

# "Die, Clippy, die!!"

## Zur Evolution von User Interface Agents und aktuelle Entwicklungen bei Intelligent Personal Assistants

Ulrike Exner

Fachhochschule Brandenburg  
exner@fh-brandenburg.de

Medieninformatik M. Sc.  
Modul: Smart Graphics

13. Januar 2016

### **Inhalt**

1. Einführung.....	2
2. Einordnung.....	2
3. Interface agents.....	5
3. 1 User Interfaces smart gestalten - aber wie? .....	5
3. 2 Interface agents in der Praxis.....	6
3. 3 Textorientierte Interface agents .....	7
3. 4 Textorientierte Interface agents im Web.....	9
3. 5 Aktuelle sprachbasierte personal assistants .....	10
4. Literaturverzeichnis.....	12
5. Abbildungsverzeichnis.....	14

# 1. Einführung

Künstliche Intelligenz (AI) fasziniert die Menschen schon lange. Aus der Popkultur ist sie nicht wegzudenken: In zahlreichen futuristischen Filmen und Büchern helfen AI-Agenten mit mal mehr, mal weniger guten Absichten ihren Erfindern bei der Lösung von Problemen oder im alltäglichen Leben. Ob es die eigenmächtigen Raumschiffcomputer "HAL 9000" in Kubricks "2001: A Space Odyssey" und "Auto" im Disney-Film WALL-E sind oder das verführerische Betriebssystem "Samantha" im Film "Her": Die Sehnsucht nach verlässlichen intelligenten Agenten mit scheinbar eigener Persönlichkeit scheint groß zu sein.

In der Realität lassen wir uns heute durch Apples "Siri" oder Microsofts "Cortana" immerhin schon den Weg zeigen oder einen Restauranttisch buchen, eine Anbindung an Fantasien aus Science-Fiction-Filmen scheint aber noch in weiter Zukunft zu liegen. Doch die Evolution von intelligenten Agenten ist eine Geschichte mit Höhen und Tiefen. Ein berühmter Vertreter einer einfachen Form von *interface agents* ist Microsofts "Clippy the Paperclip" - eine tragische Figur auf dem Experimentierfeld der Künstlichen Intelligenz, der schon vor Beginn ihres Untergangs mit Häme und Spott begegnet wurde. Neue Entwicklungen versprechen innovative *agents*, die sich das Internet zu Nutze machen. Die jeweiligen Entwicklungslinien sollen in dieser Ausarbeitung aufgezeigt und diskutiert werden.

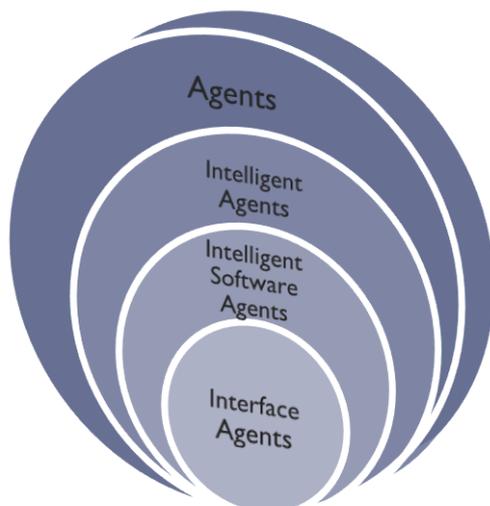


Abbildung 1: Agents - Konzeptionelle Betrachtung, eigene Darstellung.

im Bereich virtuelle Realität als computergenerierte, interaktive virtuelle Umgebung und medienpsychologische Ansätze zu menschlicher Interaktion. Daher folgt hier eine kurze Begriffseinordnung.

## 2. Einordnung

Unglücklicherweise existieren in der Forschungsliteratur eine Vielzahl an Begriffen um den Themenkomplex *user agents*. Zahlreiche verwandte und auch synonyme Bezeichnungen fanden ihren Weg in die Literatur, nicht zuletzt aufgrund der Interdisziplinarität des Forschungsthemas zwischen Informatik, Informationswissenschaft, Psychologie, Soziologie und Human-Computer-Interaction. Eng verknüpft mit Untersuchungen zum vorliegenden Thema sind Forschungsthemen

## I. agents

Das Wort *agent*<sup>1</sup> (deutsch: *Agent*) leitet sich laut Duden vom lateinischen *agere* (tun, handeln, machen) ab und bezeichnet daher etwas, das agiert. In der Künstlichen Intelligenz wird der Begriff erweitert und folgendermaßen beschrieben:

*"Ein Agent besitzt Sensoren, mit denen er seine Umgebung wahrnehmen kann, und Aktuatoren, durch die er handelt."*<sup>2</sup>

Agents können also von unterschiedlicher Gestalt sein und sich in ihrer Komplexität stark voneinander unterscheiden. Man unterscheidet zwischen *menschlichen Agenten*, *Hardware- bzw. Roboter-Agenten* und *Softwareagenten*. Ein menschlicher Agent (z. B. Reiseverkehrskauffrau) hat Augen, Ohren und andere Organe als Sensoren sowie Hände, Füße, Mund und andere Körperteile als Aktuatoren. Ein Roboter-Agent könnte Kameras und Infrarotentfernungsmesser als Sensoren verwenden und verschiedene Motoren als Aktuatoren.<sup>3</sup> In vorliegender Arbeit sollen jedoch nur *software agents* von Bedeutung sein: Wie unterscheiden sich *software agents* von *software*?

