

## **Chatbots in der praktischen Fachlexikographie und Terminologie**

---

Chatkommunikation im Sinne eines interaktiven, textbasierten Gesprächs von Internetnutzern als Teil des Internets ist in verschiedenen Benutzungszusammenhängen und für verschiedenste Anwendungen von Marketing bis Freizeit belegt. Als Chatpartner kommen neben anderen Internetnutzern aber auch Computer in Betracht, und auch diese Kommunikationsform ist sowohl in der Wirtschaft als auch im Privatgebrauch bekannt. Der Erfolg eines Chatroboters begründet sich dabei wesentlich in seiner Fähigkeit, einen Dialog mit dem Chatpartner zu führen und sinnvolle Aussagen zu machen. Als Wissensbasis für diese Kommunikation ist neben regelbasierten Verfahren auch ein Rückgriff auf fachlexikographische / terminologische Daten denkbar – nicht zuletzt in einer Fachkommunikation. Der vorliegende Beitrag versucht diese Problematik einzugrenzen und konzipiert Randbedingungen einer möglichen Umsetzung.

### **1 Chatkommunikation**

Chatkommunikation ist im Berufsleben, in der Wissenskommunikation und Bildung und in der Freizeit längst etabliert und belegt; zugehörige Systeme werden mittlerweile auch für Kunden aus dem Unternehmensbereich vertrieben (Bradner et al., 1999; Bremmer, 2005; Döring and Pöschl, 2005; Holzhauser, 2003; Medienpädagogischer Forschungsverbund Südwest, 2007b; Niemann, 2003; Puck and Exter, 2005; Samuelsen, 2003; Zumbach and Spraul, 2005). Dabei ist diese computerbasierte Kommunikation bereits seit der Kindheit der Netzwerke und Personalcomputer keine Neuheit (Weizenbaum, 1966) und heute durchaus gleichwertig zu anderen netzbasierten Kommunikationsformen zu sehen (Storrer, 2001). In ihrer Funktionalität liegt Chatkommunikation zwischen gesprochener Sprache und Schriftlichkeit (Runkehl et al., 1998; Storrer, 2001; Willand, 2002, S. 56) und ist letztlich damit eine Vermischung beider Kommunikationsformen, die auch konzeptionell noch nicht abgeschlossen ist (Beißwenger and Storrer, 2005).

Zur Typisierung der Chatkommunikation als computervermittelter Kommunikation bieten sich Fragen der Technologie, der Nutzung, der Thematik, sowie der Rollenverteilung an (Puck and Exter, 2005, S. 293; van Eckert, 2005, S. 351f.). Dabei ist zu beachten, dass die Kommunikationstechnologie lediglich einen Rahmen setzt, dessen Grenzen existieren, die aber einen breiten und immer breiteren Raum für verschiedenste Kommunikationsformen ermöglichen (Beißwenger, 2005, S. 63f.). Von besonderem Interesse ist hier der Textchat (kein Video, keine gegenseitige Grafikanimation – s.u.), d.h.

die Benutzer schreiben einander und sehen einander nicht. Zudem wird kein themenbezogener Moderator eingeschaltet, so dass die Chatkommunikation frei verläuft und daher kann auch keine Unterscheidung zwischen On-Topic-Chat und Off-Topic-Chat getroffen werden.

Als Chatkommunikation wird hier lediglich der mitunter so bezeichnete „synchrone Chat“ (Crystal, 2006, S. 11 ff.) verstanden, bei dem die Benutzer binnen kurzer Zeit – annähernd also gleichzeitig – lesen und antworten. Dabei kann keine genaue Zeitfrist für das Ende einer annähernden Gleichzeitigkeit festgemacht werden. Trotzdem besteht ein wesentlicher Unterschied zu einer Kommunikation, in der erst Stunden oder Tage später geantwortet wird oder werden kann, und die hier als Forum (Funktionalität kommunikativ ähnlich Bulletin Board oder UseNet; teilweise auch ähnlich ListServ) bezeichnet wird.

