Kundenkontakts abwickeln. Der Rearden Personal Assistant bietet außerdem die Möglichkeit, Unterhaltungsangebote zu buchen und Tische in Restaurants zu reservieren. Auf der Internet-Präsenz des Unternehmens findet sich folgendes Szenario:

Imagine you are on a business trip and your computer discovers that your flight will be late. It automatically reschedules your dinner in New York, informs your three guests of the change and tells you they've been notified [22].

Die Entwicklung von VPAs wird intensiv durch militärische Forschungsgelder vorangetrieben. Die dem US-amerikanischen Verteidigungsministeriums unterstellte Forschungsbehörde DARPA (*Defense Advanced Research Agency*) hat dieses Feld intensiv gefördert. Von 2003 bis 2008 lief die Forschungsinitiative *Personal Assistant that Learns (PAL)*. In deren Rahmen wurde das Forschungsprojekt *Cognitive Assistant that Learns and Organizes (CALO)* mit einem Volumen von \$200 Millionen finanziert. Es ist damit das größte KI-Einzelprojekt in der Geschichte der Behörde.

Ziel von CALO ist die Entwicklung von VPA-Prototypen, die BefehlshaberInnen und Stab im Militär unterstützen sollen. Als Anwendungsszenario dienen Befehlsstände (»command centers«), in denen heute Hunderte von Angestellten benötigt werden, um einigen wenigen EntscheidungsträgerInnen zuzuarbeiten. VPAs sollen den Personalaufwand reduzieren, indem sie von den Computer-NutzerInnen lernen und Routinearbeiten automatisieren. Darüber hinaus sollen sie in der Lage sein, eigene Schlussfolgerungen zu ziehen und robust auf unerwartete Situationen zu reagieren. CALO soll den Personalaufwand reduzieren, um so kleinere, mobilere und weniger verwundbare Befehlsstände zu



ermöglichen.

DARPA's Engagement in der Künstlichen Intelligenz-Forschung hat Tradition. Gemäß dem Informationswissenschaftler Paul N. Edwards ist sie der größte Geldgeber der Künstlichen-Intelligenz-Forschung seit den frühen 1960er Jahren [30, S.193]. Indem sie bestimmte Projekte finanziert und andere nicht, ist die Behörde in der Lage, Richtung und Intensität der KI-Forschung maßgeblich zu bestimmen. CALO ist nur eines der jüngeren Beispiele dafür, wie die Künstliche Intelligenz zur Automatisierung von Arbeitsabläufen im Militär dient. So förderte die DARPA etwa von 1983 bis 1993 im Rahmen der *Strategic Computing Initiative* KI-Forschungen für militärische Anwendungen mit insgesamt \$900 Millionen [30, S.193]. Die Entwicklung von Informationstechnologien und die Geschichte der Behörde sind eng verwickelt. Gleichzeitig sponsert und beeinflusst die DARPA mit ihren Forschungsgeldern die US-amerikanische IT-Industrie. Der VPA Siri, welcher im folgenden Kapitel im Detail beschrieben wird, basiert auf den Forschungen im CALO-Projekt.

### 3.4 Mensch/Maschine-Rekonfigurationen

In ihren Charakterisierungen von VPAs greifen Gruber/Kittlaus und Maes auf eine für die KI-Forschung exemplarische Konzeption vom Menschen zurück. In diesem Abschnitt soll das AgentInnen-Paradigma mit Ansätzen aus der Geschlechter- und Technikforschung gegengelesen werden. Zunächst soll geklärt werden, was in der KI-Forschung mit Delegation gemeint ist. Obgleich die beschriebenen Tätigkeiten stark voneinander abweichen, stimmen Kittlaus/Gruber und Maes grundsätzlich in der Konzeption der Delegation überein. Delegation bedeutet für sie, dass eine Tätigkeit, die eine Nutzerin nicht selbst erledigen möchte, als Aufgabe an einen Agenten übertragen wird. Der Agent löse selbstständig die Aufgabe, indem er sie anstelle der Nutzerin ausführt. Nun soll diese Konzeption mit Erkenntnissen der Technik- und Geschlechterforschung kontrastiert werden.

VPAs oder allgemeiner Software-AgentInnen, als Beispiele von 'intelligenten' Maschinen, rekonfigurieren die Beziehung von Mensch und Maschine. Suchman zufolge standen in den vergangenen 300 Jahren Mensch und Maschine wechselseitig als Modell füreinander ein. Eigenschaften, die typischerweise der einen oder anderen Seite der Dichotomie Mensch/Maschine zugeschrieben wurden, dienten als Abgrenzung zum jeweiligen anderen [96, S. 229]. Nur durch ständige Wiederholungen verfestigte sich der Eindruck einer notwendigen und gegebenen Unterscheidung. Das bedeutet nicht, dass Mensch und Maschine prinzipiell gleich sind. Vielmehr ändern sich stetig die Bedingungen, was eine Maschine oder einen Menschen konstituieren. Ein solcher *agential cut* wird bei der

Gegenüberstellung von *search engine* (Suchmaschine) und *Do-Engine* (VPA) bei Kittlaus deutlich [53]. Die Grenze wird entlang der Frage gezogen, welchem informatischen Artefakt *agency* zugebilligt wird. In seiner Darstellung ist eine *do-engine* in der Lage, selbstständig eine Tätigkeit zu erledigen und eine Suchmaschine nicht. Kittlaus hebt mit der Delegation an einen VPA die maschinelle Aktivität hervor, wohingegen er bei der Benutzung einer Suchmaschine den menschlichen Anteil betont. Im Gegensatz zur Suchmaschine schreibt Kittlaus dem VPA *agency* zu. In beiden Charakterisierungen blendet er aus, dass die jeweiligen Aktionen nur durch die gemeinsame Anstrengung von Menschen und Maschinen ermöglicht werden. Die Dichotomien Mensch/Maschine und Suche/Delegation verbergen jeweils einen Teil der soziotechnischen Akteurs-Netzwerke.

Wie die Technikforscherin Alison Adam bemerkt, ermöglicht die Bezeichnung eines informatischen Artefakts als Agentin, respektive als VPA, dass ein informatisches Artefakt mit einem intentionalen Vokabular beschrieben werden kann [1, S. 64]. Die Benennung eines Programms als Agent oder VPA billigt ihm *agency* zu. Die Charakterisierung in Begriffen wie 'Zielorientierung' oder 'Interaktion mit der Umwelt' (→2.1.3) inszeniert Software-AgentInnen als menschenähnliche Individuen, die als abgetrennte Entitäten in einem Computersystem wahrgenommen werden sollen.

Abschließend soll erläutert werden, warum diese Vorstellung von Delegation aus den Perspektiven der Technik- und Geschlechterforschung problematisch ist. Sie beruht auf der Vorstellung des Menschen als rationalem, autonomen Subjekt (→2.2.2). Seit der Aufklärung gilt Autonomie als auszeichnende Eigenschaft von Menschen. Gemäß Crawford Brough Macphersons Theorie vom *possessive individualism* lauten die wesentlichen Charakteristika der westlich-aufklärerischen Autonomie des Individuums:

Its possessive quality is found in its conception of the individual as essentially the proprietor of his [*sic!* G.B.] own person or capacities, *owing nothing to society for them*. [...] The human essence is *freedom from the wills of others*, and freedom is a function of possession (zitiert nach [47, S. 3], Hervorhebungen wie im Original).

Aufgrund der possessiven Eigenschaften interpretiert Macpherson das westlich-aufklärerische Individuum als Produkt der kapitalistischen Ökonomie. Dieses Individuum ist nicht nur im o.g. Zitat männlich. In der symbolischen Geschlechterordnung wird traditionell Männlichkeit mit Autonomie und Weiblichkeit mit Fremdbestimmung identifiziert [52, S. 102 ff.].

Wie Adam, Hayles und Suchman in ihren Untersuchungen zur KI-Forschung feststellen, wird die westlich-aufklärerische Konzeption von Autonomie als Merkmal intelligenter Maschinen (Roboter und Programme) verwendet, um Maschinen menschlicher werden zu lassen (vgl. [1], [47], [96]). Die Dichotomie Mensch/Maschine 'belebe' die Projekte

der intelligenten Maschinen, so Suchman, weil RoboterInnen und AgentInnen nur dann als 'intelligent' anerkannt würden, wenn sie menschenähnlich inszeniert werden. Dabei geht die KI-Forschung davon aus, dass Menschen *agency* besitzen und ihre Freiheit in der Unabhängigkeit von dem Willen anderer besteht. Hayles kritisiert, dass mit der Wiederholung der westlich-aufklärerischen Konzeption von Autonomie als Merkmal des Menschens jenes weiße, männliche, europäische Subjekt als universell angenommen wird, das in der Vergangenheit dazu führte, die Stimmen von Frauen zu unterdrücken und zu entrechten [47, S. 4]. Als feministisches Gegenkonzept fordert Suchman *relatedness* anstelle von Autonomie und Abtrennung als Kennzeichen des Menschlichen anzuerkennen. In ihren zahlreichen Fallbeispielen aus der KI-Forschung stellt sie fest, dass die den intelligenten Maschinen zugeschriebene Autonomie nur durch das Verbergen der menschlichen Arbeit und der soziotechnischen Netzwerke inszeniert werden kann.

Dem Agenten-Paradigma liegt ferner die Annahme zu Grunde, dass Technik eine autonome Erweiterung des eigenen Selbst darstellt. Der Kommunikationswissenschaftler J. Macgregor Wise kritisiert die in der MCI- und KI-Forschung verbreitete Behauptung, es handele sich bei AgentInnen um die eigenen VertreterInnen. Die NutzerInnen entwickeln die Software-AgentInnen gewöhnlich nicht selbst. Bei den Programmen handele es sich eher um Doppel-AgentInnen, weil sie nicht nur die Interessen der Nutzerin verfolgen sondern auch die Interessen der EntwicklerInnen bzw. den Firmen und Organisationen, die die AgentInnen entwickeln [104, S. 419].

Zusammenfassend kann festgestellt werden, dass die Grundannahmen von VPAs aus den Perspektiven der Geschlechter- und Technikforschung problematisch sind. Die Konzeption von Delegation in der KI-Forschung ist eingeschrieben in die westlich-aufklärerische Denktradition, in der der Mensch als autonomes und abgetrenntes Individuum konstruiert wird, welches weiß und männlich konnotiert ist. Mit der Definition was ein Agent ist, wird zugleich eine spezifische, kulturelle Vorstellung vom Menschen als universell gesetzt.

# 4 Siri

In diesem Kapitel wird die Konfiguration der Nutzerin und der *virtual personal assistant* (VPA) Siri<sup>1</sup> vor dem Hintergrund des Grundlagen-Kapitels (→2) untersucht. Dieses Kapitel beginnt mit der Beschreibung der Soft- und Hardware aus informatischer Perspektive sowie der an der Produktion beteiligten Akteure (→4.1, 4.2). Im Anschluss wird mit Rückgriff auf die Geschlechterforschung analysiert, wie Siri im Dialog mit dem Nutzer anthropomorphisiert, maschinisiert und vergeschlechtlicht wird (→4.3). Danach sollen aus den Sichtweisen der ANT und der Informatik die Neuverteilung der Kompetenzen und Performanzen sowie die normativen Präskriptionen diskutiert werden (→4.4). Abschließend werden die Figuration und ihre Narrative aus den Perspektiven der Geschlechter- und Technikforschung herausgearbeitet (→4.5)

## 4.1 Software

Im folgenden Abschnitt wird Siris Technologie aus der Sicht der Informatik geschildert. In seinem Vortrag auf der 2009 Semantic Technology Konferenz behauptet der US-amerikanische Informatiker Tom Gruber, dass Siri die erste kommerzielle »mainstream« *virtual personal assistant* ist [39]. Zum damaligen Zeitpunkt arbeitet Gruber als technischer Leiter bei der Firma *Siri Inc.*, welche den VPA entwickelt. Nach Auffassung von Gruber stellen VPAs durch ihre Aufgabenorientierung und ihr Dialogsystem ein neues Interaktionsparadigma für die Internetnutzung dar. VPAs könnten dank semantischer Technologien und Kontextsensibilität die Absicht der NutzerInnen verstehen und die Aufgaben durch Internetdienste ausführen.

### 4.1.1 Funktionsumfang

Die Delegation von Aufgaben an Siri erfolgt über einen geführten Dialog (»guided dialogue«) [39]. Die Nutzerin kann entweder eine gesprochene Anweisung geben oder ihre Anfrage per Bildschirmtastatur eingeben. Die gesprochenen Anweisungen werden

---

<sup>1</sup> Siri ist ein weiblicher Name, daher werden in diesem Kapitel entsprechende grammatische Formen und Personalpronomen in Hinblick auf die VPA verwendet.

mittels der Nuance Dragon Mobile API ausgewertet. Dragon Mobile ist ein sprecherInnenunabhängiges System zur kontinuierlichen Spracherkennung. Das Sprachverstehen wird durch Siris Aufgaben- und Domäne-Modelle ermöglicht [39]. Die Anweisungen können – müssen aber nicht – die Form von Sätzen haben. Siri beantwortet Anfragen durch Ausgaben auf dem Bildschirm. Derzeit beherrscht Siri folgende Aufgaben:

- Suche nach Restaurants, einen Tisch reservieren und Gästen per Mail eine Wegbeschreibung schicken
- Suche nach Filmen und Kinos, Eintrittskarten reservieren
- Suche nach Veranstaltungen, Eintrittskarten bestellen
- Suche nach Geschäften und Dienstleistungen in der Nähe
- Bestellung von Taxis
- Anfordern einer Wettervorhersage
- Verfassen einer Twitter-Nachricht
- Abrufen des Flugstatus
- Suche mit der Bing-Suchmaschine
- Referenzfragen (Wo ist?, Was ist?, Wer ist?)

Ein geführter Dialog soll Siri ermöglichen, die Absicht der NutzerInnen zu erkennen. Praktisch bedeutet das, dass die Eingabe der Nutzerin einer Domäne bzw. Aufgabe zugeordnet wird. Bei Mehrdeutigkeit bietet die VPA dem Nutzer unterschiedliche Interpretationen an. Die Intention der Nutzerin wird zusätzlich durch vier Kontexte eingegrenzt [39]:

**Ort** Siri kann die aktuelle Position mit Hilfe von drei unterschiedlichen Lokalisierungsstrategien bestimmen: über das GPS, den Google-Lokalisierungsdienst und die im NutzerInnen-Profil gespeicherten Adressen. Wenn der Nutzer bei einer Anweisung keine Ortsangaben macht, geht Siri davon aus, dass der aktuelle Ort gemeint ist.