## II. software agents

In den meisten Fällen geben wir uns damit zufrieden, wenn wir Computerprogramme bzw. *software* als folgsame, fantasielose Gefährten betrachten, die unsere Anweisungen wortwörtlich nehmen. Sehr viele moderne Anwendungen setzen aber einige höhere Prozeduren voraus und verlangen von der Software, dass sie selbst entscheidet, was sie tun muss, um eine bestimmte Aufgabe zu erfüllen.<sup>4</sup> Ein wichtiger Begriff hierbei ist die Delegation nach Kay<sup>5</sup>: Hierbei informiert der Nutzer den Computer über das Ziel einer Aufgabe und der Computer führt selbstständig die notwendigen Arbeitsschritte durch. Computerprogramme, die dazu in der Lage sind, werden als *agents* bzw. *software agents* bezeichnet:

*"[...] agents are simply computer systems that are capable of autonomous action in some environment in order to meet their design objectives."*<sup>6</sup>

## III. intelligent software agents

Was *intelligent software agents* (häufig auch verkürzt als *intelligent agents* bezeichnet) von einfachen Computerprogrammen und *software agents* unterscheidet, ist vor allem ihre Fähigkeit, nicht nur autonom zu operieren, d.h. ohne direkte Anweisung per Tastatur- oder Maustaste tätig zu werden. Darüber hinaus führen sie ihre Pflichten auf intelligente Weise aus, also kontext- und nutzerabhängig.<sup>7</sup> Zudem sind sie in der Lage, ihre Umgebung

---

<sup>1</sup> Die meisten Fachbegriffe werden in dieser Ausarbeitung nicht ins Deutsche übersetzt, um die Anwendbarkeit bzw. Nachvollziehbarkeit auf die zumeist englische Forschungsliteratur in diesem Bereich zu gewährleisten.

<sup>2</sup> Vgl. Russell und Norvig 2012, S. 60.

<sup>3</sup> Russell und Norvig 2012, S. 60.

<sup>4</sup> Vgl. Wooldridge 2002, S. 3.

<sup>5</sup> Vgl. Kay 1984.

<sup>6</sup> Wooldridge 2002, S. 8.

<sup>7</sup> Vgl. Dehn und van Mulken 2000, S. 1f.

wahrzunehmen, über einen längeren Zeitraum beständig zu sein, sich an Änderungen anzupassen sowie Ziele zu erzeugen und zu verfolgen<sup>8</sup>.

Russell und Norvig 2012, Wooldridge 2002 u.a. verwenden hierfür die Termini *autonomous agents*, *smart agents*, *intelligent agents*, *rational agents* und *bots* weitestgehend synonym - alle Begriffe stehen dabei den Konzepten der Künstlichen Intelligenz sehr nahe.

Die Bandbreite unterschiedlicher Beispiele und Typen im Bereich *intelligent software agents* ist sehr groß. Es folgen nur einige Beispiele, nach denen *intelligent agents* typisiert werden können und die in der Literatur recht häufig genannt werden.

Eine häufig zitierte Klassifikation nach Nass et al. 1995 unterscheidet zwischen *Informationsagenten*, *Kooperationsagenten* und *Transaktionsagenten*<sup>9</sup>. **Informationsagenten** unterstützen bei der Suche nach Informationen in einem Netzwerk. Sie müssen in der Lage sein, Informationsquellen zu finden, die Information zu extrahieren und die Ergebnisse anschaulich aufzubereiten. Einen ähnlichen Ansatz verfolgen **Data-mining agents**. Sie helfen dem Nutzer, sich in einer großen Menge an Daten zurechtzufinden und aufgrund einer Suchanfrage relevante Informationen zu "schürfen". Dabei suchen sie nicht nur in einzelnen Datenbanken, sondern beschaffen Daten nach dem Datawarehouse-Prinzip aus mehreren Quellen. **Kooperationsagenten** lösen Problemstellungen durch Kommunikation und Kooperation mit anderen Agenten. Man spricht dann auch von *Multi-Agentensystemen*. **Transaktionsagenten** lösen Aufgaben in transaktionsorientierten Systemen, bei denen ein großes Maß an Sicherheit, Robustheit und Datenschutz notwendig ist.

**User agents** (oder *personal agents*, *web agents*) sind häufig mit dem World Wide Web verknüpft und übernehmen verschiedenste Aufgaben für ihren Auftraggeber, dazu gehört das Prüfen von neuankommenden E-Mails und deren Sortierung nach Nutzervorgabe oder auch das Bereitstellen von personalisierten Newsstreams. Im engeren Sinne bezeichnet *user agent* meist ein Client-Programm, mit dem ein Netzwerkdienst genutzt werden kann, also z. B. Webbrowser, E-Mail-Programme, RSS-Reader oder IRC-Clients.

### III. interface agents & personal assistants

*Interface agents* schließlich sind eine weitere Spezialform von *intelligent software agents* und dienen dazu, eine grafische Benutzeroberfläche (*graphical user interface, GUI*) in der Anwendung einfacher und unterhaltsamer zu gestalten. Das Besondere an ihnen ist, dass sie damit Teil der Benutzeroberfläche selbst werden. In der Forschungsliteratur existieren zahlreiche weitere Bezeichnungen für diesen Typus, u.a. *social*, *embodied*, *conversational*, *lifelike*, *animated*, and *personified agents*.<sup>10</sup> Zudem ist auch mit dem Begriff *intelligent user interface* zumeist eine einfache Form von *interface agents* gemeint. Einen wichtigen Hinweis auf die Historie geben Case und Twyman 2015, die in diesem Bereich stark unterscheiden nach *interface agents*, die

---

<sup>8</sup> Vgl. Russell und Norvig 2012, S. 25.

<sup>9</sup> Vgl. Nass et al. 1995, S. 21.