Bei der Chatkommunikation handelt es sich prinzipiell gesehen (d.h. in der überwiegenden Mehrheit aller bekannten Fälle) um eine 1:1- bzw. serielle Kommunikation, und zwar ungeachtet der zugrunde liegenden (Netzwerk- bzw. Software-)Technologie. Das bedeutet, dass Eingabe und Ausgabe einander abwechseln, wenn sinnvolle Kommunikation produziert werden soll (Holmer and Wessner, 2005). Damit ist die Chatkommunikation in ihrer Wirkung halbduplex, d.h. es steht nur ein Kanal zum Senden und zum Empfangen der Nachricht zur Verfügung, obwohl die zugrunde liegende Netzwerktechnologie vollduplex verfügbar ist. Genaugenommen steht aber auch dieser ein Kanal der seriellen Kommunikation bei der oftmals üblichen Nutzung eines Browsers als Software auf technischer Ebene durch die beteiligten Netzwerkprotokolle nicht dauerhaft zur Verfügung sondern wird bei jeder Eingabe und Ausgabe erneut aufgebaut. Durch die allgemein übliche Sitzungsverwaltung wird das an sich verbindungslose Protokoll jedoch in einer Sitzung zusammengefasst und damit eine Quasi-Channel erstellt. Eine inhaltliche Persistenz von Chatverläufen wird in einem gewissen Umfang i.d.R. softwareseitig gegeben und somit können frühere Beiträge am Bildschirm noch gelesen werden (Holmer and Wessner, 2005, S. 195). Gegenteilige Beispiele ohne Rückgriff auf den Chatverlauf sind jedoch ebenso zahlreich aufzufinden.

Die Verbindung von Fachlexikographie (oder auch Terminologie) in Theorie sowie Praxis und Chat ist relativ wenig belegt. Dies ist vermutlich in der Tatsache begründet, dass die Chatkommunikation, die fachlexikographische oder terminologische Daten oder Informationen (zur Differenzierung siehe Geeb 2003, S. 416) beinhaltet, allgemein als Wissenschafts- oder Fachkommunikation behandelt wird und auch in diesem Zusammenhang in der Chatkommunikation erfasst wird. Der Fokus verschiebt sich jedoch unmittelbar, wenn der Chatpartner nicht mehr ein Mensch sondern eine Maschine ist, die dem Nutzer am Browser auf Grundlage von fachlexikographischen oder terminologischen Daten antwortet und Informationsbedarfe abarbeitet. Durch die lexikographische und terminologische Datenbasis z.B. eines fachlexikographischen Nachschlagewerks stehen Antwortmöglichkeiten zur Verfügung, die automatisiert auf bestimmte Fragen ausgegeben werden können. Damit ist die Funktionalität eines Chatbots (s.u.) grob umrissen, wobei der Chatbot als maschineller Chatpartner i.d.R. noch mit einer Identität versehen wird um eine größtmögliche Nähe zum bereits durch die zwischen-

liegende massive Technik distanzieren Benutzer (Stein, 2005) zu schaffen. Namensgebung und auch bildliche Erscheinung sind dabei wichtige Elemente und tragen nicht unwesentlich zum Erfolg des Bots bei.

### 2 Chatbot

*Chatbots* (= *Chat Roboter*, auch *Chatterbot* oder *Lingubots* genannt; benachbart: *Avatar*) sind Softwareagenten, die über ein Interface durch eine Kommunikation in natürlicher Sprache den Zugriff auf eine Wissensbasis bieten. Im Gegensatz zu Suchmaschinen verfügen Sie über eine i.d.R. graphisch dargestellte Persönlichkeit und einen hinterliegenden Charakter (Trogemann, 2003, S. 282ff.) mit außersprachlichen Ausdrucksformen (Mimik, Gestik, z.T. auch Stimmführung). Durch einen nur teilweise vorhersehbaren Verlauf des Diskurses und eine mögliche Lernfähigkeit geht im Gegensatz zu früheren Dialogsystemen (Nake, 1988, S. 27ff.) ein Teil der Autonomie über den Dialog von dem System und dem Entwickler des Bots an den Benutzer über. Es ist dabei umstritten, ob der Dialog mit der Maschine in Anlehnung an zwischenmenschliche Kommunikation überhaupt diesen Begriff der Kommunikation verdient oder lediglich eine Form der Interaktion darstellt (Cyraneck, 1988).