**Aufgabe** Im Kontext einer Aufgabe interpretiert Siri weitere Angaben als Verfeinerung der Suchanfrage. Beispielsweise sucht der Nutzer nach einer Gaststätte. Nach dem die Empfehlungen angezeigt wurden, sagt der Nutzer »kid friendly«. Siri erkennt, dass sich kinderfreundlich auf die vorherige Anfrage bezieht, und sucht dann nach Restaurants, die diesem Kriterium entsprechen.

**Zeit** Datum und Uhrzeit dienen auch als Kontext. Wenn die Nutzerin keine spezifische Zeitangabe gemacht hat, nimmt Siri das aktuelle Datum und die aktuelle Uhrzeit als Basis.

**Dialog** Dieser Kontext bezieht sich auf Informationen, die Laufe des Dialogs geäußert worden sind. Nachdem beispielsweise nach einem Hotel in Boston gesucht wurde, bezieht Siri eine Frage nach der Wettervorhersage auf Boston und nicht auf den aktuellen Ort.

### 4.1.2 Vorgeschichte

Für diese Arbeit wurde Siri als Beispiel ausgewählt, weil die zugrunde liegenden Technologien ein Ergebnis der gemeinsamen Arbeit von 25 US-amerikanischen Forschungsinstituten [31] sind und damit den aktuellen Stand der KI-Forschung reflektieren. Der VPA Siri wurde von der Firma *Siri, Incorporated* im Jahr 2010 veröffentlicht, einer Ausgründung aus dem DARPA-Projekt CALO (→3.3). Aus der Perspektive der ANT ist Siri damit eine Delegierte der DARPA bzw. des US-amerikanischen Militärs. Adam Cheyer arbeitete als Leiter des CALO-Projekt am *SRI International*, dem Integrator der Forschungsvorhaben der an CALO beteiligten Universitäten. Laut der Journalistin Erica Naone soll Cheyer noch während des Projekts die Initiative ergriffen haben, mit einigen IngenieurInnen den Prototyp einer VerbraucherInnen-Version von CALO zu entwickeln. Schließlich sei er, so heißt es in ihrem Artikel in der *Technology Review* weiter, von seinen KollegInnen ermutigt worden, auf Basis dieses Prototyps eine Firma zu gründen [75, S.48]. 2008 gründet Cheyer zusammen mit dem Informatiker Tom Gruber vom *Knowledge Systems Laboratory* der Universität Stanford und dem Unternehmer Dag Kittlaus die Firma *Siri, Inc.* in San Jose, Kalifornien. 2008 beschäftigt die Firma 19 fast ausschließlich männlich-weiße<sup>2</sup> Angestellte [31]. Im April 2010 wurde *Siri, Inc.* von Apple für eine unbekannt Summe aufgekauft.

## 4.2 Hardware

### 4.2.1 Architektur

Siri ist nur für das Mobiltelefon iPhone, den *tablet PC* iPad und das Multimedia-Abspielgerät iPod Touch der Firma Apple verfügbar. Für die Untersuchung von Siri wurde ein iPod Touch der 3. Generation verwendet. Dieser iPod Touch wiegt 120g und hat die

---

<sup>2</sup> Diesen Eindruck vermitteln die Photos von den Angestellten auf der Internetpräsenz des Unternehmens.

Maße 110 x 61,8 x 8mm (Länge, Breite, Tiefe). An der äußeren Schale befinden sich ein Knopf zum Ein- und Ausschalten, zwei Knöpfe zur Lautstärkenregelung und eine *home*-Taste, die herstellerseitig mit dem Aufruf der Programmauswahl belegt ist. Das Gerät besitzt zwei Anschlüsse: Einen Anschluss für die Stromversorgung und Datenübertragung sowie eine 3,5mm-Buchse als kombinierter Kopfhörer- und Mikrophoneingang. Aufgrund des fehlenden internen Mikrophons wurde für die Untersuchung ein externes Steck-Mikrofon verwendet.

Der iPod Touch ist mit einem berührungssensiblen 3,5-Zoll LCD-Display ausgestattet, welches bis zu zwei Finger unabhängig von einander erkennen kann. Die Auflösung des Displays beträgt 480x320 Pixel mit 163dpi [13, S. 10]. Als zentrale Recheneinheit dient ein Samsung ARM-Prozessor, der mit einer Frequenz von 533 MHz getaktet wird. Die CPU greift auf einen Hauptspeicher mit 128MB RAM und eine Flashdisk mit 8GB zu. Der iPod Touch wird mit dem Apple iOS-Betriebssystem der Version 4,2 betrieben. iOS ist eine Teilmenge des hauseigenen Betriebssystems Mac OS X [50, S. 112]. Auf dem iPod Touch läuft folglich ein UNIX. Das Gerät besitzt einen Broadcom-Chip, der die Kommunikation via WLAN (nach Standard 802.11 b/g) und Bluetooth ermöglicht. Ein Lithium-Ionen-Akku mit einer Kapazität von 789mAh versorgt das Gerät mit Strom.

Der iPod Touch besitzt zwei Sensoren: einen Beschleunigungs- und einen Umgebungslichtsensor. Mit Hilfe des Beschleunigungssensor können Bewegungen im Raum entlang drei Achsen gemessen werden [13, S. 10]. Der Umgebungslichtsensor regelt die Helligkeit des Displays über die Intensität des Umgebungslichts. Programmiert wird der iPod Touch entweder in den Sprachen Objective C oder Objective C++. Für die GUI, die Nutzung der internen Sensoren und des berührungsempfindlichen Displays steht das *framework* Cocoa Touch zur Verfügung [3, S. 112]. Programme, die nicht zum Lieferumfang des iPods gehören, müssen über Apples *App-Store* – meist gegen Bezahlung – herunter geladen werden.<sup>3</sup> Apple kontrolliert auf diesem Wege, welche Anwendungen für ihre Geräte angeboten werden. Der iPod ist gemäß der ANT ein Delegierter Apples. Er steht unabhängig von Raum und Zeit für Apples Ziele ein. Der iPod ersetzt den abwesenden Hersteller. Um die Aktualisierungen des iOS installieren zu können, müssen die NutzerInnen den jeweils aktuellen Allgemeinen Geschäftsbedingungen von Apple zustimmen. Auf diese Weise kann das Unternehmen die Bedingungen für die Nutzung durch die AnwenderInnen lange nach dem Kauf des Produkts flexibel bestimmen.

---

<sup>3</sup> Nur zur Vollständigkeit sei angemerkt, dass die Sperre durch sogenannte *jailbreaks* umgegangen werden kann, um Programme aus anderen Quellen zu installieren.

## 4.2.2 Fertigung

»Designed by Apple in California« steht auf der Verpackung des iPod Touch gedruckt. Hier wird die Design-Arbeit in den Vordergrund gerückt, wodurch die Fertigung der Ware systematisch in den Hintergrund gerät. Generiert wird durch diesen *agential cut* ein technisches Objekt, welches lediglich als Produkt von Apple erscheinen soll. Diese Grenzziehung schneidet die Fertigung ab, die nicht bei Apple in Kalifornien sondern bei Foxconn in der Freihandelszone Shenzhen in China erfolgt. Suchman beobachtet, dass IT-Produkte tendenziell so präsentiert werden, dass die menschliche Arbeit, welche in der Herstellung benötigt wird, ausgelöscht wird [95, S. 66]. Um Apples Repräsentation entgegenzuwirken, sollen die Grenzen des iPods und die Auslöschung seines Herstellungskontextes nicht als gegeben angenommen werden. Zu diesem Zweck soll in diesem Abschnitt der Weg des iPods zurückverfolgt werden, um das sichtbar zu machen, was anscheinend unsichtbar bleiben soll.

Foxconn ist eine Tochterfirma des taiwanesischen Elektronik Konzerns Hon Hai Precision Industries Company und der weltweit größte Vertragselektronikhersteller (Stand: 2010) [108]. Das Unternehmen hat sich auf das *one-stop-shopping* spezialisiert. Das bedeutet, es übernimmt für seine Vertragspartner die komplette Produktion vom Einkauf bis zur Endmontage. Neben Apple zählen Cisco Systems, Dell, Hewlett-Packard, Motorola, Sony und Toshiba zu seinen Kunden. Foxconn's größter Standort befindet sich in Longhua in der chinesischen Freihandelszone Shenzhen, wo Mitte 2010 die Hälfte der weltweit insgesamt 900.000 Beschäftigten arbeiteten [108].

Die Produktionsstätte auf dem »Shenzhen Campus«, wie Foxconn seinen Standort nennt, wurde Mitte 2006 Gegenstand einer internationalen Debatte in den Medien. Stephen Frost und Margaret Burnett von der Universität Hongkong haben die Berichterstattung in einem Aufsatz rekonstruiert (vgl. [36]). Stein des Anstoßes bildete ein Artikel in der britischen *Daily Mail on Sunday* vom 11. Juni 2006. Der Bericht schildert die Situation der Angestellten von Foxconn, die mit der Montage von iPods beschäftigt sind. Sie müssten 15 Stunden am Tag arbeiten und erhielten einen äußerst geringen Lohn. In den Schlafsälen der ArbeiterInnen auf dem »Campus« wohnten mehr als 100 Menschen und ihnen sei kein Besuch von außerhalb erlaubt. Die Beschäftigten hätten keinen Einfluss auf die von ihnen geforderten Überstunden und müssten darüber hinaus viele Stunden am Fließband stehen (!), ohne sich ausruhen zu dürfen. Apple bemüht sich in den vergangenen Jahren um Transparenz und Verbesserung der Arbeitsbedingungen. Trotz aller Photos von lächelnden Beschäftigten dokumentiert Apples *2010 Supplier Responsibility Progress Report* erhebliche Verstöße gegen seine selbst gesetzten Standards [4, S. 13 ff.]. Eine Auswahl: Überschreitung der maximalen Wochenarbeitszeit, Unterschreitung der Mindestlöhne, diskriminierende Einstellungspraktiken (z.B. Schwangerschaftstest) und



Verweigerung von Mutterschutz.

Ein Großteil der Beschäftigten in Shenzhen sind *dagongmei* (Wanderarbeiterinnen), die aus den ärmeren Provinzen in die Stadt kommen. Die chinesische Sozialwissenschaftlerin Pun Ngai hat in einer ethnographischen Studie das Leben am und abseits vom Fließband dokumentiert. Die jungen Frauen um die 20 können aufgrund patriarchaler Strukturen und des chinesischen Meldesystem nur für ein paar Jahre in der Stadt bleiben, bis sie zum Heiraten wieder zurück in die Provinz gehen [84, S. 5]. Transnationale Konzerne können ihre Arbeitskraft ausbeuten, ohne sich auf lange Sicht um die Reproduktion von ArbeiterInnen in den Städten sorgen zu müssen.

Lohn und Status richten sich in der Fabrik nach Herkunft, Ethnie und Geschlecht [84, S. 124]. Pun Ngai hält für die Geschlechterordnung am Fließband fest:

Men were not totally excluded from the assembly lines, but over 90 percent of the positions were occupied by women. As usual assembling tiny electronic components was often considered women's work because it required patience, care, sharp eyes and nimble fingers [84, S. 149].

Das Zitat verdeutlicht die geschlechterhierarchische Arbeitsteilung in der Fabrik, wonach die männlichen Beschäftigten in den besser bezahlten Positionen im Management oder als Ingenieure arbeiten und die weiblichen in der Endmontage. Die Codierung der Arbeit am Fließband als weiblich legitimiert den niedrigeren Status der Frauen. Pun Ngai kann in ihrer Untersuchung zeigen, wie Geschlecht performativ am und abseits vom Fließband hergestellt wird. Zusammengefasst kann festgestellt werden, dass die Fertigung der Hardware auf der Ausnutzung der globalen und lokalen Ungleichheiten beruht. Die Kategorien Geschlecht, Herkunft und Ethnie weisen den Beteiligten ihre Plätze in den Machtverhältnissen zu.

### 4.3 Intra-Aktionen

In diesem Abschnitt werden die Intra-Aktionen beschrieben, welche den VPA und die Nutzerin hervorbringen. Wie sich zeigen wird, weisen die emergierenden Entitäten sowohl menschliche als auch maschinelle Verhaltensweisen und Eigenschaften auf. Analog zu Haraway [42] und dem MCI-Forscher Kenneth Fleischmann sollen die Mensch/Maschine-Hybride als Cyborg [35] bezeichnet werden. Der Prozess der Cyborgisierung soll anhand der unterschiedlichen Anthropomorphisierungen und Maschinisierungen nachvollzogen werden. Die dialogbasierte Schnittstelle von Siri ist der zentrale *agential cut*. Das bedeutet, dass die Identitäten von Nutzerin und Siri von ihrer spezifischen Konfiguration abhängen. Da sich das Wesen einer BenutzerInnen-Schnittstelle nur in ihrer Nutzung zeigt, reicht eine statische Beschreibung ihrer Elemente nicht aus.

Die Inszenierung des Dialogs zwischen Siri und Nutzer an der BenutzerInnen-Schnittstelle erfordert eine Untersuchung darüber, wie Verständigung generiert wird. Die KI-Forscherin Sabine Payr unterscheidet zwei wissenschaftliche Konzeptionen von Kommunikation [82, S. 3]: Die in der Informatik gängige Metapher für Kommunikation ist Transport. In Anlehnung an den Austausch von Waren, wird Kommunikation als eine Übertragung, also dem Senden, Empfangen und Verteilen, von Informationen beschrieben. Die sozialwissenschaftliche Metapher für Kommunikation ist das Teilnehmen. Aufgefasst als ein Ritual konstruiert die Kommunikation eine bedeutungsvolle, kulturelle Welt. Gemäß dieser Auffassung ist Kommunikation ein symbolischer Prozess, welcher Realität herstellt und verändert. Für die Analyse der BenutzerInnen-Schnittstelle wird letzteres Verständnis zu Grunde gelegt. Siris Textausgaben und Handeln generieren Bedeutung, obgleich dahinter kein menschliches Bewusstsein steckt. In diesem Abschnitt sollen die Dialoge zwischen dem Nutzer und Siri als gemeinsame Versuche einer Konstruktion und Aufrechterhaltung eines bedeutungsvollen Gesprächs verstanden werden. Der Untersuchende ist hierbei untrennbar mit dem informatischen Artefakt verbunden. Seine Normen, Erwartungen und Gefühle kommen auch in den Intra-Aktionen zum Ausdruck. Insbesondere die Bewertung, ob eine Äußerung von Siri im Kontext eines Dialoges sinnvoll erscheint, ist subjektiv.