<sup>10</sup> Vgl. Swartz und Nardi, S. 2.

einen Menschen abbilden oder diesem nachempfunden sind (*embodied*) oder eben nicht (*disembodied*).<sup>11</sup>

Mit der Weiterentwicklung der Anwendungen in diesem Bereich haben sich auch die verwendeten Fachbegriffe verändert. Seit der Entwicklung web- und besonders sprachbasierter interface agents hat sich der Terminus *online assistant* bzw. *personal assistant* durchgesetzt. Aber auch zu diesem Begriff existieren eine Vielzahl nah verwandter bzw. synonyme Bezeichnungen, allen voran die *intelligent virtual assistants*<sup>12</sup>, aber auch *virtual butlers* oder *intelligent personal agents* sowie mit Fokus auf Agenten auf mobilen Endgeräten die Bezeichnungen *mobile companions* oder *mobile virtual assistant*<sup>13</sup>.

## 3. Interface agents

### 3.1 User Interfaces smart gestalten - aber wie?

Früher wie heute ist einer der wichtigsten Aspekte im Bereich Human-Computer-Interaction<sup>14</sup> das Design von *User Interfaces*.<sup>15</sup> Doch wie wird *User Interface* aktuell definiert? Unter einer Benutzeroberfläche oder Benutzerschnittstelle (*User Interface*) versteht man die Art und Weise, wie Befehle und Daten in den Computer eingegeben werden. Die Benutzeroberfläche ist somit die Schnittstelle zwischen Computer und Mensch. Sie wird von dem Betriebssystem oder den Anwendungsprogrammen bereitgestellt und soll dem Benutzer das Arbeiten mit dem Computer erleichtern.<sup>16</sup>

In Vergangenheit haben sich verschiedene Typen von User Interfaces herausgebildet, die z. T. schon angesprochen worden sind. Eine klare Abgrenzung ist dabei in der Praxis häufig nicht möglich. Einige wichtige Typen werden im Folgenden kurz zusammengefasst.<sup>17</sup>

1. *Command line interfaces* sind eine klassische Möglichkeit, mit dem Computer über eine Kommandozeile zu interagieren. Eingaben werden über die Tastatur in Form von Zeichenketten gegeben.
2. *Graphical user interfaces (GUI)* akzeptieren Eingaben über Tastatur und Maus und ermöglichen eine graphische Ausgabe über den Computermonitor.
3. *Web-based user interfaces / web user interfaces (WUI)* als Sonderform der grafischen user interfaces verarbeiten Ein- und Ausgaben über Websites, die dem Nutzer über

---

<sup>11</sup> Vgl. Case und Twyman 2015, S. 1

<sup>12</sup> Seit dem Jahr 2000 findet die International Conference on Intelligent Virtual Agents (IVA) jährlich statt. Zuletzt gastierte sie 2015 in Delft, Niederlande: <http://iva2015.tudelft.nl/>.

<sup>13</sup> Vgl. Gilbert et al. 2011.

<sup>14</sup> "Human-computer interaction (HCI) is a cross-disciplinary area (e.g., engineering, psychology, ergonomics, design) that deals with the theory, design, implementation, and evaluation of the ways that humans use and interact with computing devices. Interaction is a concept to be distinguished from another similar term, interface. Roughly speaking, interaction refers to an abstract model by which humans interact with the computing device for a given task, and an interface is a choice of technical realization (hardware or software) of such a given interaction model. Thus, the letter I in HCI refers to both interaction and interface, encompassing the abstract model and the technological methodology" Kim 2015, S. 1

<sup>15</sup> Vgl. Anton Nijholt 2004, S. 465.

<sup>16</sup> Vgl. IT Wissen 2015.

<sup>17</sup> Vgl. El-Bakry et al. 2010, S. 191.

einen Browser angezeigt werden. Bekannte Technologien zur Realisierung sind Java, AJAX, .NET, usw.

4. *Touch user interfaces* sind eine Sonderform von grafischen user interfaces, die einen Touchscreen-Bildschirm als kombinierte Ein- und Ausgabeschnittstelle verwenden.
5. *Voice user interfaces* als Sprachdialogsysteme nehmen die natürliche Sprache des Menschen als Eingabe entgegen und geben ebenfalls eine akustische o. a. Ausgabe zurück.

Wie können *User Interfaces* nun "smart" designet werden, um Menschen Hilfestellung und Ratschlag anzubieten, wenn sie diese benötigen? Es zeichnet sich ein Trend ab hin zu *Social Interfaces* ab, die Mensch-zu-Mensch-Kommunikationseigenschaften betonen und den Computer zu einem *Social Actor* werden lassen. Soziale Charakteristika menschlicher Kommunikation werden dabei imitiert. Beim Umgang mit dem Computer wird nur allzu häufig deutlich, wie vor allem Alltagsnutzer in menschliche Verhaltensmuster geraten und von der Maschine vor Ihnen gleichsam menschliche Reaktionen erwarten. Interaktionsdesign basiert demnach immer auf den Gegebenheiten sozialer Interaktion und dem, was Menschen emotional stimuliert.<sup>18</sup>

Ein wichtiger Aspekt dabei ist der *Anthropomorphismus*, also lt. Duden die Übertragung menschlicher Eigenschaften auf Nichtmenschliches.<sup>19</sup> Wie stark *menschlich* User Interfaces tatsächlich ausgeprägt waren und sind, variiert und bildete auch gewisse Trends der Vergangenheit ab. In der Praxis gibt es Beispiele, in denen die menschliche Sprache akustisch verwendet wird, der Agent als vermenschlichte Figur auftritt, also auch visuell in Erscheinung tritt oder beide Erscheinungen in Kombination umgesetzt wurden.

### 3. 2 Interface agents in der Praxis

Das Ziel, menschliche, personalisierte und unterhaltsame Benutzerschnittstellen zu schaffen, hat gegen Ende der Neunziger und zu Beginn der 2000er Jahre zum Entstehen vermenschlichter *embodied conversational agents* geführt. Intelligente Agenten, die die menschliche Gestik, Mimik und Sprache imitieren, variieren dabei von einfachen Cartoon-ähnlichen 2D-Darstellungen, bis hin zu virtuellen Personen in 3D-Welten und intelligenten Androiden in der realen Welt.<sup>20</sup> Das Ziel eines *intelligent user interfaces* ist es also, dem Computer ein "Gesicht" zu geben, mit dem der Nutzer sprechen kann.