Die Wissensbasis eines Chatbots kann idR. durch die Kommunikation mit den Benutzern und eventuell zusätzlichem redaktionellen Eingriff erweitert werden, d.h. der Chatbot lernt oder wird trainiert. Der Chatbot hat dabei im Verhältnis zu anderen Suchwerkzeugen durch Kommunikationsfähigkeit und Vermenschlichung eine hohe Grundakzeptanz bei geübten und ungeübten Benutzern (Wirth, 2003, S. 124f.).

Chatbots sind nicht zuletzt in ihrer moralischen, ethischen und rechtlichen Konsequenz umstritten. So wird teilweise belegt, dass die anthropomorphe Metapher durch diese starke Personalisierung eines Maschineninterfaces beim Benutzer zu hohe Erwartungen in der Kommunikation/Interaktion mit der Maschine hervorruft (Lindner, 2003, S. 9ff.). Es wird konstatiert, die Einführung menschlicher Züge in das Maschineninterface sei für Kommunikationszweck schädlich (Braun, 2000, S. 83 ff.). Zudem sei es dem Benutzer durch die Verwendung eines Chatbot nicht möglich, die Reaktion der Maschine vorherzusagen oder mehrfach erneut exakt abzurufen, worin ein erhebliches Usabilityproblem besteht (Shneidermann, 1997).

Die Historie der Chatbots geht bis in die 60er Jahre zurück, in denen ELIZA als erster bekannter Chatbot Fragen durch Rückfragen und vorgefertigte Antworten beantworten konnte (Weizenbaum, 1966). Chatbots wird neben einer Vereinfachung der Informationssuche auch ein Potential für Einsparungen in Unternehmen zugeschrieben (Braun, 2000, S. 36ff.). Schließlich bieten die Chatbots auch im Bereich des Customer Relationship Managements unternehmensrelevante Informationen durch eine verbesserte Analysemöglichkeit von Informationsbedarfen und Suchpfaden (Morphy, 2001; von Wendt, 2003, S. 46ff.). Trotz aller Begeisterung für Chatbots und ihre Wertigkeit in der Mensch-Maschine-Kommunikation (Wolff, 2003) werden sie nicht nur positiv gewertet und haben z.B. in dem Einsatz als unterstützende Lernagenten im Fremdsprachener-

werb nur wenig Erfolg gezeigt (Jia, 2007) – ganz im Gegensatz zur Chatkommunikation Mensch-Mensch in diesem Bereich (Darhower, 2002).

Der Chatbot dient in vielen bekannten Formen in der Regel der Lösung konkreter Wissens- oder Produktanfragen und damit einer spezifischen Chatsituation, die sich z.B. vom Alltagsplaudern oder Chatflirt deutlich in Kommunikationsform (Informationsbedarf) und Kommunikationsmittel (Denotation und Konnotation) unterscheidet. Zweifellos ist allerdings in der Kommunikation mit dem Chatbot mit einer im Chat generell bekannten informellen Kommunikationsform (Harnoncourt et al., 2005, S. 165ff.) zu rechnen. Die Regelmäßigkeit von themenbezogenen Chats (Döring and Pöschl, 2005, S. 154ff.) ist hier nicht anzusetzen, da der Austausch alleine zwischen Benutzer und Bot erfolgt.