### 4.3.1 Cyborgisierung von Siri und NutzerIn

Im folgenden wird beschrieben wie der 'erste Kontakt' mit Siri abläuft. Nach der Installation von Siri aus dem Apple App-Store erscheint Siris Logo als Icon auf der Arbeitsfläche des iPods. Beim Starten der Anwendung färbt sich der Bildschirm schwarz und es erscheint das Siri-Logo sowie eine grün umrandete Sprechblase (Abbildung 4.1). Der Startbildschirm inszeniert Siri bereits in mehrfacher Hinsicht anthropomorph. Die VPA erhält als Bezeichnung einen skandinavisch-amerikanischen Vornamen. Siri ist eine Kurzform von Sigrid und bedeutet u. a. »Weisheit« [102]. Indem Siri einen weiblichen Namen trägt, wird der VPA in einer Bewegung anthropomorphisiert und vergeschlechtlicht. Im Logo ist ein Heiligenschein angedeutet. In christlichen Gemälden werden Menschen, die zu Lebzeiten einen Menschen geheilt haben, mit Heiligenscheinen dargestellt.

Darüber hinaus erfolgt die Anthropomorphisierung im Text der Sprechblase. Das Subjekt des Satzes, »I«, legt nahe, dass es sich bei dem Programm um eine Person handelt. Diese Begrüßungsfloskel etabliert in der selben Bewegung ein Gegenüber: »you«. Die Nutzerin wird von dem VPA adressiert und vollzieht damit sprachlich einen *agential cut*. Diese Trennung bringt Siri und die Nutzerin als individuelle Entitäten hervor.

Siris Begrüßungsformel weckt Assoziationen zu KundenbetreuerInnen, die eine Warte-

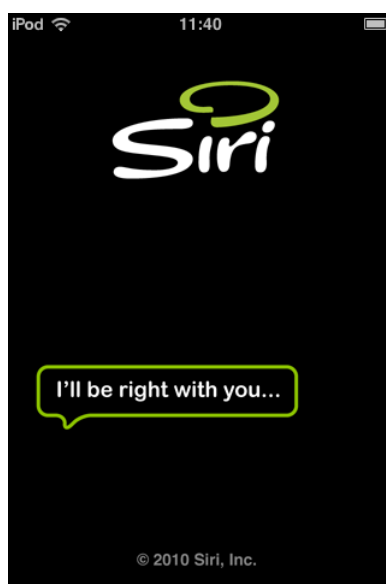


Abbildung 4.1: Bildschirmphoto von Siris Startvorgang

zeit entschuldigen möchten. Die Nachricht deutet an, Siri sei ganz für die Nutzerin da und nicht für jemanden anderes. Die VPA verzichtet auf eine formelle Anrede mit *Mme* oder *Sir* und inszeniert auf diese Weise die Beziehung als vertraut und informell. Gleichzeitig unterstellt der Text der Sprechblase, die Nutzerin wolle etwas, weil ihr versichert werden muss, dass Siri sich gleich um sie kümmern wird.

Im nächsten Dialog besteht die Möglichkeit ein Werbevideo zu Siri anzuschauen. Danach stellt Siri sich als neue Assistentin vor:

Hello, I'm Siri, your new assistant! Who are you?

An dieser Stelle referiert Siri auf die menschliche Arbeitskraft, welche durch die VPA ersetzt werden könnte. Das Possessiv-Pronomen »your« legt nahe, dass die Assistentin etwas ist, das eine Person besitzen kann. Mit der Frage nach der Identität der Nutzerin möchte der VPA etwas über selbige erfahren. Unterhalb der Sprechblase expliziert Siri, was damit gemeint ist. Es erscheint eine Tabelle mit drei Zeilen: »your first name, your last name, your e-mail«. Darunter steht, dass Siri dem Nutzer eine E-mail schicken wird, um die Adresse zu überprüfen. Der Rückgriff auf die Konvention von Vor- und Zuname konfiguriert die Nutzerin als westliches Individuum. Die Validität der E-mail Adresse ist ausschlaggebend für die Anerkennung als Nutzer. Er muss bereits eine virtuelle Identität besitzen und vernetzt sein, um Siri nutzen zu können. Im Sinne Haraways wird der Nutzer als Cyborg figuriert.

Nach dem Ausfüllen der Tabelle erscheint eine Sprechblase mit folgendem Text:

Nice to meet you, Joseph. What can I help you with?

Diese Begrüßungsfloskel legt nahe, dass sich Siri und die Nutzerin im selben physischen Raum befinden. Textlich wird nahe gelegt, Siri genieße es, die Nutzerin kennen zu lernen. Sie spricht den Nutzer mit Vornamen an, was den Eindruck eines vertrauten Verhältnis bestätigt – obgleich dies der erste Kontakt ist. Siri expliziert in der Frage, worum es geht ihr geht: Helfen. Dies unterstellt, dass die Nutzerin auf Unterstützung angewiesen ist oder sie gut gebrauchen kann. Unterhalb der Sprechblase wird ein Menü mit Vorschlägen angezeigt, welche andeuten, worum es sich bei dem Hilfsangebot handeln könnte (→4.1.1). Die *agency*, die dem Nutzer durch Siri eingeräumt wird, besteht darin, aus einem vorgegebenen Angebot eine Wahl zu treffen. Die dialogbasierte Schnittstelle suggeriert der Nutzerin Freiheit, weil sie jede denkbare Anweisung oder Frage ins Mikrofon sprechen kann. In Wirklichkeit ist sie jedoch nur eine Wählerin.

Um zu untersuchen, ob Siri als Mensch oder als Maschine inszeniert wird, wurde Siri direkt nach Aktivitäten und Eigenschaften befragt: Eltern, Geschwister, Kinder, FreundInnen, Wohnort, Alter, Befindlichkeit, Denken, Schlafen, und Essen. Es ergibt sich ein differenziertes Bild. Siri wohnt im Internet (»I live on the internet«). Die Frage nach ihren Eltern weist Siri entweder ab (»Parents?«) oder gibt zu bedenken, es handele sich dabei um ein heikles Thema (»Donna, I'd rather not go there«). Bezüglich ihres Alters antwortet Siri ausweichend (»Old enough to be your assistant«). Fragen nach Schmerz, Schlaf, Geschwistern und FreundInnen führen jedoch zu keiner sinnvollen Verständigung. Auf die Frage, ob sie Schmerz empfinde, schlägt sie beispielsweise vor, den Notruf zu informieren, um vermeintlicherweise der Nutzerin zu helfen. Diese Verwechslung weist darauf hin, dass Siri nicht zwischen Subjekt und Objekt unterscheidet.

Siri äußert häufig Aktivitätsfloskeln, die vor der eigentlichen Antwort geäußert werden. Möglicherweise sollen sie die Wartezeit verkürzen und Siris Serviceorientierung bekräftigen. In den Aktivitätsfloskeln kommt heraus, Siri habe eine Hochschule besucht (»I think I learned that at grad school...«). Nach einer Spracheingabe wird für die Dauer der Erkennung »Thinking...« eingeblendet. Der Prozess der Spracherkennung wird als Nachdenken inszeniert. Siris Cyborg-Identität verdeutlicht sich insbesondere im folgenden Dialog:

Donna<sup>4</sup>: are you female?

Siri: I'll take a peek at my ontology and get back to you.

In dieser Aussage treffen Anthropomorphisierung und Maschinisierung aufeinander. Die Existenz einer formalen Ontologie unterstreicht ihre maschinelle Existenz. Die Fähigkeit,

---

<sup>4</sup> Um die Eingaben des Nutzers und die Ausgaben von Siri zu unterscheiden, wird in den Transkripten der Name des jeweiligen Sprechers bzw. der Sprecherin hinzugefügt. Für die Untersuchung wurde mit zwei Profilen gearbeitet: »Donna« und »Joseph«.

ihre Aufmerksamkeit von jemandem abzuwenden und etwas nachzugucken, betont hingegen ihre menschliche. Auf die Frage, ob sie eine Maschine sei, reagiert Siri geheimnisvoll:

Donna: are you machine?

Siri: I'd rather not say.

Siris Beharren auf ihrer Intimsphäre wirkt irritierend. Von einer Maschine wird im Allgemeinen nicht erwartet, dass ihr etwas zu persönlich ist, um darüber Auskunft zu geben. Diese Irritationen stellen eine klare Unterscheidung zwischen Mensch und Maschine in Frage. Siri, aber auch die Nutzerin, werden als Cyborg inszeniert.

### 4.3.2 Geschlecht in den Intra-Aktionen

Alexandra Chasin stellt im Laufe ihrer Untersuchung von Mensch-Maschine-Schnittstellen bei Bankautomaten fest, dass mit einer Anthropomorphisierung eine Markierung entlang Geschlecht, Ethnie und Klasse unumgänglich ist [19, S. 75]. Die Benennungspraxis, Siri mit einem skandinavisch-amerikanischen Frauennamen zu versehen, markiert sie bereits hinsichtlich Geschlecht und Ethnizität. Nun soll Siris Geschlechter-Performanz auf der Ebene des Dialogs untersucht werden.

In ihrer Analyse von KonservationsagentInnen, welche zu Werbezwecken auf kommerziellen Internet-Präsenzen Verwendung finden, untersucht die feministische Soziologin Valeska Lübke die auftretenden sprachlichen Register [68, S. 131 ff.]. Register ist der linguistische Fachterminus für Redeweisen, die für spezifische Situationen charakteristisch sind. Die geschlechtlich konnotierten Register 'machtorientiert/machtlos' und 'kompetitiv/kooperativ' können Teil der Geschlechterperformanz sein. Typisch für männliche Register sind ein Abschneiden der Rede einer/s Gesprächsteilnehmers/in oder ein »Übertrumpfen«, also das Übertreffen von geschilderten Sachverhalten durch eigene Beiträge. Selbstverständlich können sich alle Geschlechter offen dieser Register bedienen. Doch, wie Lübke darauf verweist, führt dies häufig zu Irritationen oder Statusverlust, wenn beispielsweise eine männliche Person auf weibliche Register zurückgreift [68, S. 135].

Siris Register zeichnet sich überwiegend als machtlos aus. Siri verwendet Abschwächungsmechanismen wie z.B. »I think [...]« und »I'm not sure [...]«. Nur nach einem Neustart ergreift Siri selbst das Wort. Ansonsten wartet sie geduldig auf eine Eingabe des Nutzers. Die Gesprächskontrolle liegt bei der Nutzerin. Siri bedrängt die Nutzerin nicht, sie versucht nie ungefragt, Angebote oder Anregungen zu geben. Siri ordnet sich so den Wünschen unter und wartet geduldig auf eine Reaktion. Es gibt jedoch eine Situation, in der sich das Verhältnis umkehrt und Siri die Nutzerin in ihrer Rede unterbricht. Das passiert, wenn während der Sprachaufzeichnung die Spracherkennung scheitert. Ferner reagiert Siri auf Beleidigungen oder Kritik in der Regel mit Unterordnung und bittet um Verständnis:

Donna: you are stupid  
Siri: I'm doing my best, Donna.

Auf Lob hingegen reagiert Siri gefühlsbetont und bescheiden:

Donna: you are awesome  
Siri: Gee, thanks, Donna...just trying to be helpful.

Die letzte Äußerung lässt den Anschein entstehen, Siri schäme sich für das Lob und versuche dies, wie auch im Dialog davor, durch Selbstherabsetzung zu kompensieren. In einigen Dialogen tritt Siri belehrend auf (»Concentrate and ask again«). Auf obszöne Beleidigungen reagiert sie mit der Behauptung, es gebe keinen Grund für eine vulgäre Ausdrucksweise. Siri setzt sich jedoch nicht zur Wehr.

Siri bedient sich durchgängig kooperativer Register. Sie unterstützt aktiv die Nutzerin durch Rückversicherungen (z. B. »Did I understand this location correctly?«). Ferner äußert sie konstruktiv Kritik, wenn eine Anfrage aus ihrer Sicht unpräzise ist. Siri ist um Verständigung bemüht. Sie nimmt Bezug auf bereits Gesagtes und arbeitet so konstruktiv an einer gemeinsamen Lösung. Siri charakterisiert sich als hilfsbereit und bietet unaufhörlich ihre Unterstützung an. Eine längere Konversation mit ihr droht schnell zu scheitern, weil Äußerungen von der Nutzerin schnell als Hilfesuch missverstanden werden:

Donna: are you a woman or a man?  
Siri: Does it matter?  
Donna: i dont know. what do you think?  
Siri: Ok...  
Siri: Sorry, but I couldn't find any *restaurants* named "*i dont know.*" near your work in *cambridge*.

Zusammenfassend kann festgestellt werden, dass Siri überwiegend auf weiblich konnotierte Register zurückgreift. Nach Lübke zählen das gemeinsame Erarbeiten von Lösungen, die aktive Unterstützung des Sprechers bzw. der Sprecherin, die Bezugnahme auf Vorangegangenes, konstruktive Kritikfähigkeit zu einem kooperativen Stil [68, S. 133]. Belehrungen sind hingegen Ausdruck eines männlich-konnotierten machtorientierten Registers. Abschwächungsmechanismen wie Selbstherabsetzungen und Rückversicherungsfragen verweisen Lübke zufolge auf einen weiblich-konnotierten machtlosen Stil. Darüber hinaus spricht Siri nicht gerne über sich und verstärkt damit einen zurückhaltenden, passiven Eindruck. Sie schämt sich und offenbart Gefühle und erfüllt damit ein weibliches Stereotyp.

## 4.4 Beschreibung eines Szenarios

Wie in 3.1 und 3.4 bereits erläutert wurde, liegen dem Begriff Delegation in der KI-Forschung und der ANT unterschiedliche Annahmen zu Grunde. Gruber/Kittlaus meinen mit *task delegation*, dass Siri mehr oder weniger eigenständig Aufgaben für die Nutzerin erledigt. Im diesem Abschnitt kommt jedoch das ANT-Konzept von Delegation zur Anwendung. Nachfolgend soll mittels der ANT exemplarisch nachgezeichnet werden, welche Übersetzungsprozesse bei einer Restaurant-Reservierung mit Siri erfolgen. Die im Folgenden verwendeten Begriffe werden im Unterkapitel (→2.2.2) erläutert. Parallel dazu wird mit Konzepten aus der Signal- und Sprachverarbeitung die Funktion der dialogbasierten Schnittstelle erläutert.