Die ersten *intelligent user interfaces* verwendeten dabei natürliche dialogorientierte Kommunikationsmuster, wofür das UI Design anthropomorphe Repräsentanten im 2D- oder 3D-Gewand erfunden hat, mit denen über eine textbasiertes Interaktionsmodul kommuniziert wurde,<sup>21</sup> die z. T. auch als *User Interface Avatars* bezeichnet wurden. Ein

---

<sup>18</sup> Vgl. Anton Nijholt 2004, S. 468.

<sup>19</sup> Vgl. Both 2012, S. 124.

<sup>20</sup> Vgl. André et al. 2005, S. 1.

<sup>21</sup> Vgl. Berner 2003.

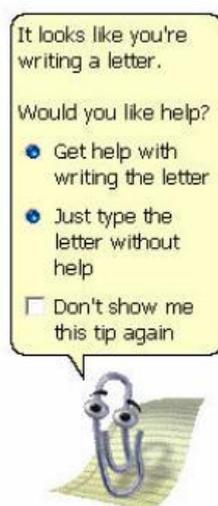
bekanntes Beispiel hierfür sind die *user interface agents* des Microsoft Office-Pakets, wie *Clippy, the Paperclip*.

Mit dem Siegeszug des World Wide Web wurden diese 2D- oder 3D-Avatare für Websites aufbereitet, wo sie Nutzern beispielsweise auf Bibliothekswebsites assistieren und dafür auf externe Ressourcen und Services des Internets zurückgreifen.

Einen entscheidenden Sprung bei der Weiterentwicklung von *interface agents* betraf die Verlagerung von *textorientierten*<sup>22</sup> zu *sprachbasierten* Schnittstellen. Die ersten erfolgreichen *personal assistants*, die auf Spracherkennung basierten, wurden mit Siri ab 2011 von Apple für das iPhone vermarktet.

### 3. 3 Textorientierte Interface agents

#### I. Clippy, the Paperclip



**Abbildung 2: Clippy versucht zu helfen,**  
[http://www.slate.com/content/dam/slate/articles/technology/technology/2012/04/120410\\_TECH\\_clippyMicrosoftWord.jpg.CROP.article250-medium.jpg](http://www.slate.com/content/dam/slate/articles/technology/technology/2012/04/120410_TECH_clippyMicrosoftWord.jpg.CROP.article250-medium.jpg)

Ende der 90er Jahre waren es vor allem die Fortschritte im Bereich der Computergrafik, die die Realisation von videobasierten, cartoonartigen oder 3D-modellierten animierten *Interface Agents* ermöglichte. Es sollte eine möglichst menschliche Verhaltensweise simuliert werden, indem Sprache, Emotionen, Gestik sowie Augen-, Kopf- und Körperbewegungen implementiert wurden. Anwendern sollte es dadurch ermöglicht werden, bekannte menschliche Interaktionsfähigkeiten anzuwenden, um die Kommunikation mit dem Computer reibungsloser zu gestalten.<sup>23</sup> Viele dieser *interface agents* wurden daraufhin menschlich gestaltet, jedoch ist die Liste tierischer oder dinglicher *interface agents*-Beispiele lang.

Einer der bekanntesten *user interface agents* ist *Clippy, the Paperclip* von Microsoft (in Microsofts Office-Programmen selbst als "Clippit", dt.: "Karl Klammer" bezeichnet), der ab 1997 in jeder Office Software Suite enthalten war. Schnell wurde Clippy zum Hassobjekt vieler Anwender und auch vieler IT-Journalisten. Nachdem Microsoft es ab 2000 ermöglichte, den Assistenten dauerhaft zu entfernen, erklärten zahlreiche (Online-) Artikel, wie man eben dies dauerhaft umsetzen konnte (z. B. PC Welt 2005: "Karl Klammer & Co. Entsorgen"<sup>24</sup>, ZDNet 1998: "Die Clippy, Die"<sup>25</sup>). Nach Clippys offiziellem Ende 2007 ließen

<sup>22</sup> Um eine Verwechslung mit den von den grafischen Benutzeroberflächen (*graphical user interfaces* bzw. GUIs) zu unterscheidenden zeichenorientierten Benutzeroberflächen (*text-based user interfaces*) oder gar Kommandozeilen (*command-line interfaces*) zu vermeiden, wird in dieser Arbeit der Begriff *textorientierte interface agents* verwendet.

<sup>23</sup> Vgl. Dehn und van Mulken 2000, S. 1f.

<sup>24</sup> Vgl. Eggeling 2005.

<sup>25</sup> Vgl. Noteboom 1998.

ihn genervte Nutzer viele weitere Tode sterben. Es gibt Hunderte Cartoons, Videos und Texte, in denen er ein grausames Ende findet.<sup>26</sup>

## II. Diskussion

Forlizzi et al. konstatiert, dass die HCI Community gespalten bleibt in Bezug auf *interface agents* wie Clippy: Viele Studien, die sich auf die Effizienz und Effektivität von solchen Interfaces beziehen, konnten nicht nachweisen, dass sie diese signifikant steigern würden.<sup>27</sup>

Mitunter wurden von Seiten der Befürworter große Hoffnungen in die Erschaffung von *interface agents* gesetzt, insbesondere versprach man sich ein positives, soziales Empfinden bei der Interaktion mit einem interface agent dieser Art und pädagogische Mehrwerte und positive Effekte auf das motivierte Lernen von Schülern und Studenten durch sogenannte *animated pedagogical agents*.<sup>28</sup>