Bei der Nutzung eines Chatbot wird prinzipiell in der 1:1-Kommunikation gearbeitet. Damit besteht eine Kommunikation immer nur zwischen dem Chatbot und einem Benutzer. So befinden sich im Chat-Raum als traditionellem Merkmal des Chats (Crystal, 2006, S. 11ff.) immer nur diese beiden Akteure. Technisch antwortet der Chatbot selbstverständlich auf viele Benutzer, aus der Sicht der Kommunikation relevant ist jedoch die Tatsache, dass das Gespräch Benutzer-Chatbot geschlossen ist und keine weiteren Benutzer (Mensch oder Maschine) daran teilnehmen. Damit stehen – realisiert durch entsprechende Prozesse auf dem Server beliebig viele Chatbots zur Verfügung, die eine weitgehend gleiche Funktionalität (Unterschiede in den Sitzungen mit dem Chatbot ergeben sich durch eventuelles interaktives Lernen des Chatbots, s.u.) und das exakt gleiche Aussehen haben. Die Bezeichnung als eine 1:1 Kommunikation ist damit gerechtfertigt und die fatale Problematik z.B. paralleler Chatbeiträge (Holmer and Wessner, 2005, S. 183ff.) entfällt. Insofern ist der Chatbot ein Chatsystem, und damit eine konkrete Ausprägung und Anwendung von Chat-Technologie (hierzu: Beißwenger 2005, S. 68ff.).

Gängige Steuerungs- und Kommunikationsprobleme des Chat wie z.B. mangelnde Zuordnung von Personen und deren Aussagen bei steigender Zahl der Roombenutzer (Chen and Sun, 2007) entfallen beim Chatbot. In sofern ist auch ein threaded Chat (Cadiz et al., 2000) in diesem Kontext nicht erforderlich. Chatbots wurden gemäß einer Studie unter deutschen Internetnutzern bereits 2001 als virtuelle Berater von 40 Prozent der Nutzer gewünscht (Braun, 2000, S. 13). In diesem Sinne hat z.B. die Staats- und Universitätsbibliothek Hamburg mit „Stella“ eine virtuelle Ansprechpartnerin geschaffen (Staats- und Universitätsbibliothek, 2005), die zudem für Ausführung, Wirkungsweise und Benutzung ausgezeichnet wurde (BUB, 2007). Bekannte Anwender von Chatbots sind aber auch Coca-Cola, Ford, Schwäbisch-Hall, Deutsche Bank, Bertelsmann, YelloStrom und viele andere mehr.

### 3 Chatbot-Technologie

Grundlagen des Chat ist das System IRC, Internet Relay Chat, das Ende der 80er Jahre bekannt wurde. Chaträume werden idR. hier von mehreren Benutzer genutzt und ein Automat (Chatbot) ist nicht vorhanden. 1:1 Sitzungen sind möglich, immer ist aber

eine gesonderte Software mit eigenem Protokoll erforderlich (IRC-Client). Später in der Entwicklung anzusiedeln ist das web- und browserbasierte Chat, das in seiner Geschwindigkeit langsamer sein kann, prinzipiell aber ähnliche Funktionalitäten bietet. Ein Chatbot arbeitet prinzipiell mit eben dieser Technologie indem jeweils abwechselnd vom menschlichen Benutzer und Chatbot geantwortet wird. Denkbar und diskutiert neben den hierbei verwendeten klassischen Protokollen sind auch Streamingprotokolle (Vronay et al., 2004). In der konkreten Ausgestaltung eines Chatbots lassen sich nach der Komplexität der Reaktion des Bots verschiedene Chatbottypen in Anlehnung an und Erweiterung zu *Braun* unterscheiden (*Braun*, 2000, S. 43ff.):

- Der Chatbot beantwortet nur Fragen, die vollständige oder beinahe vollständige Übereinstimmung mit im System hinterlegten Fragen haben; eine Standardantwort wird gegeben.
- Der Chatbot antwortet auf eng definierte Fragetypen mit einer gewissen, begrenzten Beliebigkeit innerhalb der nachgefragten Information.
- Der Chatbot arbeitet mit Mustererkennung als einer zentralen Technologie (von *Wendt*, 2003) in der Fragestellung und antwortet dann mit im System hinterlegten Antworten. Dieses *Cased-Based Reasoning* oder *Nearest-Neighbor Classification* definiert Alice (zu Alice s.u.) „For every input, we find the best matching “case” in the pattern set, and generate a reply based on the associated template“ (*Artificial Intelligence Foundation*, 2007b). Der Chatbot antwortet auf eine ihm unbekannt Frage mit einer Rückfrage und sieht die dann erhaltene Rückantwort des Benutzers als mögliche Beantwortung der vom ihm als Bot gestellten Rückfrage an. Ein Frage-Antwortpaar wurde gebildet. Die Gefahr von Falschinformation des Bots ist hier aber nicht unerheblich, denn der Bot muss – wenn kein redaktioneller Eingriff vorhanden ist – alle Aussagen als wahr annehmen.
- Der Chatbot hat verschiedene Verhaltensweisen (= Programme) für verschiedene Kommunikationsformen und Themen, die je nach initialer Eingabe durch den Benutzer genutzt werden und thementypische Fragestellungen beantworten können.