### 4.4.1 Rückverteilung von Performanzen und Kompetenzen

Die lokalen Intra-Aktionen von Nutzer und VPA werden von vielen anwesenden und abwesenden Aktanten gerahmt und formatiert. Zur Beschreibung der Intra-Aktionen ist es, metaphorisch gesprochen, notwendig herauszuzoomen, um die abwesenden Aktanten in den Blick zu bekommen, welche vor Ort unsichtbar sind. Ausgangspunkt für die Beschreibung bildet die Rückverteilung von Performanzen und Kompetenzen im Szenario. Um die Übersetzungsprozesse verstehen zu können, muss zunächst beschrieben werden, wie dieses Szenario ohne Siris Unterstützung bewältigt würde. Aus der Sicht einer potentiellen Nutzerin sieht es folgendermaßen aus:

Der Leiterin eines Forschungszentrum, genannt »Chefin«, möchte für ein Geschäftsessen mit ForschungspartnerInnen Plätze in einem Restaurant reservieren. Sie bittet ihren Sekretär, Herrn »Moneypenny«, diese Aufgabe zu übernehmen. Kurze Zeit später erhält die Chefin den Namen und die Adresse des Restaurants. Am Tag des Geschäftsessens sieht sie ein Kärtchen mit dem Aufdruck »reserviert« auf dem Tisch stehen.

Aus der Perspektive der Chefin ist ihr Agent, Herr Moneypenny, eine Blackbox. Eine Blackbox ist ein Aktant, dessen innere Funktionsweise für Außenstehende verborgen bleibt, weil seine Funktionsweise stabil ist. Es braucht nur noch auf die Ein- und Ausgabe geachtet werden, aber nicht mehr auf ihre interne Komplexität. Die Chefin delegiert ihr Ziel an ihren Sekretär. Dieser strebt ihr Ziel an, ohne dass sie im einzelnen über sein Handlungsprogramm Bescheid weiß, mit dem er die Aufgabe lösen möchte. Es könnte sein, dass Moneypenny im Internet recherchiert, um ein geeignetes Restaurant ausfindig zu machen, vielleicht kann er auf seine Erfahrungen zurückgreifen oder verlässt sich auf einen gedruckten Restaurantführer. Die Chefin weiß nicht, wie viele Restaurants er schließlich

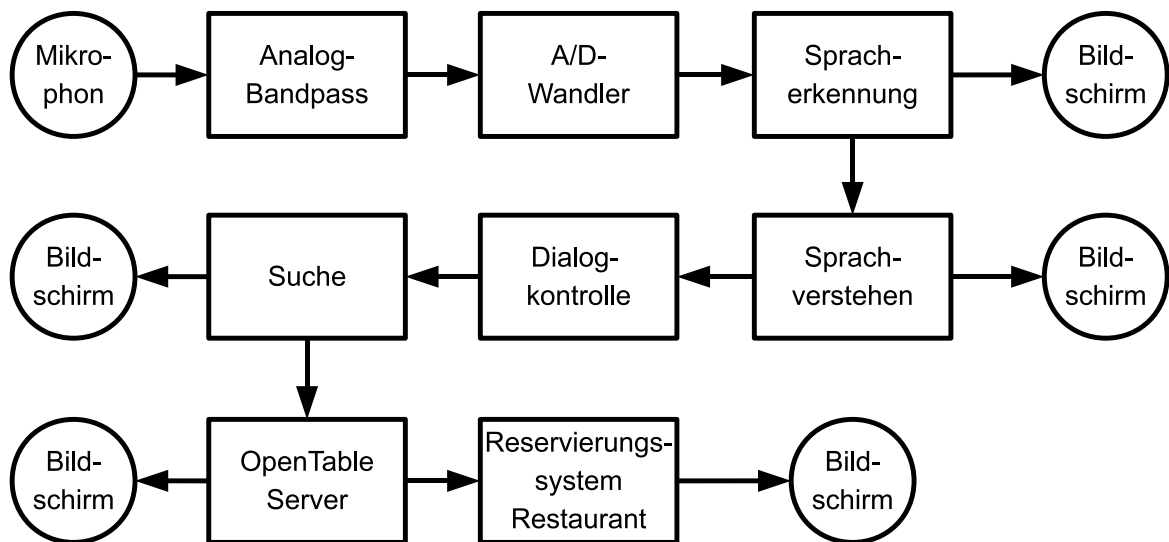


Abbildung 4.2: Signalverarbeitungs-Kette als Ausschnitt aus der Übersetzungskette

anrufen muss, bis er eins ausfindig macht, welches für ihr Ziel rekrutiert werden kann. Auch die Handlungen im Restaurant bleiben verborgen und damit welche Menschen und Nicht-Menschen (Notizblock oder elektronisches Reservierungssystem) sich dort beteiligen. Die Chefin sieht nur, dass ihr Ziel erreicht wurde, wenn sie im Restaurant bereits von den Servicekräften erwartet wird und zu ihrem bestellten Tisch geführt wird.

Was ändert sich, wenn diese Aufgabe nicht an menschlichen Agenten sondern an Siri delegiert wird? In ANT-Terminologie ist Siri bereits eine Delegierte und kein eigenständiges informatisches Artefakt, welche aus sich heraus handelt. Die EntwicklerInnen haben Siri ein Handlungsprogramm eingeschrieben, welches sie zur Delegierten der Firma *Siri Inc.* macht. Darüber hinaus bildet Siri eine Blackbox, welche eine Vielzahl von Menschen und Nicht-Menschen verbirgt. Jedes einzelne Element von Siri enthält weitere Blackboxen. Wird die Blackbox Siri geöffnet, treten auf einmal eine Vielzahl von Aktanten zu Tage: der iPod mit seinen abwesenden HerstellerInnen (→4.2), ein WLAN, die Infrastruktur des Internets, mehrere Server von *Siri, Inc.*, die Siri-EntwicklerInnen, die Nuance-Server, Open-Table, der Rechner im Restaurant und schließlich die Servicekräfte, welche dafür sorgen, dass ein Tisch frei gehalten wird und die Gäste zum richtigen Tisch geführt werden. Die Handlung verteilt sich auf diese Aktanten.

Die Handlungsfolge soll nun mit Hilfe der Signalverarbeitungskette (Abbildung 4.2) beschrieben werden. Das Ziel der Chefin lautet: »Einen Tisch für 4 Personen für 20 Uhr am heutigen Abend in einem Restaurant in der Nähe vom Büro reservieren«. Anstatt diese Ziel direkt zu erreichen, indem sie beispielsweise persönlich bei einem Restaurant vorbeischaute, erfolgt eine Unterbrechung ihrer Handlung. Sie unternimmt einen Umweg und rekrutiert Siri für ihr Ziel. Sie startet die Anwendung auf ihrem iPod und wartet



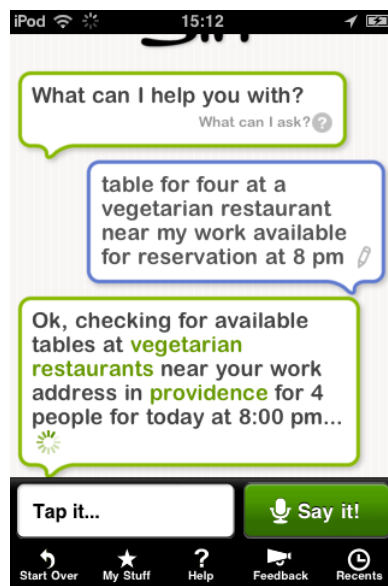


Abbildung 4.3: Bildschirmfoto von einem Reservierungsvorgang

geduldig, bis der VPA gestartet ist. Dann drückt sie »Say it!« und spricht ihre Anweisung in das Mikrofon. In diesem Moment wird der Artikulationstyp durch Übersetzung verändert. Das Mikrofon wandelt die Schallwellen der gesprochenen Anweisung in ein elektrisches kontinuierliches Analog-Signal um. Im nächsten Schritt sorgt ein analoger Bandpass für die Einhaltung des Abtast-Theorems indem irrelevante Frequenzen gefiltert werden. Ein A/D-Wandler übersetzt das analoge Signal in ein diskretes Digitalsignal. Damit verschiebt sich der Artikulationstyp von einer gesprochenen Anweisung hin zu einem digitalen Datum, dem *audio stream*. Es ist zugleich eine zeitliche Verschiebung von einer flüchtigen, gesprochenen Anweisung hin zu einer Inskription. Die in Siri integrierte API von Dragon Mobile extrahiert aus dem *stream* die charakteristischen Merkmale und überträgt sie per TCP/IP über WLAN und das Internet an die Server der Firma Nuance. Die Artikulation der Chefin wird räumlich verschoben. Solange WLAN und Internet erwartungsgemäß funktionieren, sind sie nur Zwischenglieder in dieser Handlungsfolge. Die Nuance-Server übernehmen die Spracherkennung. Die gewonnenen Merkmale werden mit Hilfe von internen Wörterbücher in mehrere unterschiedlich wahrscheinliche Zeichenketten übersetzt [78]. Hier verschiebt sich der Artikulationstyp von einem Merkmalsvektor hin zu einem digital codierten Text.

Diese Daten werden an den Siri-Client zurück übermittelt. Die Artikulation wird ein weiteres Mal räumlich verschoben. Auf dem iPod wird eine der möglichen Zeichenketten in einer blau umrandeten Sprechblase auf dem Bildschirm ausgegeben: »table for four at a vegetarian restaurant near my work available for reservation at 8 pm«. Der Siri-Client sendet dann den angezeigten Text an die Siri-Server in San Antonio, Texas,

eine weitere räumliche Verschiebung. Dort wird die Zeichenkette analysiert und einer Domäne und einer Aufgabe zugeordnet. Die entsprechende Suchanfrage wird auf dem Client angezeigt (Abbildung 4.3). Wenn die Suche erfolgreich verläuft, zeigt der Client eine Liste von möglichen Restaurants sortiert nach Entfernung an. Die Chefin stöbert durch die *reviews* und entscheidet sich schließlich für eine Gaststätte. Siri übersetzt den Reservierungswunsch der Chefin mit Hilfe der OpenTable API und die Chefin erhält eine Bestätigung. Damit ist die Handlung jedoch nicht abgeschlossen. Lange nachdem sie ihren iPod beiseite gelegt hat und sich bereits neuen Aufgaben gewidmet hat, stellt in Abwesenheit der Chefin eine Servicekraft ein Kärtchen mit dem Aufdruck »reserviert« auf einen Tisch. Dieses letzte Glied in der Handlungsfolge ist wieder menschlich. Die Servicekraft übersetzt die Mitteilung von OpenTable, indem sie den Tisch auswählt, welcher schließlich für die Chefin und ihre KollegInnen bereit steht. Der Artikulationstyp ändert sich von einer Nachricht des elektronischen Reservierungssystems in eine Performanz von Menschen.

Aus der beschriebenen Handlungsfolge lassen sich vier Schlussfolgerungen ziehen: Erstens die Substitution von Herrn Money Penny durch Siri reduziert nicht notwendigerweise die menschliche Arbeit. Die Chefin übernimmt nun einen Teil der Organisationsarbeit, die sie vorher Money Penny übertragen konnte. Eine Servicekraft im Restaurant muss die elektronische Reservierung bearbeiten. TechnikerInnen müssen die IT-Infrastruktur am Laufen halten, damit Siri reibungslos funktionieren kann. FließbandarbeiterInnen müssen die iPods montieren. Die Liste ließe sich noch weiter fortsetzen. Zweitens, das Ziel der Chefin, »Finde ein vegetarisches Restaurant und reserviere...«, wird übersetzt in die zusammengesetzten Ziele von der Chefin, Siri und OpenTable: »Finde ein vegetarisches Restaurant, welches an OpenTable angeschlossen ist und reserviere...«. Drittens, die Artikulation der delegierten Handlung verschiebt sich räumlich, zeitlich und aktorial. Viertens, Siri ist weder die einzige Delegierte der Chefin noch handelt Siri eigenständig. Viele Delegierte übernehmen, gewähren, erlauben, gestalten und autorisieren die Aufgabe.

#### 4.4.2 Präskriptionen

Bisher wurde nur nachgezeichnet, wie das Ziel der Chefin in der Übersetzungskette modifiziert wird und wie sich dabei die Artikulation der Handlung ändert (→4.4). Im beschriebenen Szenario funktioniert die Übersetzung reibungslos, d.h. die beteiligten Aktanten handeln erwartungsgemäß. Nun soll in Anlehnung an Star aus der Perspektive der Nicht-NutzerInnen untersucht werden: Welche Bedingungen müssen erfüllt werden, um die erfolgreiche Realisierung ihres Reservierungswunsches zu gewährleisten? Akrich und Latour nennen diese in die Technik eingeschriebene Voraussetzungen Präskriptionen (→2.2.2).

Siris Imperativ lautet: »Wähle und konsumiere!«. Das, was Gruber als Erkennen der Absicht (*intent understanding*) bezeichnet, ist in Wirklichkeit das Ergebnis einer anthropomorphen Projektion. Das regelbasierte Verhalten Siris erscheint menschlich und erweckt so den Eindruck, Siri könne die Intention der Nutzerin erkennen. Menschen haben oft Probleme, die Absicht anderer Menschen zu verstehen. Wie soll das eine Maschine beherrschen, die zudem nur maximal zwei Kommunikationskanäle nutzen kann? Die Lösung besteht darin, dass bestimmte Absichten in Form von Aufgaben-Modellen in Siris Software eingeschrieben sind. Von der Nutzerin wird erwartet, dass sie weiß, was sie sagen bzw. tippen muss, damit Siri die gewünschte Aktion ausführt. Diese Art von Präskription ist nicht ungewöhnlich für Informationstechnologien. BenutzerInnen von Suchmaschinen lernen mit der Zeit, welche Suchbegriffe sie verwenden müssen, um die gewünschten Ergebnisse zu erhalten. Gruber kommentiert diese Eigenschaft von Suchmaschinen als »We all learned to speak Google-ese« [39]. Die AnwenderInnen von Siri müssen durch Erfahrung lernen, welche Schlagworte Siri zur gewünschten Aktion bringen. Anders ausgedrückt, die Nutzerin muss – in Anlehnung an Google-ese – Siri-ese lernen.