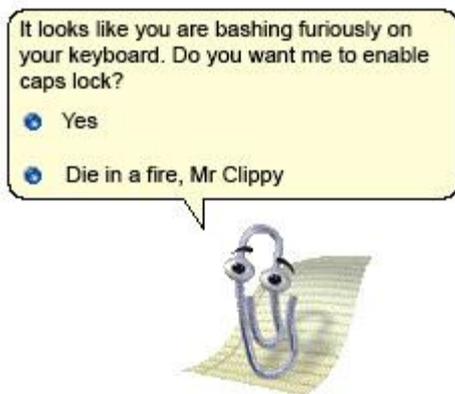


Abbildung 3: Clippy: Eine Beispiel für Satire, <http://www.extremetech.com/wp-content/uploads/2012/01/mr-clippy-angry.jpg>

Im Umfeld der Befürworter fanden eine ganze Reihe von Studien statt, die interface agents und deren Ausprägungen detailliert auf ihre Effektivität untersuchten. So fanden einige Untersuchungen heraus, dass Nutzer ihre Agenten intelligenter und vertrauenswürdiger einstufen, je realistischer sie aussahen (im Gegensatz zu einfachen 2D-Modellierungen). An dieser Stelle sei die Studie von King und Ohya 1996 exemplarisch genannt. Auch über die Ausprägung nach Geschlecht des interface agents wurde geforscht und darüber, ob ein menschliche oder besser tierische interface agents zur wählen seien (Catrambone et al. 2002).

Die Liste der Kritik und Erklärungsversuche über das Scheitern von Clippy u.a. ist jedoch ebenso lang: (Dehn und van Mulken 2000) nennen Aspekte, die die Gegner von interface agents sammelten, um darzulegen, dass die Implementierung dieser eine schlechte Idee war: Einige Verhaltensweisen könnten aufgrund des abweichenden Mentalen Modells durch den Anwender missinterpretiert werden. Besonders Bewegungen, die ein interface agent in Leerlaufzeiten vornimmt (z. B. mit dem Fuß wippen), könnten bei unerfahrenen Nutzern zum Schluss führen, dass System arbeite oder berechne gerade etwas<sup>29</sup>. Ein weiteres Argument kritisiert, dass ein interface agent nur ein weiteres Objekt sei, das die Aufmerksamkeit des Nutzers auf sich zieht und somit die kognitiven Ressourcen von Nutzern beansprucht.<sup>30</sup> Andere Forscher sind sich zudem einig darüber, dass Nutzer ihren Computer ohnehin als sozialen Agenten verstehen und in dieser

<sup>26</sup> Vgl. Horchert 2015.

<sup>27</sup> Vgl. Forlizzi et al., S. 210f.

<sup>28</sup> Siehe hierzu Clark Elliott et al. 1997 und Lester et al. 1997.

<sup>29</sup> Vgl. Wilson 1997.

<sup>30</sup> Vgl. Walker et al. 1994.

Weise mit ihm interagieren - ganz ohne, dass Interface Designer ihm künstlich ein menschlicheres Antlitz gegeben haben. Diesen Ansatz bezeichnet man auch als Computers As Social Actors (CASA)-Theorie.<sup>31</sup>

Einer der strengsten Kritiker der ersten Interface Agents war Ben Shneiderman, der den zugrundeliegenden Anthropomorphismus als "humpty dumpty syndrome"<sup>32</sup> bezeichnete und UI Designer schon früh davor warnte, dem "primitiven Drang" nachzugeben, *intelligent agents* menschlich zu gestalten:

*"The anthropomorphic styles are cute the first time, silly the second time, and an annoying distraction the third time."*<sup>33</sup>

Er konstatiert, dass sich Nutzer vom künstlichen „Zeigefinger“ eines Tutors bevormundet und demotiviert fühlen könnten und diese Agenten ohnehin wenig authentisch wirken und sogar zu Fantasieverlust führen könnten.

### 3. 4 Textorientierte Interface agents im Web

#### I. Bibliotheksassistent (online)

Die Gestaltung von web-basierten user interfaces ist heute ein komplexes und dynamisches Feld. Anfang der 2000er kamen textorientierte interface agents auf, die sich einiger Zeit großer Beliebtheit erfreuten, heute jedoch ähnlich wie offline-Vorgänger wie Clippy, the Paperclip, meist nicht mehr verwendet werden. In einigen Quellen werden interface agents dieser Art auch als *automated online assistant*, *text-based conversational agents*, *artificial conversation entities*, *chatterboxes*, *chatbots* oder einfach als *bots* bezeichnet.<sup>34</sup>

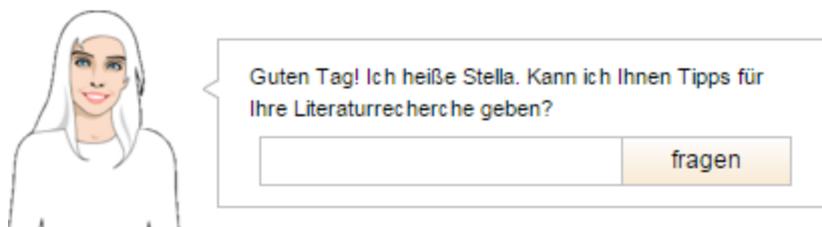


Abbildung 4: Stella, der Chatbot der Staats- und Universitätsbibliothek Hamburg, Screenshot.

Zumeist basierten Chatbots dieser Zeit auf der Artificial Intelligence Markup Language (AIML), ein XML-Dialekt. Zu erwartende Nutzereingaben müssen damit aufwändig in Taxonomien überführt werden.