Zahlreich genutzt als Chatbot ist die Technologie, auf der Alice aufbaut (*Artificial Intelligence Foundation*, 2007a; *Reichle*, 2006). Alice basiert auf einer definierten XML-Sprache für eine Mustererkennung von Fragen an den Bot und deren Antworten, AIML (*Wallace*, 2002). Durch die Trennung von Interpretation und Darstellung einerseits und Wissensbasis in AIML andererseits ist der Aufbau der Wissensbasis softwareunabhängig möglich. Die Interpretation und Darstellung können durch frei verfügbare Programme in verschiedenen Sprachen durchgeführt werden. Verschiedene Chatbots auf Basis von Alice und AIML sind belegt (*Braun*, 2000, S. 44ff.) und haben dabei u.U. auch eine gewisse Lernfähigkeit. Als eine vergleichbare Möglichkeit kann bei Pandorabots auf der Grundlage von Alice ein Bot generiert werden, der dann von dem dortigen Server gehostet wird (*Boguschewski*, 2005).

## 4 Chatbot für ein Online-Lexikon

Ein Chat eignet sich prinzipiell auch für Informations- und Beratungsgespräche (Beißwenger, 2005, S. 10), wobei die automatisierte Anwendung als Chatbot noch kaum beschrieben ist, jedoch auf allgemeine Eigenschaften kommunikativer Schnittstellen in der Informatik zurückgeführt werden kann (Schmitt, 1983, S. 22 ff.). Es gilt dabei einen möglichst hohen Grad an Interaktivität (Nake, 1988, S. 16 f.) im Dialog zwischen Nutzer und lexikographischem Nachschlagewerk zu erzielen. Dabei wird davon ausgegangen, dass sich das Diskursmuster im wesentlichen auf den Verlauf *Frage* → *Antwort* mit eventuell *n* zwischengeschalteten Rückfragen und Rückantworten reduzieren lässt. Die Beratungsfunktion einer Verbraucherzentrale eignet sich damit für einen konkreten Test mit einem Chatbot. Ausgangspunkt sei hierbei ein Online-Lexikon für junge Verbraucher, das an der generellen Beratungs- und Informationsarbeit einer Verbraucherzentrale ansetzt (Geeb, 2006; Geeb and Spree, 2005). Dieser Rahmen determiniert bereits die Domäne durch die im Lexikon vorhandenen Themen, die sich ausschließlich auf die generelle Beratungsthematik der Verbraucherzentrale bezieht. Die Wissensbasis des Chatbots lässt sich damit zumindest domänenspezifisch als Zielsetzung aus diesem fachlexikographischen Nachschlagewerk ableiten. Für den Chatbot steht dabei in Analogie zu einem korpusbasierten Training des Bots (Abu Shawar and Atwell, 2005) eine automatisierte Wissensbasis zur Verfügung, die allerdings im Gegensatz zu einem allgemeinsprachlichen Korpus als lexikographisches Nachschlagewerk eine differenzierte fachliche Informationstruktur aufweist.