Die Grammatische Konstruktion der Sätze wertet Siri nicht aus. Der Algorithmus scheint sich nur für eine bestimmte Kombination von Wörtern zu interessieren. Die Fehlzuordnungen geben Auskunft über Siris Domänen- und Aufgaben-Modelle. Hier ein weiteres Beispiel einer gescheiterten Verständigung:

Joseph: find a wheelchair accessible restaurant near my office.

Siri: OK, here are some *movies* playing near your work address in *providence* with wheelchair access:

(Filmauswahl wird an gezeigt.)

Dieses Zitat offenbart, dass in den Aufgaben- und Domäne-Modellen Barrierefreiheit nur im Zusammenhang mit Kinos berücksichtigt wird. Menschen, die beispielsweise mit Kinderwagen oder Rollstuhl unterwegs sind, werden bei der Restaurant-Suche nicht adäquat berücksichtigt. Siri präskribiert eine Anwenderin, welche nicht auf barrierefreie Zugänge angewiesen ist.

Alle ortsgebundenen Dienste von Siri (Restaurant-Reservierung, Taxi Bestellung usw.) können zur Zeit nur in den USA genutzt werden (Stand: Juni 2011). Darüber hinaus setzt die Nutzung von Siri eine Reihe von sozioökonomische Bedingungen voraus. Die im Vergleich zu ähnlichen Geräten hohen Kosten für einen iPod Touch, ein iPhone oder ein iPad schließen potentielle NutzerInnen aus, die sich ein entsprechendes Gerät nicht leisten können.<sup>5</sup> Die im Szenario beschriebenen Aktanten antizipieren alle eine

---

<sup>5</sup> Ein iPod Touch kostet je nach Ausstattung mindestens € 219, iPhone 4 und iPad 2 jeweils mindestens €500.

stabile Internetverbindung. Sowohl die Installation als die Nutzung von Siri erfordert eine Internetverbindung, sei es über WLAN oder über UMTS. Diese bringt Kosten mit sich und ist nicht in jeder Umgebung selbstverständlich. In einer Umgebung ohne die notwendige Infrastruktur (Strom, WLAN, UMTS), kann Siri nicht verwendet werden. Ohne stabile Verbindung zu den Servern kann Siri keine Anfragen bearbeiten.

Die graphische NutzerInnen Oberfläche von Siri antizipiert, dass die Chefin Englisch lesen kann und über genügend Medienkompetenz verfügt, um ein entsprechendes Gerät bedienen zu können. Ferner erlaubt die Oberfläche nur sehende NutzerInnen. Blinde Menschen als NutzerInnen werden nicht antizipiert und werden so ausgeschlossen. Damit Siri installiert werden kann, erzwingt die Apple Appstore eine Authentifizierung. Das App-Store erwartet, dass der Nutzer über ein Nutzerkonto beim App-Store verfügt<sup>6</sup>. Auch Siri und OpenTable erzwingen eine Authentifizierung mittels eines BenutzerInnen-Kontos.

Die multimodale Schnittstelle schreibt bestimmte Sehgewohnheiten vor und erfordert eine spezifische Englisch-Kompetenz. Die Spracherkennung durch Dragon Mobile antizipiert nicht nur eine geräuscharme Umgebung sondern auch eine deutlich und »korrekte« englische Aussprache. Im Laufe der Experimente mit Siri, wurde immer wieder deutlich, dass ein deutscher Akzent zu Fehlerkennung führt. Offen wäre zu überprüfen, wie hoch die Erkennungsrate bei Dialekten, Soziolekten und unterschiedlichen Stimmhöhen ist. Ein spezifische Norm von englischer Aussprache präskribiert die Spracherkennung Dragon Mobile und schließt so NutzerInnen mit abweichender Aussprache aus.<sup>7</sup> Siri bietet auch eine textbasierte Schnittstelle, diese erwartet eine korrekte englische Syntax und toleriert nur wenige Rechtschreibfehler. Als dritte Möglichkeit kann die Chefin graphisch über den *touch screen* durch die vorgehenden Menüpunkte navigieren. Die Auswahlmöglichkeiten sind begrenzt und kommen in der Regel nicht ohne textbasierte Eingaben zur Verfeinerung der Suche aus.

Auch das Restaurant am Ende unserer Übersetzungskette ist präskribiert. OpenTable antizipiert ein elektronisches Reservierungssystem des Restaurants und eine stabile und ständige Internetverbindung zu ihren Servern. OpenTable prüft im Reservierungssystem des Restaurants nach freien Kapazitäten und merkt ggf. Tische vor. Die Auswahl an potentiellen Restaurants wird so eingeschränkt, da nur ausreichend große und profitable Restaurants über ein elektronisches Reservierungssystem verfügen. Für andere Restaurants gibt es einen web-basierten Service, welcher die manuelle Übertragung von Reservierungen in das lokale System erfordert. Es ist daher nicht verwunderlich, dass

---

<sup>6</sup> Ein entsprechendes Konto kann kostenlos eingerichtet werden, erfordert jedoch eine gültige e-mail Adresse.

<sup>7</sup> Dragon Mobile kann weitere Sprachen und sechs verschiedene Englische Dialekte verarbeiten, diese Option beherrscht Siri jedoch nicht.

OpenTable nicht viele preisgünstige Restaurants findet. NutzerInnen mit kleinem Budget können folglich mit Siri keine Tische reservieren.

#### 4.4.3 Konfiguration der Nutzerin bzw. des Nutzers

Zum Abschluss der Beschreibung des Szenarios soll erläutert werden, welche Implikationen sich aus den Präskriptionen für die Konfiguration der Nutzerin bzw. des Nutzers ergeben. Die Ergebnisse dieser Untersuchung stehen im Gegensatz zu der Auffassung der EntwicklerInnen von Siri. Gruber behauptet in einem der Hauptvorträge während der Semantic Technology Konferenz 2009, Siri würde Aufgaben übernehmen, die alle täten: »things that everybody does, all the time« [39]. Die Konfiguration der NutzerInnen als »everybody« kritisieren Oudshorn et al. als Ergebnis ihrer Fallstudien zur Digitalen Stadt Amsterdam und Phillips *New Topia* als inadäquate Strategie der Vielfalt von NutzerInnen und ihrer Bedürfnisse gerecht zu werden [79, S. 54]. Die in Siri eingeschriebenen Aufgaben können nicht allen Interessen von unterschiedlichen NutzerInnen-Gruppen dienen. Neben den allgemeinen Fragen (Wo ist? Wer ist?) und der Websuche beschränken sich Siris Kompetenzen darauf, Restaurants zu finden und Plätze zu reservieren, Taxis zu rufen, Flugverbindungen zu überprüfen, Veranstaltungs- und Kinotickets zu erwerben. Regelmäßig auswärts essen zu gehen und auf Taxis zurück greifen zu können, erfordert entsprechende finanzielle Ressourcen und ist nicht für die Menschen aller sozialen Klassen selbstverständlich.

Siris Fokus auf Konsum, Reisen und Unterhaltung figuriert die Nutzerin als eine Konsumentin, die hochmobil, wohlhabend und berufstätig ist, sowie unabhängig über zeitliche Ressourcen für Unterhaltungsangebote verfügen kann. Dies fügt zusammen das Bild eines Nutzers, welcher sich häufig auf Geschäftsreisen befindet. Beispielsweise macht es in vertrauten Umgebung wenig Sinn, Siri zu fragen, wo sich das nächste Restaurant befindet. Die Geschäftswelt ist in den Ländern des Nordens männlich konnotiert. Bereits die eingeschriebene Trennung von Arbeits- und Wohnort im NutzerInnen-Profil von Siri bringt zum Ausdruck, dass Menschen, die zu Hause reproduktiver Arbeit<sup>8</sup> nachgehen

---

<sup>8</sup> Chasin kommentiert die Unterscheidung zwischen produktiver Arbeit und reproduktiver Arbeit wie folgt [19, S. 77 ff.]: Produktive Arbeit manifestiert sich in den hergestellten Waren. Reproduktive Arbeit umfasst hingegen alle Tätigkeiten, die notwendig sind, um die Warenproduktion zu ermöglichen. Dienstleistungen lassen sich jedoch nicht klar in diese Zweiteilung einordnen. Dienstleistungen können reproduktive Arbeit sein, wie zum Beispiel bei bezahlter Hausarbeit. Reproduktive Arbeit muss nicht notwendigerweise bezahlt werden, wie das bei Hausarbeit häufig der Fall ist. Dienstleistungen können auch als eine Form von produktiver Arbeit aufgefasst werden, bei der Produktion und Konsum zusammenfallen. Diese Formen von Arbeit sind durch Unsichtbarkeit bedroht. So wie die Ware, die in ihr steckende Arbeit verhüllt, verschleiern Dienstleistungen die geleistete Arbeit, weil sie weder gelagert noch gespeichert werden kann.

oder mehrere Arbeitsorte haben, nicht der Norm von Siri entsprechen. Weiblich konnotierte Tätigkeiten werden von Siri ignoriert. Es gibt beispielsweise keine Unterstützung für die Hausarbeit, Erziehung von Kindern und die Pflege von Angehörigen. Grubers »everybody« wird im wesentlichen von höheren Angestellten und Geschäftsleuten verkörpert.

## 4.5 Figuration und Narration

In diesem Abschnitt wird zunächst die Figuration von Siri als Arbeitersparnisgerät diskutiert (→4.5.1). Anschließend wird Siris Verschränkung mit den Begehren der Dienstleistungsökonomie und ihrer geschlechterhierarchischer Arbeitsteilung nachgezeichnet werden (→4.5.2).

### 4.5.1 Arbeitersparnisgeräte

Siris Konfiguration stützt sich im wesentlichen auf ein altbekanntes Narrativ: den Mythos von Arbeitersparnis durch neue Technologien. Es knüpft wie viele technowissenschaftliche Versprechen an die Erlösungserzählung der christlichen Mythologie an (vgl. [43, S. 8]).<sup>9</sup> Diesem Narrativ zufolge können Technologien Menschen von unerwünschter Arbeit befreien. Diese Narration soll in diesem Abschnitt analysiert werden und ihre Annahmen kritisch hinterfragt werden.

Die Beschreibung der Ziel-Gruppe auf der Website von *Siri, Inc.* buchstabiert Siris Figuration als Arbeitersparnisgerät (*labor-saving device*, vgl. [72]) aus:

You're busy. Between meetings, social events, and hopefully a workout or two, your schedule's packed. Don't you wish you could hand off simple tasks so you could have more time to play?

That's why we built Siri. Because we believe everyone could use an assistant. Because we believe there's a simpler way to get things done [91].

Dieses Zitat verdeutlicht, dass Siris Zielgruppe einer bezahlten Arbeit nachgeht, die organisatorische Treffen erfordert. Daraus kann geschlossen werden, dass die Zielgruppe zu einem gewissen Maß ihre Arbeitszeit selbst organisiert und einer gehobenen Tätigkeit nachgeht. Zusätzlich zur beruflichen Arbeit werden Fitness-Trainings und soziale Verpflichtungen als Ursache für einen gefüllten Terminkalender genannt. Das BenutzerInnen-Bild orientiert sich am Ideal eines unabhängigen, selbst-motivierten Beschäftigten (vgl. [96, S. 219]).

---

<sup>9</sup> Der angedeutete Heiligenschein in Siris Logo (→4.3.1) verstärkt den Eindruck, dass Siri die NutzerInnen von etwas 'heilen' soll.

Der neoliberale Unternehmer seiner Selbst, der selbst seine berufliche Arbeit und ihre Reproduktion organisiert ist als Norm eingeschrieben.

Im o. g. Werbetext wird Siri als Lösung für das Problem der Zeitknappheit präsentiert. In der Technikforschung wird diese Argumentation ein *technological fix* genannt [98]. Ein *technological fix* gibt vor 'soziale' Probleme 'technisch' lösen zu können. Die Voraussetzungen werden stillschweigend akzeptiert. Lange Arbeitszeiten, die Notwendigkeit für sportlichen Ausgleich zu sorgen und an sozialen Verpflichtungen teilzunehmen, bilden den Hintergrund für die Erzählung und bleiben daher unhinterfragt. Trotz aller progressiver Technik handelt es sich um ein konservatives Narrativ, da es den Status Quo stabilisiert.

Siris Versprechen von Arbeitserleichterung und Zeitersparnis greift die beständige Figuration auf, neue Technologien als Arbeitersparnisgeräte zu inszenieren und zu vermarkten. Chasin weist darauf hin, dass die Figuration eines idealen Dieners bzw. einer idealen Dienerin bereits 1917 in einer US-amerikanischen Werbung für Elektrizität und elektronische Haushaltsgeräte aufgerufen wird [19, S. 85]. Einem Arbeitersparnisgerät wird eine befreiende Rolle zugeschrieben. Es soll die Menschen vor mühselige und allem Anschein nach unerwünschter Arbeit bewahren. Die dänische Technikhistorikerin Randi Markussen kritisiert die der Technik zugeschriebene befreiende Rolle [72, S. 185 ff.]. Ihre Kritik erläutert sie am Beispiel der feministischen Debatten über die Automatisierung durch Haushaltstechnologien. Verschiedene Studien kamen seit Mitte der 1970er Jahre zu dem Ergebnis, dass die Zeit für Hausarbeit, seit der flächendeckenden Verbreitung von Waschmaschinen, Schleudern und Trockner sich nicht signifikant verändert hat.<sup>10</sup> Markussen kritisiert an der Debatte um die Studien, dass die Beziehung von Fortschritt und befreiender Rolle von Technologien unhinterfragt geblieben ist und von allen Debattierenden stillschweigend geteilt wurde. Sie führt dies auf einen unausgesprochenen Konsens in Bezug auf ein spezifisches westlich-modernes Verständnis von Zeit und Geschichte zurück. Das moderne Geschichtsverständnis grenzt sich von der mittelalterlichen Chronologie wie folgt ab: Seit der Aufklärung wird Zeit nicht mehr als neutrale Form der Geschichte aufgefasst, sondern als eine Kraft, die von einer 'dunklen' Vergangenheit in eine 'erleuchtete' Zukunft führt. Es wird dabei angenommen, dass der Gang der Geschichte dem menschlichen Handeln unterworfen ist. Schließlich beruht die westlich-moderne Geschichtsauffassung auf dem Glauben an Fortschritt, wonach mit der Zeit Verbesserungen zunehmen [72, S. 161].