Websites, die in der Vergangenheit interface agents nutzten, wie z. B. buy.com, extempo.com, ananova.com, mysimon.com haben längst andere Wege der Benutzerführung gewählt. Auch Bibliotheken im amerikanischen Raum interessierten sich Anfang der 2000er Jahre für diese neuer Art des Auskunftsdienstes. In Deutschland wurde der "ASKademicus" als "Chatbot" als digitaler Auskunftsdienst in der Universitätsbibliothek Dortmund 2004

<sup>31</sup> Siehe dazu Nass et al. 1995 sowie Nass et al. 1994.

<sup>32</sup> Shneiderman 1993, S. 331.

<sup>33</sup> Shneiderman 1995, S. 13ff.

<sup>34</sup> Vgl. Rubin et al. 2010, S. 497.

zuerst angeboten. Im Allgemeinen wurden in Bibliotheken solche interface agents implementiert, um als "niedrigschwelliges Angebot den Einstieg in die Informationsrecherche und -beschaffung zu erleichtern und Aufmerksamkeit für elektronische Medien wie Datenbanken und Online-Zeitschriften zu wecken."<sup>35</sup> Des Weiteren betrachtete man diese als Ergänzung zu der persönlichen Beratung an der Auskunftstheke und versprach sich eine Verbesserung der Nutzung der Website und einen Imagewandel.

## II. Diskussion

Der Bibliotheksbereich fachlicherseits zeigte sich begeistert von Chatbots. Noch 2011 gab es ein Projekt zur Implementierung eines neuen agents in der Universitätsbibliothek Hildesheim<sup>36</sup>, obwohl einige Schwächen solcher Systeme bereits hinlänglich bekannt waren.

Im Allgemeinen wurden bestehende Chatbots gut oder sehr gut evaluiert. Diverse Studien benennen eine Reihe an Vorteilen und sprechen eine eindeutige Empfehlung zur Implementierung von Chatbots aus.<sup>37</sup> Zu den Errungenschaften von Chatbots zählen die *Steigerung der Diversität*, da ein Zugang zu den Bibliotheksservices auch für Menschen (mit einer Behinderung) möglich wird, denen ein physischer Zugang zur Bibliothek erschwert ist, *Imagewandel*, *Freude an der Nutzung der Bibliothek* und das *automatische Beantworten von oft gestellten Fragen* zur Entlastung der Bibliotheksmitarbeiter. Zudem wiederholen die Befürworter Argumente, die bereits Verfechter von Clippy und Co. vorbrachten, z. B. das eine Vermenschlichung der Web-Auskunft sich psychologisch positiv auf Nutzer auswirken würden.

Neben den *organisatorischen und technischen Hürden zur Implementierung* in zumeist öffentlich finanzierten Bibliotheken ist der Hauptkritikpunkt die *Akzeptanz durch die Zielgruppe*. Viele Nutzer "missbrauchen" einen Chatbot und richten Fragen an sie, die nichts mit dem ursprünglichen Sinn und Zweck des als Informationssystem konzipierten agents zu tun haben. Auch Beleidigungen und Versuche, das System in die Irre zu führen, sind nicht selten. Dies sind als Reaktionen zu betrachten auf nicht ausgereifte Systeme, Systeme die nicht ernst genommen oder als Marketing-Gag wahrgenommen werden.<sup>38</sup>

## 3. 5 Aktuelle sprachbasierte personal assistants

### I. Siri, Cortana & Co.

In einer zunehmend mobilen Welt war die Entwicklung sprachbasierter Dienste eine logische Konsequenz aus Mangel an komfortablen Tastaturen und großen Bildschirmen. Populäre Vertreter moderner *virtual assistants* sind Apples Siri, Microsofts Cortana und Googles *Conversational Search*. Diese Assistenten bieten die Möglichkeit, in natürlicher Sprache mit

---

<sup>35</sup> Christensen 2008, S. 19ff.

<sup>36</sup> Vgl. Reichle 2011.

<sup>37</sup> Vgl. Rubin et al. 2010, S. 512f.

<sup>38</sup> Vgl. Rubin et al. 2010, S. 514f.

dem Smartphone zu kommunizieren und aufgrund dessen verschiedene Befehle ausführen zu lassen (z. B. Fragen nach dem Wetter, dem richtigen Weg, Termine eintragen lassen, usw.). Diese Entwicklung war freilich nicht möglich ohne verbesserte Spracherkennungssoftware. Kommerzielle Speech-to-text-Anwendungen befinden sich bereits seit Mitte der 90er Jahre in Entwicklung, die frühen Systeme waren jedoch wenig praxistauglich. Erst im Jahr 2007 stellte Ford SYNC einen nützlichen Spracherkennungsdienst für seine Fahrzeuge vor.



Abbildung 5: Nutzung eines sprachbasierten personal assistants, [http://www.wired.com/images\\_blogs/gadgetlab/2013/03/130318\\_siri\\_0078.jpg](http://www.wired.com/images_blogs/gadgetlab/2013/03/130318_siri_0078.jpg)

Anfrage unter Zuhilfenahme diverser dritter Webservices (z. B. Wolfram Alpha, Wikipedia, Apple Maps, usw.) verarbeitet und beantwortet werden. Siri fokussiert sich also auf die Lösung von Aufgaben anstatt Informationen zu repräsentieren und bezieht den Kontext (Standort, Zeit,..) des Nutzers ein. Herzstück des assistants ist Nuance als Spracherkennungssoftware, welche Sprache in Text (bei der Eingabe) und Text in Sprache (bei Siris Antwort) umwandelt. Siri verfügt darüber hinaus über eine semantische "Autovervollständigung".

Apple's Siri (Speech Interpretation and Recognition Interface) war ab 2011 Bestandteil des iPhone 4S und erfüllt die Funktionen eines heutigen modernen persönlichen Assistenten. Dafür bedient es sich in erster Linie einer sprachlichen Nutzerschnittstelle, um auf die Stimme des Besitzers reagieren zu können. Über eine Client-Server-Verbindung werden die komprimierten Sprachdaten verschlüsselt an einen Apple-Server gesendet. Dort kann die

## II. Diskussion

Aktuelle Entwicklungen wie Siri und Co. zeigen: Anthropomorphismus bleibt ein wichtiges Stichwort. Siri kommt als sympathische Freundin daher, hat die "schlechten Manieren" von Clippy abgelegt und arbeitet nur nach klarer Aufforderung. Im Vergleich zu textbasierten *interface agents* ist die Bedienung durch das sprachbasierte Interface nahezu mühelos geworden. Weiterentwicklungen sind also eher im Bereich der Künstlichen Intelligenz zu erwarten. Der *interface agent* der Zukunft ist ein sich selbst entwickelnder, vorausschauender Assistent, ein "Global Brain", das unbegrenzte Funktionalitäten und Informationsquellen heranzieht, so dass der Umgang mit Informationen noch selbstverständlicher wird. Viv ist beispielsweise ein neuer "superintelligent, self-learning, voice-enabled agent", der momentan im Silicon Valley entwickelt wird.<sup>39</sup>

<sup>39</sup> <http://venturebeat.com/2014/08/12/in-the-world-of-intelligent-software-agents-the-real-hal-may-be-viv/>, zuletzt abgerufen am 30.12.2015

Heutige Visionen über die Erweiterbarkeit solcher Interfaces beziehen die Wahrnehmung, Erkennung und Reaktion von nonverbaler menschlicher Kommunikation und Emotion<sup>40</sup> mit ein, also beispielsweise Mimik, Gestik oder sogar Geruch.<sup>41</sup> Weitere Forschungsthemen (2014/2015) gehen dabei beispielsweise auf die Entwicklung intelligenter Agenten für das Internet of Things bzw. das Smart Home ein<sup>42</sup>.

## 4. Literaturverzeichnis

Alan, Alper T.; Ramchurn, Sarvapali D.; Rodden, Tom; Costanza, Enrico; Fischer, Joel; Jennings, Nicholas R. (2015): Managing energy tariffs with agents: a field study of a future smart energy system at home. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers. ACM, S. 1551–1558.

André, Elisabeth; Dorfmueller-Ulhaas, Klaus; Rehm, Matthias (2005): Engaging in a Conversation with Synthetic Characters Along the Virtuality Continuum. In: David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell et al. (Hg.): Smart Graphics, Bd. 3638. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), S. 1–12. Online verfügbar unter <https://www.informatik.uni-augsburg.de/lehrestuehle/hcm/publications/2005-SG/36380001.pdf>, zuletzt geprüft am 24.11.2015.

Anton Nijholt (2004): Where computers disappear, virtual humans appear. In: *Computers and Graphics* 28, S. 465–476. Online verfügbar unter <http://wwwhome.ewi.utwente.nl/~anijholt/artikelen/compgraph2004.pdf>, zuletzt geprüft am 24.11.2015.

Berner, Uwe (2003): Optimization Strategies for a Scalable Avatar. In: Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, Andreas Butz, Antonio Krüger und Patrick Olivier (Hg.): Smart Graphics, Bd. 2733. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), S. 222–230.

Both, Göde (2012): Better Living Through Siri? Arbeitersparnis, Geschlecht und Virtuelle Assistent\_innen. In: *Bulletin Zentrum Für Transdisziplinäre Geschlechterstudien/Humboldt-Universität Zu Berlin* 40, S. 123–138.

Case, J. Eric; Twyman, Nathan W. (2015): Embodied Conversational Agents: Social or Nonsocial? In: Proceedings on 48th Hawaii International Conference on System Sciences (HICSS). IEEE, S. 491–496.

Catrambone, Richard; Stasko, John; Xiao, Jun (2002): 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, S. 166–171.

Christensen, Anne (2008): Virtuelle Auskunft mit Mehrwert: Chatbots in Bibliotheken. Berlin: IfBB (Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft, 222). Online verfügbar unter <http://opus.kobv.de/zlb/volltexte/2008/1764/pdf/h222.pdf>, zuletzt geprüft am 24.11.2015.

---

<sup>40</sup> Vgl. Kazemifard et al. 2014.

<sup>41</sup> Vgl. Vira und Vasiljevs 2014, S. 26.

<sup>42</sup> Vgl. Alan et al. 2015.