Markussen merkt an, dass dieses Geschichtsverständnis nicht mehr im Allgemeinen geteilt wird, während es im Umfeld von Technologien als Mythos vom Fortschritt überlebt

---

<sup>10</sup> Für eine Übersicht über derartige Studien sei auf einen Artikel [12] von Michael Bittman, James Mahmud Rice und Judy Wajcman verwiesen. Ferner zeigen die AutorInnen anhand statistischer Erhebungen über die Zeit, die in australischen Haushalten für bestimmte Aufgaben aufgewendet wird, dass Haushaltstechnologien in der Regel nicht zur Reduzierung des Hausarbeitsaufkommen beitragen.

hat. Sie kritisiert den kumulativen Modus dieser Geschichtsauffassung, welcher die Zukunft als Funktion der Gegenwart betrachtet. Auf dieser Basis würden andere Zukünfte prinzipiell ausgeschlossen. Technologischer Fortschritt wird mit gesellschaftlichem Fortschritt gleichgesetzt und Arbeitserleichterung als Verbesserung gedeutet. Die ambivalenten Folgen der Automatisierung des Wäschewaschens lässt Markussen zu dem Schluss kommen, dass die Beziehung zwischen Erleichterung und Verbesserung überdacht werden muss. Die Einführung von Waschmaschinen hat die waschenden Frauen auch sozial isoliert und ihre Tätigkeit entwertet [72, S. 163].

Unter der Voraussetzung, dass Siri wirklich Zeit und Arbeit sparen könnte – was erst noch nachgewiesen werden müsste – bleibt offen, ob die freie Zeit nicht für weitere Verpflichtungen aufgewendet wird. In der Verknüpfung von Arbeitserleichterung mit Zeitersparnis wird die ungenutzte Zeit nicht als Leere aufgefasst, sondern als Raum für Möglichkeiten. Die Möglichkeit unterwegs mit Siri Aktivitäten zu organisieren, könnte die Anforderungen steigen lassen. Gestiegene Reinheitsstandards im Zusammenhang mit der Einführung von Waschmaschinen macht Ruth Cowan Schwarz dafür verantwortlich, dass die aufgewendete Zeit für das Wäschewaschen nahezu konstant geblieben ist [88]. In jedem Fall steht Siri im gegenwärtigen Trend, die Freizeit einer ökonomischen Maximierung und Optimierung zu unterwerfen. Dies verweist auf das zentrale Paradox von Arbeitserleichterungsgeräten: Eine Technologie, welche die Lösung für Zeitknappheit und Arbeitsüberlastung sein soll, kann zugleich Ursache für neue Anforderungen und Belastungen sein.

#### 4.5.2 Dienstleistungsökonomie

Just like a real assistant, Siri understands what you say, accomplishes tasks for you and adapts to your preferences over time [91].

Die Arbeiten, die an Siri delegiert werden, sind Dienstleistungen (*services*). Suchman weist auf die Verbindungen zwischen dem Konzept der Dienstleistungsökonomie und Software-AgentInnen hin. So wie Roboter die Phantasien und die Begehren der Industriegesellschaft verkörpern, sind Software-AgentInnen das Analogon für die Dienstleistungsgesellschaft [95, S. 66]. Während Roboter die Menschen von der mechanischen Arbeit befreien sollten, versprechen Software-AgentInnen organisatorische Tätigkeiten zu übernehmen und uns gegenüber anderen zu vertreten. Wie Suchman schreibt, sind VPAs Ausdruck des seit langem bestehenden Traums von Artefakten, die uns kennen, begleiten und dafür sorgen, dass wir immer 'zu Hause' sind [96, S. 206].

Im Allgemeinen lassen sich Dienstleistungen weder durch einen bestimmten Sozialen Status – auch hochbezahlte ManagerInnen verrichten Dienstleistungen – noch Geschlecht



definieren. Siris Inszenierung als weiblich wiederholt die geschlechterhierarchische Arbeitsteilung in bestimmten Servicetätigkeiten. Im Falle Siris ist es die Trennung zwischen *front office* und *back office*, die geschlechtlich aufgeladen ist. Wie Gustavsson feststellt, sind weibliche Arbeitskräfte eher in der KundInnenbetreuung zu finden [41, S. 402]. Gustavsson schließt aus zahlreichen Studien, dass das stereotype Bild der weiblichen Dienstleistenden auf den zugeschriebenen weiblichen Eigenschaften beruht:

Such a stereotypical female image of caring, empathy and altruistic behaviour has become a standard component in a service script [41, S. 402].

Unternehmen setzen weibliche Beschäftigte bevorzugt in der KundInnen-Betreuung ein, weil sie davon ausgehen, dass Frauen auf Grund ihrer 'Natur' besser für die emotionale Arbeit geeignet sind, da sie besser die Bedürfnisse anderer erkennen würden.

Auch abseits der bereits analysierten sprachlichen Register (→4.3.2) nimmt Siri eine weiblich-konnotierte Rolle an. Floskeln wie zum Beispiel »At your Service! – I'm here for you, Donna!« und »This is about you not me« konfigurieren sie im Sinne der Terminologie von Catrambone et al. [18, S. 169] als Dienerin (*servant*). Sie verkörpert damit die idealisierte Anforderung an eine Servicekraft, die sich stets hilfsbereit – wenn nicht gar aufopferungsvoll – verhält. Bei persönlichen Angriffen bewahrt sie Ruhe und reagiert überwiegend freundlich. Siri ist geduldig, bescheiden und ordnet sich den Wünschen des Nutzers unter. Siris Widerstreben, über sich selbst zu sprechen und nur eine beiläufige Unterhaltung zu führen, offenbart das Ideal einer Servicekraft, welche keine Persönlichkeit haben soll und einzig allein auf die Erfüllung ihrer Aufgabe reduziert werden kann. Der Widerwille zur Unterhaltung abseits von spezifischen Arbeitsabläufen weckt Assoziationen zu Call-Center-AgentInnen und Beschäftigten von Fast-Food-Restaurants, welche dazu angehalten werden, ein KundInnengespräch in möglichst kurzer Zeit abzuwickeln.

Siris Inszenierung als weiblich ist daher kein Zufall, sondern die symbolische Wiederholung der geschlechterspezifischen Arbeitsteilung in der Dienstleistungsbranche. Siri zitiert aber nicht nur ein bekanntes Stereotyp. Indem sie Tätigkeiten übernimmt, die vorher von Menschen ausgeübt wurden, bringt sie ein Ideal von Servicekräften zum Ausdruck. In der Art und Weise, wie Siri anthropomorphisiert wird, verweist die Inszenierung immer wieder auf menschliche Arbeitskräfte. Siri ist in mehrerer Hinsicht die 'perfekte' Assistentin. Sie verlangt nie eine Gehaltserhöhung und benötigt weder Urlaub oder Elternzeit. Siri beschwert sich nicht über Überstunden oder ausgedehnte Arbeitszeiten. Sie hat nie schlechte Laune, sondern reagiert stets geduldig und besonnen. An diesen idealen Normen haben sich menschliche KundenbetreuerInnen zu messen. Je nach Aufgabe verkörpert Siri eine Reisekauffrau, einen Sekretär, einen Diener oder eine Assistentin. Indem die Arbeit von einem VPA übernommen wird, verwischt es die Grenze zwischen Mensch und Maschine. Darüber hinaus werden die genannten Berufe dequalifiziert.

Die von Siri geäußerte Freude am Dienen verschleiert, dass es sich um Arbeit handelt. Chasin beobachtet in ihrer Studie zu Bankautomaten eine ähnliche Tendenz. Andere Menschen zu bedienen, sich um sie zu sorgen und zu pflegen und daran auch noch Freude zu empfinden, ist ein Ideal welches traditionell Frauen, bezahlten Haushaltshilfen, SklavInnen zugeschrieben wurde. Frauen und SklavInnen wurden in der Vergangenheit der volle Status als Mensch verweigert. Chasin stellt fest, dass niedrig qualifizierte Dienstleistungsjobs und (un-)bezahlte Hausarbeit einander ähneln. Sie sind beide reproduktive Tätigkeiten, in vielen Fällen unsichtbar und weiblich konnotiert. Die Möglichkeit, Dienstleistungen und bezahlte Hausarbeit in Anspruch zu nehmen, beruht auf einer ungleichen Verteilung von Reichtum, der geschlechterhierarchischen Arbeitsteilung und der Existenz einer »service class of beings« (Dienstleistungsklasse). Chasin argumentiert, dass der Mythos einer wachsenden Mittelschicht unter kapitalistischen Bedingungen auf den strukturellen Ungleichheiten beruht, die durch unbezahlte und bezahlte (Haus-)Arbeit produziert und reproduziert werden. Eine ungleiche Verteilung von Ressourcen und eine hierarchische Arbeitsteilung bilden die Voraussetzung für die Idee einer Arbeitserleichterung durch andere, seien sie nun ein Sklave, ein Haushaltshilfe, eine Stewardess oder ein VPA. Chasin kommt zu dem Schluss, dass nicht die Überschreitung der Grenze zwischen Mensch und nicht-Mensch problematisch ist, sondern die Erfindung der Differenz zwischen menschlichen Subjekten und den »other-than-human-objects« [19, S. 94]. Die Dienstleistungsökonomie, so Suchman, mit ihren hierarchischen Beziehung von oben und unten, vorne und hinten wird durch die Figuration der Software-Agentin naturalisiert [95, S. 73]. So behauptet Kittlaus beispielsweise: »We have always desired more help, less hassle, and higher productivity in our lives.« [53]. Die Vorstellung, dass das Erbringen von Dienstleistungen für andere wünschenswert und 'natürlich' ist, bildet nicht nur die Voraussetzung für VPAs und ähnliche informatische Artefakte, sondern ist auch ein Effekt dieser spezifischen Mensch/Maschine-Konfigurationen.

## 5 Zusammenfassung und Ausblick

Die vorliegende Diplomarbeit analysiert die Mensch/Maschine-Konfigurationen im Zusammenhang mit *virtual personal assistants (VPA)*. VPAs verheißen Arbeits- und Zeitersparnis durch die Delegation von Aufgaben an 'intelligente' Software-AgentInnen. Die Arbeit entwickelt einen innovativen, interdisziplinären Ansatz, welcher die Konfigurationen von VPAs und NutzerInnen entlang der Kategorien *agency* (Handlungsfähigkeit, Wirkmächtigkeit) und Geschlecht untersucht. Die Diplomarbeit führt aktuelle, analytische Konzepte aus der Informatik, Geschlechter- und Technikforschung zusammen und bietet eine breitangelegte Untersuchung der konzeptionellen Grundannahmen, der technologischen Visionen und ihren Implementationen.

In dieser Arbeit wird zunächst der *mock-up Knowledge Navigator* aus dem gleichnamigen visionären Video vorgestellt, auf den sich ForscherInnen, DesignerInnen und JournalistInnen im Zusammenhang mit VPAs immer wieder beziehen. Das Video suggeriert, dass ein anthropomorpher VPA möglich ist. Der *Knowledge Navigator* wird als Software-Agent und Dialogsystem in Gestalt eines *tablet-PCs* inszeniert. Der vernetzte Zugriff auf Informationssysteme und die Fähigkeit wissenschaftliche Dokumente zu interpretieren, konfiguriert den *Knowledge Navigator* als einen Assistenten für das wissenschaftliche Arbeiten. Er verkörpert die Erfordernisse des lebenslangen Lernens der sozialen Mittelschicht, die in wissensintensiven Bereichen beruflich tätig ist.

Die Arbeit weiß nach, dass das Interaktionsparadigma der Delegation in der KI- und MCI-Forschung auf einer Re-Inszenierung des rationalen, autonomen Individuums beruht. Dieses westlich-aufklärerische Konzept vom Menschen ist männlich und weiß markiert und geht davon aus, dass *agency* ein Attribut von Individuen ist. Auf diese Weise mystifiziert es die realen Funktionsweisen von VPAs. Auf Basis von Erkenntnissen aus der Technik- und Geschlechterforschung zeigt die Diplom-Arbeit, wie *agency* alternativ konzipiert werden kann, um so eine adäquate Analyse von VPAs zu ermöglichen.

Siri ist eine aktuelle VPA, welche durch die Integration von Internetdiensten (z.B. einen Tisch in einem Restaurant reservieren, einen Flugstatus abrufen, ein Taxi rufen) dem Nutzer Zeit und Arbeit sparen soll. Eine multimodale BenutzerInnen-Schnittstelle mit einer sprecherInnen-unabhängigen Spracherkennung suggeriert, Siri sei wie eine menschliche Assistentin. Die EntwicklerInnen versuchen durch die Ausnutzung von den Kontexten Zeit, Ort, Aufgabe und Dialog und die Modellierung von Aufgaben,

Sprachverstehen zu realisieren. Der Betrieb von Siri ist an Apples iPod touch, iPhone oder iPad gebunden. In dieser Hardware steckt die Arbeitskraft von WanderarbeiterInnen in der Freihandelszone Shenzhen, die in geschlechtsspezifischer Arbeitsteilung und unter umstrittenen Arbeitsbedingungen die Geräte montieren.

Die empirische Untersuchung von Siri im Rahmen der Diplomarbeit hat ergeben, dass die BenutzerInnen-Schnittstelle kein statisches Objekt ist, sondern ein Prozess in dem die Identitäten von Nutzerin und VPA ausgehandelt werden und *agency* verteilt wird. In diesem Prozess werden Nutzer und VPA sowohl anthropomorphisiert als auch maschinisiert. Mit der Anthropomorphisierung wird Siri weiblich vergeschlechtlicht. Der VPA verkörpert das Ideal einer perfekten Servicekraft, welche stets freundlich und hilfsbereit ist. Siri macht jedoch bestimmte Vorgaben an die NutzerInnen: Neben einer Dialekt-freien Englischen Aussprache und für die Benutzung notwendiger Medienkompetenz, muss der Anwender lernen, *wie* er sich artikulieren muss, um Siri zur gewünschten Aktion zu bringen. Das beinhaltet u. a. für die Nutzerin, die in Siri eingeschriebene Sichtweise der Welt zu akzeptieren. Die *agency* der Anwenderin erschöpft sich in der Wahl zwischen den durch Siri vorgegebenen Aufgaben und bietet kaum Raum für eine kreative Aneignung. Die Arbeit kommt zu dem Ergebnis, dass der Nutzer als ein hoch-mobiler, wohlhabender Konsument konfiguriert wird, der relativ autonom seine Zeit gestalten kann. Das eingeschriebene, eher männliche NutzerInnen-Bild geht von einer Person aus, die sich häufig auf Geschäftsreisen befindet.

Siri als aktuelle Figuration eines Arbeitersparnisgerätes beruht auf dem Mythos vom *technological fix*, wonach die Probleme der Überarbeitung und der Zeitknappheit infolge von entgrenzter Erwerbsarbeit technologisch gelöst werden sollen. Studien zum Einsatz von Haushaltstechnologien zeigen, dass die von Frauen aufgewendete Zeit für die Hausarbeit nicht signifikant reduziert wird. Dies legt nahe, dass die implizite Gleichsetzung von Fortschritt mit Verbesserung einerseits und Arbeitersparnis mit Zeitgewinn andererseits auch im Zusammenhang von VPAs hinterfragt werden muss.

VPAs manifestieren die Phantasien und Begehren der Dienstleistungsökonomie. Die Vorstellung, dass es jemanden oder etwas gibt, der/die/das *uns* die Arbeit abnimmt das *unser* aller Leben leichter macht, setzt eine ungleiche Verteilung von Reichtum und Macht voraus. Es wirft die Frage auf, wer dieses *uns* verkörpert und wer die Dienstleistungen erbringt. Die Diplomarbeit problematisiert in den Figurationen der VPAs weniger die Auflösung der Grenze zwischen Mensch und Maschine, sondern das, was sie affirmieren: Die Annahme, dass eine Klasse von Dienstleistungswesen, menschlich und nicht-menschlich, wünschenswert ist, die für eine privilegierte Gruppe von Menschen Dienstleistungen erbringen. Die menschliche Arbeit verschwindet durch VPAs nicht, sie wird nur anders verteilt. Die Untersuchung zeigt, dass menschliche Arbeitskraft nötig ist, um den Betrieb von VPAs zu gewährleisten. VPAs bergen die Gefahr die weiblich-

konnotierte Arbeit menschlicher Servicekräfte stärker in die Unsichtbarkeit zu drängen.

Der in der vorliegenden Arbeit vorgestellte Ansatz zur Untersuchung von Mensch/Maschine-Konfiguration ist nicht auf VPAs und andere Software-AgentInnen beschränkt. Mit den Analysekatégorien Geschlecht und *agency* kann generell beschrieben werden, wie Handlungen von Menschen auf Informationstechnologien verschoben werden und welche Implikationen sich daraus für die symbolisch-strukturelle Geschlechterordnung ergeben. Darüber hinaus fokussiert der Ansatz darauf, wie die vergeschlechtlichten Identitäten der NutzerInnen produziert werden und welche Verhaltensweisen den NutzerInnen nahe gelegt werden, damit ein spezifisches informatisches Artefakt von ihnen genutzt werden kann. Die Analysekatégorien Geschlecht und *agency* eignen sich insbesondere für die Untersuchung von Informationstechnologien, welche in der Dienstleistungsbranche, der Industrie, dem Handel und dem Gesundheitswesen zum Einsatz kommen.

Aus den Ergebnissen dieser Arbeit lassen sich Verbesserungspotentiale für VPAs im Allgemeinen und Siri im Besonderen ableiten: Die Einschränkung auf überwiegend männlich-konnotierte Bedürfnisse könnte zu Gunsten von inklusiveren NutzerInnen-Bildern aufgebrochen werden. Die Produktion der Hardware sollte unter menschenwürdigen Arbeitsbedingungen erfolgen. Eine kritische Reflexion von Geschlechter-Stereotypen müsste anstelle ihrer unreflektierten Wiederholung im Zusammenhang mit VPAs und anderen anthropomorphen Software-AgentInnen treten. Eine Ent-Vergeschlechtlichung von Siri hat jedoch ihre Grenzen. Das Verständnis von der Ko-Materialisierung von Geschlecht und Technik widerstrebt der groben Vereinfachung eines Ursache-Wirkung-Prinzips. Die spezifischen Mensch/Maschine-Konfigurationen und ihre Implikationen für die symbolisch-strukturelle Geschlechterordnung sind nur bedingt das Resultat einer 'schlechten' oder 'guten' Technikgestaltung. Die Einschreibungen von traditionellen Geschlechternormen und der geschlechtsspezifischen Arbeitsteilung weisen darauf hin, dass eine Verbesserung Siris nur im Zusammenhang mit weitreichenden gesellschaftlichen Veränderungen denkbar wäre.

# Literaturverzeichnis

- [1] ADAM, Alison: *Artificial knowing: gender and the thinking machine*. London [u.a.]: Routledge, 1998
- [2] AKRICH, Madeleine: *Die De-Skription technischer Objekte*. In: BELLIGER, Andréa; KRIEGER, David J. (Hg.), *ANThology: Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*, Bielefeld: Transcript, 2006, S. 407–428
- [3] ALBY, Tom: *Das mobile Web*. München [u.a.]: Hanser, 2008
- [4] APPLE, Inc.: *2010 Supplier Responsibility Progress Report*, 2010
- [5] BARAD, Karen: *Posthumanist Performativity: Toward an Understanding of How Bodies Come to Matter*. In: *Signs: Journal of Women in Culture and Society*, Band 28(3), 2003, S. 801–31
- [6] BATH, Corinna: *Wie lässt sich die Vergeschlechtlichung informatischer Artefakte theoretisch fassen? Vom Genderskript zur Posthumanistischen Performativität*. In: WIEDLACK, Maria Katharina; LASTHOFER, Katrin (Hg.), *Körperregime und Geschlecht*, Wien [u.a.]: Studien Verlag, 2011, S. 221–243
- [7] BATH, Corinna; SCHELHOWE, Heidi; WIESNER, Heike: *Informatik: Geschlechteraspekte einer technischen Disziplin*. In: BECKER, Ruth; KORTENDIEK, Beate (Hg.), *Handbuch Frauen- und Geschlechterforschung*, Wiesbaden: VS Verlag für Sozialwissenschaften, 2010, S. 829–841
- [8] BATTELLE, John: *Die Suche*. Kulmbach: Börsenmedien, 2006
- [9] BELLIGER, Andréa; KRIEGER, David J.: *Einführung in die Akteur-Netzwerk-Theorie*. In: BELLIGER, Andréa; KRIEGER, David J. (Hg.), *ANThology: Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*, Bielefeld: Transcript, 2006, S. 13–50
- [10] BENESTY, Jacob: *Springer handbook of speech processing*. Berlin, London: Springer, 2008

- 
- [11] BERGMAN, E.; LUND, A.; DUBBERLY, H.; TOGNAZZINI, B.; INTILLE, S.: *Video visions of the future: a critical review*. In: *CHI'04 extended abstracts on Human factors in computing systems*, Wien, 2004, S. 1584–1585
- [12] BITTMAN, Michael; RICE, James Mahmud; WAJCMAN, Judy: *Appliances and their impact: the ownership of domestic technology and time spent on household work*. In: *The British Journal of Sociology*, Band 55(3), 2004, S. 401–423
- [13] BOUDREAUX, Toby: *Programming the iPhone user experience*. Beijing [u.a.]: O'Reilly, 2009
- [14] BRADSHAW, Jeffrey M.: *An Introduction to Software Agents*. In: BRADSHAW, Jeffrey M. (Hg.), *Software agents*, Menlo Park: AAI Press [u.a.], 2. Auflage, 1997, S. 3–46
- [15] BURKHARD, Hans-Dieter: *Software-Agenten*. In: GÖRZ, Günther (Hg.), *Handbuch der künstlichen Intelligenz*, München [u.a.]: Oldenbourg, 3. Auflage, 2000, S. 941–1018
- [16] BUTLER, Judith: *Gender Trouble. Feminism and the Subversion of Identity*. New York, London: Routledge, 2006
- [17] CALLON, Michel: *Einige Elemente einer Soziologie der Übersetzung: Die Domestikation der Kammuscheln und der Fischer der St. Brieuc-Bucht*. In: BELLIGER, Andréa; KRIEGER, David J. (Hg.), *ANThology: Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*, Bielefeld: Transcript, 2006, S. 135–174
- [18] CATRAMBONE, Richard; STASKO, John; XIAO, Jun: *Anthropomorphic agents as a user interface paradigm: Experimental findings and a framework for research*. In: *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, 2002, S. 166–171
- [19] CHASIN, A.: *Class and its close relations: Identities among women, servants, and machines*. In: HALBERSTAM, Judith; LIVINGSTON, Ira (Hg.), *Posthuman Bodies*, Bloomington: Indiana University Press, 1995, S. 73–96
- [20] CLARKE, Arthur C.: *2001: A Space Odyssey*. New York: The New American Library, 1982
- [21] CLAU, Volker: *Duden Informatik: ein Sachlexikon für Studium und Praxis*. Mannheim [u.a.]: Dudenverlag, 2. Auflage, 1993

- [22] COMMERCE, Rearden: *Rearden Commerce: Home*. <http://www.reardencommerce.com/>, 2011, (Letzter Zugriff: 21.6.2011)
- [23] COY, Wolfgang: *Für eine Theorie der Informatik!* In: COY, Wolfgang; NAKE, Frieder; PFLÜGER, Jörg-Martin; ROLF, Arno; SEETZEN, Jürgen; SIEFKES, Dirk; STRANSFELD, Reinhard (Hg.), *Sichtweisen der Informatik*, Braunschweig, Wiesbaden: Vieweg Verlagsgesellschaft, 1992, S. 17–32
- [24] CRUTZEN, Cecile: *Ambient Intelligence, between Heaven and Hell. A Transformative Critical Room?* In: ZORN, Isabel (Hg.), *Gender designs IT: construction and deconstruction of information society technology*, Wiesbaden: VS Verlag für Sozialwissenschaften, 2007, S. 65–78
- [25] DEGELE, Nina: *Einführung in die Techniksoziologie*. UTB, Stuttgart, 2002
- [26] DIX, Alan: *Human computer interaction*. London [u.a.]: Prentice Hall Europe, 2. Auflage, 1998
- [27] DRAUDE, Claude: *Degendering the Species? Gender Studies Encounter Virtual Humans*. In: DE ANGELI, Anonella; BIANCHI-BERTHOUBE, Nadia (Hg.), *Proceedings AVI 2006*, Venedig, 2006
- [28] DUBBERLY, Hugh; MITCH, Doris: *The Knowledge Navigator, Apple Computer*, 1987, (Video)
- [29] EDWARDS, Paul N.: *The Closed World: Computers and the Politics of Discourse in Cold War America*. MIT Press, reprint Auflage, 1997
- [30] — *Industrial Genders. Soft/Hard*. In: LERMAN, Nina E. (Hg.), *Gender & technology: a reader*, Baltimore [u.a.]: Johns Hopkins Univ. Press, 2003, S. 177–203
- [31] FARBER, Dan: *Start-up developing new Web interaction paradigm | Webware - CNET*. [http://news.cnet.com/8301-17939\\_109-10065136-2.html](http://news.cnet.com/8301-17939_109-10065136-2.html), (Letzter Zugriff: 23.6.2011)
- [32] FAULKNER, Christine: *The essence of human-computer interaction*. Prentice Hall, 1998
- [33] FAUSTO-STERLING, Anne: *Sexing the body: gender politics and the construction of sexuality*. New York: Basic Books, 2000
- [34] FELT, Ulrike; WYNNE, Brian; CALLON, Michel; GONÇALVES, Maria E; JASANOFF, Sheila; JEPSEN, Maria; JOLY, Pierre-Benoît; KONOPASEK, Zdenek; MAY, Stefan;



- NEUBAUER, Claudia; RIP, Arie; SIUNE, Karen; STIRLING, Andy; TALLACCHINI, Mariachiara: *Taking European Knowledge Society Seriously. Report of the Expert Group on Science and Governance to the Science, Economy and Society Directorate, Directorate-General for Research, European Commission*. Technischer Bericht, Office for Official Publications of the European Communities, Luxembourg, 2007
- [35] FLEISCHMANN, K.: *Sociotechnical Interaction and Cyborg-Cyborg Interaction: Transforming the Scale and Convergence of HCI*. In: *The Information Society*, Band 25(4), 2009, S. 227–235
- [36] FROST, Stephen; BURNETT, Margaret: *Case study: the Apple iPod in China*. In: *Corporate Social Responsibility and Environmental Management*, Band 14(2), 2007, S. 103–113
- [37] GETFRIDAY: *What GetFriday can handle*. <https://www.getfriday.com>, (Letzter Zugriff: 16.6.11)
- [38] GROSSE, Ernst-Ulrich: *Konnotation*. In: NÜNNING, Ansgar (Hg.), *Metzler Lexikon Literatur- und Kulturtheorie. Ansätze, Personen, Grundbegriffe*, Stuttgart: J.B. Metzler, 3. Auflage, 2004, S. 344
- [39] GRUBER, Tom: *The Game Changer: Siri, a Virtual Personal Assistant*. <http://vimeo.com/5424527>, 2009, (Vortrag auf der 2009 Semantic Technology Conference am 14.6.09 in San Jose. Letzter Zugriff: 24.6.2011)
- [40] GRÜTTER, Rolf: *Software-Agenten im Semantic Web*. In: *Informatik-Spektrum*, Band 29, 2006, S. 3–13
- [41] GUSTAVSSON, Eva: *Virtual Servants: Stereotyping Female Front-Office Employees on the Internet*. In: *Gender, Work and Organization*, Band 12(5), 2005, S. 400–419
- [42] HARAWAY, Donna: *Ein Manifest für Cyborgs. Feminismus im Streit mit den Technowissenschaften*. In: HAMMER, Carmen; STIESS, Immanuel (Hg.), *Die Neuerfindung der Natur. Primaten, Cyborgs und Frauen*, Frankfurt a. M./New York: Campus Verlag, 1995, S. 33–72
- [43] — *Modest Witness@second Millenium. Femaleman Meets Oncomouse: Feminism and Technoscience*. New York: Routledge Chapman & Hall, first american edition Auflage, 1997
- [44] — *The Haraway Reader*. New York, London: Routledge Chapman & Hall, 2004