- Clark Elliott; James C. Lester; Jeff Rickel (1997): Integrating Affective Computing Into Animated Tutoring Agents. In: Proceedings of the IJCAI Workshop on Animated Interface Agents: Making Them Intelligent, S. 113–121.
- DEHN, DORIS M.; van MULKEN, SUSANNE (2000): The impact of animated interface agents. A review of empirical research. In: *International Journal of Human-Computer Studies* 52 (1), S. 1–22. DOI: 10.1006/ijhc.1999.0325.
- Eggeling, Thorsten (2005): Karl Klammer & Co. Entsorgen. Online verfügbar unter <http://www.pcwelt.de/ratgeber/Co-Entsorgen-269047.html>, zuletzt geprüft am 24.11.2015.
- El-Bakry, Hazem M.; Riad, Alaa M.; Abu-Elvoud, Mohamed; Mohamed, Samaa; Hassan, Ahmed E.; Kandel, Mahmoud S.; Mastorakis, Nikos (2010): Adaptive User Interface for Web Applications. In: Proceedings of the 4th WSEAS International Conference on Business Administration (ICBA'10). WSEAS, S. 20–22.
- Forlizzi, Jodi; Zimmerman, John; Mancuso, Vince; Kwak, Sonya: How interface agents affect interaction between humans and computers. In: Ilpo Koskinen und Turkka Keinonen (Hg.): the 2007 conference. Helsinki, Finland, S. 209–229. Online verfügbar unter <http://repository.cmu.edu/cgi/viewcontent.cgi?article=1038&context=hcii>.
- Gilbert, Mazin; Arizmendi, Iker; Bocchieri, Enrico; Caseiro, Diamantino; Goffin, Vincent; Ljolje, Andrej et al. (2011): Your mobile virtual assistant just got smarter! In: INTERSPEECH, S. 1101–1104.
- Horchert, Judith (2015): icrosofts Office-Assistent: Was wurde eigentlich aus Karl Klammer? Online verfügbar unter <http://www.spiegel.de/netzwelt/web/microsoft-office-was-wurde-eigentlich-aus-karl-klammer-a-1033286.html>, zuletzt geprüft am 24.11.2015.
- IT Wissen (2015): Benutzeroberfläche. UI (user interface). Online verfügbar unter <http://www.itwissen.info/definition/lexikon/Benutzeroberflaeche-UI-user-interface.html>, zuletzt geprüft am 24.11.2015.
- Kay, Alan (1984): Computer Software. In: *Scientific American* 251 (3), S. 53. Online verfügbar unter <https://frameworker.files.wordpress.com/2008/05/alan-kay-computer-software-sciam-sept-84.pdf>, zuletzt geprüft am 24.11.2015.
- Kazemifard, Mohammad; Ghasem-Aghaee, Nasser; Koenig, Bryan L.; Ören, Tuncer I. (2014): An emotion understanding framework for intelligent agents based on episodic and semantic memories. In: *Autonomous agents and multi-agent systems* 28 (1), S. 126–153.
- Kim, Gerard Jounghyun (2015): Human-Computer Interaction. Fundamentals and Practice. Hoboken: CRC Press. Online verfügbar unter <http://gbv.ebib.com/patron/FullRecord.aspx?p=1693405>, zuletzt geprüft am 24.11.2015.
- King, William Joseph; Ohya, Jun (1996): The representation of agents. In: Proceedings of the Conference Companion on Human Factors in Computing Systems. Vancouver, British Columbia, Canada. New York, NY, USA, S. 289–290.
- Lester, James C.; Converse, Sharolyn A.; Kahler, Susan E.; Barlow, S. Todd; Stone, Brian A.; Bhogal, Ravinder S. (1997): The Persona Effect: Affective Impact of Animated Pedagogical Agents. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM (CHI '97), S. 359–366. Online verfügbar unter <http://doi.acm.org/10.1145/258549.258797>, zuletzt geprüft am 24.11.2015.
- Nass, Clifford; Moon, Youngme; Fogg, B. J.; Reeves, Byron; Dryer, D.Christopher (1995): Can computer personalities be human personalities? In: *International Journal of Human-Computer Studies* 43 (2), S. 223–239. DOI: 10.1006/ijhc.1995.1042.

- Nass, Clifford; Steuer, Jonathan; Tauber, Ellen R. (1994): Computers are social actors. In: Beth Adelson, Susan Dumais und Judith Olson (Hg.): the SIGCHI conference. Boston, Massachusetts, United States, S. 72–78.
- Noteboom, N. (1998): Die Clippy, Die. ZDNet AnchorDesk. Online verfügbar unter [http://www.zdnet.com/anchordesk/story/story\\_2589.html](http://www.zdnet.com/anchordesk/story/story_2589.html), nicht mehr verfügbar.
- Reichle, Meike (2011): Entwicklung eines prototypischen Chatbots für die Universitätsbibliothek Hildesheim: Universitätsbibliothek Hildesheim.
- Rubin, Victoria L.; Chen, Yimin; Thorimbert, Lynne Marie (2010): Artificially intelligent conversational agents in libraries. In: *Library Hi Tech* 28 (4), S. 496–522.
- Russell, Stuart J.; Norvig, Peter (2012): Künstliche Intelligenz. Ein moderner Ansatz. 3., aktualisierte Aufl. München [u.a.]: Pearson (Pearson-Studium).
- Shneiderman, Ben (1993): A nonanthropomorphic style guide: overcoming the Humpty Dumpty syndrome. In: *Sparks of Innovation in Human-Computer Interaction* 16 (7), S. 331–335.
- Shneiderman, Ben (1995): Looking for the bright side of user interface agents. In: *ACM interactions* 2 (1), S. 13–15.
- Swartz, Luke; Nardi, Bonnie A.: Why people hate the paperclip: Labels, appearance, behavior, and social responses to user interface agents. Stanford University. Online verfügbar unter <http://xenon.stanford.edu/~lswartz/paperclip/paperclip.pdf>, zuletzt geprüft am 24.11.2015.
- Vira, Inese; Vasiljevs, Andrejs (2014): The Development of Conversational Agent Based Interface. In: *HTL Baltic* 14, S. 46–53.
- Walker, Janet H.; Sproull, Lee; Subramani, R. (1994): Using a Human Face in an Interface. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM (CHI '94), S. 85–91.
- Wilson, Michael (1997): Metaphor to Personality: the Role of Animation in Intelligent Interface Agents. In: Proceedings of the IJCAI Workshop on Animated Interface Agents: Making Them Intelligent.
- Wooldridge, Michael (2002): Intelligent Agents: The Key Concepts. In: G. Goos, J. Hartmanis, J. van Leeuwen, Vladimír Mařík, Olga Štěpánková, Hana Krautwurmová und Michael Luck (Hg.): Multi-Agent Systems and Applications II, Bd. 2322. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), S. 3–43.

## 5. Abbildungsverzeichnis

Abbildung 1: Agents - Konzeptionelle Betrachtung, eigene Darstellung.....	2
Abbildung 2: Clippy versucht zu helfen.....	7
Abbildung 3: Clippy: Eine Beispiel für Satire.....	8
Abbildung 4: Stella, der Chatbot der Staats- und Universitätsbibliothek Hamburg.....	9
Abbildung 5: Nutzung eines sprachbasierten personal assistants.....	11