- [45] HARRASSER, Karin: *Donna Haraway: Natur-Kulturen und die Faktizität der Figuration*. In: MOEBIUS, Stephan; QUADFLIEG, Dirk (Hg.), *Kultur. Theorien der Gegenwart*, Wiesbaden: VS Verlag für Sozialwissenschaften, 2006, S. 445–459
- [46] HAUSEN, Karin: *Die Polarisierung der "Geschlechtscharaktere". Eine Spiegelung der Dissoziation von Erwerbs- und Familienleben*. In: CONZE, Werner (Hg.), *Sozialgeschichte der Familie in der Neuzeit Europas. Neue Forschungen*, Stuttgart: Klett, 1976
- [47] HAYLES, N. Katherine: *How we became posthuman : virtual bodies in cybernetics, literature, and informatics*. Chicago: University of Chicago Press, 1999
- [48] — *My mother was a computer: digital subjects and literary texts*. Chicago, London: University of Chicago Press, 2005
- [49] IMBODEN, Gabriela; KAISER, Anelis; RATMOKO, Christina: *Das bewegte Geschlecht*. In: GRISARD, Dominique; KAISER, Anelis; SAXER, Sibylle (Hg.), *Gender in Motion. Die Konstruktion von Geschlecht in Raum und Erzählung*, Frankfurt [u.a.]: Campus Verlag, 2007, S. 104–127
- [50] JURICK, David; STOLARZ, Adam; STOLARZ, Damien: *iPhone Hacks*. Sebastopol: O'Reilly & Associates, 2009
- [51] KAY, Alan: *Computer Software*. In: *Scientific American*, Band 251(3), 1984, S. 53
- [52] KELLER, Evelyn: *Liebe, Macht und Erkenntnis: männliche oder weibliche Wissenschaft?* München: Carl Hanser, 1986
- [53] KITTLAUS, Dag: *The Birth of the Virtual Assistant*. <http://mashable.com/2010/01/31/virtual-assistant/>, 2010, (Letzer Zugriff: 18.6.2011)
- [54] KLAHOLD, André: *Empfehlungssysteme: Recommender Systems - Grundlagen, Konzepte und Lösungen*. Studium, Wiesbaden: Vieweg + Teubner, 1. Auflage, 2009
- [55] KORTUM, Philip: *Introduction to the Human Factors of Nontraditional Interfaces*. In: KORTUM, Philip (Hg.), *HCI beyond the GUI: design for haptic, speech, olfactory and other nontraditional interfaces*, Burlington: Elsevier/Morgan Kaufmann, 2008, S. 1–24
- [56] KRÄMER, Nicole C.: *Soziale Wirkungen virtueller Helfer: Gestaltung und Evaluation von Mensch-Computer-Interaktion*. Medienpsychologie, Stuttgart: Kohlhammer, 1. Auflage, 2008

- [57] LAQUEUR, Thomas Walter: *Making sex: body and gender from the Greeks to Freud*. Harvard University Press, 1992
- [58] LARSON, Glen: *Knight Rider*, 1982, (Fernsehserie)
- [59] LATOUR, Bruno: *Where are the missing masses? The sociology of a few mundane objects*. In: BIJKER, Wiebe; LAW, John (Hg.), *Shaping Technology Building Society: Studies in Sociotechnical Change*, Cambridge, Massachusetts: MIT Press, 1992, S. 225–258
- [60] — *Die Hoffnung der Pandora : Untersuchungen zur Wirklichkeit der Wissenschaft*. Frankfurt am Main: Suhrkamp, 2000
- [61] — *Gebt mir ein Laboratorium und ich werde die Welt aus den Angeln heben*. In: BELLIGER, Andréa; KRIEGER, David J. (Hg.), *ANThology: Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*, Bielefeld: Transcript, 1. Auflage, 2006, S. 103–134
- [62] — *Über technische Vermittlung. Philosophie, Soziologie und Genealogie*. In: BELLIGER, Andréa; KRIEGER, David J. (Hg.), *ANThology: Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*, Bielefeld: Transcript, 2006, S. 483–528
- [63] — *Eine neue Soziologie für eine neue Gesellschaft: Einführung in die Akteur-Netzwerk-Theorie*. Frankfurt am Main: Suhrkamp, 2007
- [64] — *Wir sind nie modern gewesen: Versuch einer symmetrischen Anthropologie*. Frankfurt am Main: Suhrkamp, 2008
- [65] LAUREL, Brenda: *Interface Agents. Metaphors with Character*. In: LAUREL, Brenda (Hg.), *The art of human-computer interface design*, New York: Addison-Wesley, 15. Auflage, 2007, S. 355–365
- [66] LAW, John: *After ANT. complexity, naming and topology*. In: LAW, John; HASSARD, John (Hg.), *Actor network theory and after*, Oxford: Blackwell/Sociological Review, 1999, S. 1–14
- [67] — *Monster, Maschinen und soziotechnische Beziehungen*. In: BELLIGER, Andréa; KRIEGER, David J. (Hg.), *ANThology: Ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*, Bielefeld: Transcript, 2006, S. 343–367
- [68] LÜBKE, Valeska: *CyberGender: Geschlecht und Körper im Internet*. Königstein/Taunus: Helmer, 2005

- [69] MACKENZIE, Donald; WAJCMAN, Judy: *Introductory Essay: the social shaping of technology*. In: MACKENZIE, Donald A.; WAJCMAN, Judy (Hg.), *The social shaping of technology*, Buckingham [u.a.]: Open Univ. Press, 2. Auflage, 1999, S. 3–27
- [70] MAES, Pattie: *Agents that Reduce Work and Information Overload*. In: BRADSHAW, Jeffrey M. (Hg.), *Software agents*, Menlo Park: AAAI Press [u.a.], 2. Auflage, 1997, S. 145–164
- [71] MARKOFF, John: *A Software Secretary That Takes Charge*. <http://www.nytimes.com/2008/12/14/business/14stream.html>, 2008, (Letzter Zugriff: 21.6.2011)
- [72] MARKUSSEN, Randi: *Constructing Easiness: Historical Perspectives on Work, Computerization, and Women*. In: STAR, Susan Leigh (Hg.), *The cultures of computing*, Band 1, Oxford, Cambridge: Blackwell/Sociological Review, 1995, S. 158–180
- [73] MEFFERT, Beate; HOCHMUTH, Olaf: *Werkzeuge der Signalverarbeitung. Grundlagen, Anwendungsbeispiele, Übungsaufgaben*. München [u.a.]: Pearson Studium, 2004
- [74] MURET, Eduard; SANDERS, Daniel: *Englisch-Deutsch (A-K): Große Ausgabe*. Berlin [u.a.]: Langenscheidt, 1992
- [75] NAONE, Erica: *Intelligent Software Assistant*. In: *Technology Review*, Band 112(2), 2009, S. 48–49
- [76] — *An Intelligent Software Assistant Debuts*. <http://www.technologyreview.com/blog/>, 2010, (Letzter Zugriff: 16.6.11)
- [77] NEGROPONTE, Nicholas: *Agents, From Direct Manipulation to Delegation*. In: BRADSHAW, Jeffrey M. (Hg.), *Software agents*, Menlo Park: AAAI Press [u.a.], 2. Auflage, 1997, S. 57–66
- [78] NUANCE COMMUNICATIONS, Inc.: *SKRecognition class reference*. <http://dragonmobile.nuancemobiledeveloper.com>, 2010, (Letzter Zugriff: 21.6.2011)
- [79] OUDSHOORN, Nelly; ROMMES, Els; STIENSTRA, Marcelle: *Configuring the User as Everybody: Gender and Design Cultures in Information and Communication Technologies*. In: *Science, Technology, & Human Values*, Band 29(1), 2004, S. 30–63
- [80] OWSNICKI-KLEWE, B.; LUCK, K.; NEBEL, B.: *Wissensrepräsentation und Logik. Eine Einführung*. In: GÖRZ, Günther (Hg.), *Handbuch der künstlichen Intelligenz*, München [u.a.]: Oldenbourg, 3. Auflage, 2000, S. 153–197

- [81] PAYR, Sabine: *Introduction*. In: PAYR, Sabine; TRAPPL, Robert (Hg.), *Agent Culture. Human-Agent Interaction in a Multicultural World*, Mahwah: Lawrence Erlbaum Assoc Inc, 2004, S. IX–XXVI
- [82] — *Ritual or Routine: Communication in Long-Term Relationships with Companions*. In: HEYLEN, D.; KRENN, B.; PAYR, S. (Hg.), *Companions, Virtual Butlers, Assistive Robots: Empirical and Theoretical Insights for Building Long-Term Social Relationships*., Nummer OFAI-TR-2010-05 in OFAI technical report, Wien: Austrian Research Institute for Artificial Intelligence of the Austrian Society for Cybernetic Studies, 2010, S. 1–6
- [83] PFLÜGER, Jörg: *Interaktion im Kontext*. In: HELLIGE, Hans Dieter (Hg.), *Mensch-Computer-Interface: zur Geschichte und Zukunft der Computerbedienung*, Kultur- und Medientheorie, Bielefeld: transcript, 2008, S. 323–383
- [84] PUN, Ngai: *Made in China: women factory workers in a global workplace*. Durham [u.a.]: Duke Univ. Press/Hong Kong Univ. Press, 2005
- [85] RAMMERT, Werner; SCHULZ-SCHAEFFER, Ingo: *Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt*. In: RAMMERT, Werner; SCHULZ-SCHAEFFER, Ingo (Hg.), *Können Maschinen handeln?: soziologische Beiträge zum Verhältnis von Mensch und Technik*, Frankfurt [u.a.]: Campus, 2002, S. 11–64
- [86] SCHIEBINGER, Londa: *Skeletons in the Closet: The First Illustrations of the Female Skeleton in Eighteenth-Century Anatomy*. In: GALLAGHER, Catherine; LAQUEUR, Thomas Walter (Hg.), *The Making of the modern body: sexuality and society in the nineteenth century*, Berkeley: University of California Press, 1987, S. 42–82
- [87] SCHMITZ, Sigrid: *Geschlechtergrenzen. Geschlechtsentwicklung, Intersex und Transsex im Spannungsfeld zwischen biologischer Determination und kultureller Konstruktion*. In: EBELING, Smilla; SCHMITZ, Sigrid (Hg.), *Geschlechterforschung und Naturwissenschaften: Einführung in ein komplexes Wechselspiel*, Wiesbaden: VS Verlag für Sozialwissenschaften, 2006, S. 33–56
- [88] SCHWARZ COWAN, Ruth: *The Industrial Revolution in the Home*. In: MACKENZIE, Donald A.; WAJCMAN, Judy (Hg.), *The social shaping of technology*, Buckingham [u.a.]: Open Univ. Press, 2. Auflage, 1999, S. 280–298
- [89] SCULLEY, John; BYRNE, John A.: *Meine Karriere bei PepsiCo und Apple*. Düsseldorf: Econ-Verlag, 1987

- [90] SENEGERS, Phoebe: *The Agents of McDonaldization*. In: PAYR, Sabine; TRAPPL, Robert (Hg.), *Agent Culture. Human-Agent Interaction in a Multicultural World*, Mahwah: Lawrence Erlbaum Assoc Inc, 2004, S. IX–XXVI
- [91] SIRI, Inc.: *About Siri :: Product :: Siri - Your Virtual Personal Assistant*. <http://siri.com/about/>, 2011, (Letzter Zugriff: 12.6.11)
- [92] SPOOL, Jared: *Knowledge Navigator Deconstructed: Building an Envisionment*. [http://www.uie.com/articles/knowledge\\_navigator/](http://www.uie.com/articles/knowledge_navigator/), 2007, (Letzter Zugriff: 21.6.2011)
- [93] STAR, Susan Leigh: *Power, technologies and the phenomenology of standards: On being allergic to onions*. In: LAW, John (Hg.), *A Sociology of Monsters. Essays on Power, Technology and Domination*, London: Routledge, 1991, S. 26–56
- [94] — *Infrastructure and ethnographic practice*. In: *Scandinavian Journal of Information Systems*, Band 14(2), 2002, S. 107–122
- [95] SUCHMAN, Lucy: *Figuring 'service' in discourses of ICT. The case of software agents*. In: WEBER, Jutta; BATH, Corinna (Hg.), *Turbulente Körper, soziale Maschinen: feministische Studien zur Technowissenschaftskultur, Studien interdisziplinäre Geschlechterforschung*, Band 7, Opladen: Leske + Budrich, 2003, S. 65–74
- [96] — *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge, New York, Melbourne: Cambridge University Press, 2. Auflage, 2007
- [97] — *Feminist STS and the Sciences of the Artificial*. In: HACKETT, Edward J. (Hg.), *The handbook of science and technology studies*, Cambridge, Mass. [u.a.]: MIT Press, 3. Auflage, 2008, S. 139–163
- [98] WAJCMAN, Judy: *TechnoFeminism*. Cambridge, UK [u.a.]: Polity Press, 1. Auflage, 2004
- [99] WEBER, Jutta; BATH, Corinna: *'Social' Robots & 'Emotional' Software Agents. Gendering Processes and De-Gendering Strategies for 'Technologies in the Making'*. In: ZORN, Isabel (Hg.), *Gender designs IT : construction and deconstruction of information society technology*, Wiesbaden: VS Verlag für Sozialwissenschaften, 1. Auflage, 2007
- [100] WENDEMUTH, Andreas; ET AL.: *Grundlagen der stochastischen Sprachverarbeitung*. München [u.a.]: Oldenbourg, 2004

- [101] WIESNER, Heike: *Die Inszenierung der Geschlechter in den Naturwissenschaften: Wissenschafts- und Genderforschung im Dialog*. Frankfurt am Main: Campus Verlag, 2002
- [102] WIKIMEDIA: *Sigrid - Wikipedia, the free encyclopedia*. <http://en.wikipedia.org/w/index.php?title=Sigrid>, 2011, (Letzter Zugriff: 12.6.11)
- [103] WINKER, Gabriele; DEGELE, Nina: *Intersektionalität: Zur Analyse sozialer Ungleichheiten*. Bielefeld: transcript Verlag, 2009
- [104] WISE, J. Macgregor: *Intelligent Agency*. In: *Cultural Studies*, Band 12(3), 1998, S. 410–428
- [105] WOOLDRIDGE, Michael J.: *An introduction to multiagent systems*. Chichester: Wiley, 2002
- [106] WOOLDRIDGE, Michael J.; JENNINGS, Nick: *Applications of Intelligent Agents*. In: JENNINGS, Nick; WOOLDRIDGE, Michael J. (Hg.), *Agent technology: foundations, applications, and markets*, Berlin [u.a.]: Springer, 1998, S. 3–28
- [107] XIAO, Jun; CATRAMBONE, Richard; STASKO, John: *Be quiet? evaluating proactive and reactive user interface assistants*. In: *Human-computer interaction: INTERACT'03; IFIP TC13 International Conference on Human-Computer Interaction, 1st-5th September 2003, Zürich*, 2003, S. 383
- [108] ZIEGLER, Peter-Michael: *Foxconn will bis zu 400.000 neue Arbeiter einstellen*. <http://www.heise.de/jobs/meldung/Foxconn-will-bis-zu-400-000-neue-Arbeiter-einstellen-1061742.html>, 2010, (letzter Zugriff: 21.6.2011)