Aus syntaktischer und semantischer Sicht wäre der Umgang mit besonderen Ausdrucksformen, die im Chatbereich außerhalb von Bots sonst bekannt sind wie Emoticons und Abkürzungen (Diekmannshenke, 2005) im Chatbot theoretisch realisierbar. Eine große Gruppe der Chatbenutzer ist regelmäßiger Nutzer dieser Emoticons (Husmann, 1998, S. 74). Andererseits ist dieser umfangreiche Gruppencode (Husmann, 1998, S. 34ff.) dem nicht unbedeutenden Anteil der anzusetzenden Chatbotnutzer, der die Chatgepflogenheiten des Internets nicht kennt (Medienpädagogischer Forschungsverbund Südwest, 2007a), unbekannt und wäre für diese Gruppe in den Antworttexten des Chatbots damit ein Kommunikationshindernis. Eine Realisierung der Emoticons in Frage und Antwort würde außerdem eine intensive Pflege eines entsprechenden Datenbestandes erfordern. Zudem ist der Mehrwert in der konkreten Benutzerintention „Sachinformation“ (Geeb, 1998, S. 58ff.) nicht wirklich erkennbar, denn diese Chatausdrucksformen dienen insbesondere der Sozialisierung und Abgrenzung von Gruppen (Crystal, 2006, S. 171ff.), nicht aber dem Informationsbedarf, der durch den Chatbot abgedeckt werden soll. Dies gilt um so mehr, als ein wesentlicher Teil der Zielgruppe das Medium Internet zur Zeit (Shuli and Nielsen, 2002, 110 f.) in der Hauptsache für Informationszwecke nutzt (Medienpädagogischer Forschungsverbund Südwest, 2007b, 44 f.). Die Emoticons sollten damit im Sinne der *substitution normalized form* wie aus Alice bekannt ersatzlos aus der Eingabe gelöscht werden (Abu Shawar and Atwell, 2002).

Trotzdem muss aber mit einer informellen Ausdrucksweise und auch mit dem Verzicht auf gewissen grammatikalische Regeln wie z.B. Groß-/Kleinschreibung gerechnet werden (Crystal, 2006, S. 176f.; Puck and Exter, 2005, S. 297f.), und sogar Tippfehler müssen in Kauf genommen und bei der der Anwendung des Gesprächs auf die (lexikographische) Wissensbasis bedacht werden (Willand, 2002, S. 59).

Kohärenzbildung in computervermittelter Kommunikation durch Training / Schulung (Zumbach and Spraul, 2005, S. 380) ist in diesem Fall kaum durchführbar, da die Benutzer ständig wechseln und eine vollständig heterogene und unbekannte Gruppe bilden. Auch scheint die Benutzung eines Chatbots prinzipiell auf Grund des natürlichsprachlichen Interfaces keine weiteren Erläuterung zu erfordern (Wirth, 2003, S. 124f.). Tatsächlich ist Chatten mit seiner Chatiquette ein relativ etabliertes Kommunikationssystem, das aber nicht in seinem vollständigen Umfang in einem Chatbot realisiert wird – und diese Grenzen und Unterschiede zum normalen Chat müssen im Sinne der Benutzerfreundlichkeit vermittelt werden. Ein entsprechendes Training kann und muss sich in diesem Fall auf eine Kurzanweisung beschränken. Diese Kurzanweisung kann in Textform gegeben werden, sollte aber zur Erhöhung der Akzeptanz und Beachtung auch als Kurzvideo (z.B. Flash) mit menschlicher Stimme dargeboten werden oder multimedial von der Chatfigur und damit dem Chatbot selbst geäußert werden. Geschriebene Anweisungen und Hilfstexte werden in der Regel nur wenig beachtet, insbesondere im Internet und bei einer jüngeren Zielgruppe wird eine intuitive Benutzung des Systems erwartet.

Durch die Kommunikation mit dem Computer, der im vorliegenden Konzept in relativ natürlicher Sprache antworten soll, werden voraussichtlich zahlreiche Nutzer eine gewisse emotionale Bindung mit diesem virtuellen Gesprächspartner entwickeln (Reeves and Nass, 1996). Somit kommt der graphischen und evtl. akustischen Ausgestaltung der Chatbotfigur eine wichtige Bedeutung im Sinne einer ästhetischen Dimension von Information zu (Hentschläger and Wiener, 2002). Die Mehrzahl der aktiven Chatbots ist weiblich (auch männliche sind belegbar: z.B. Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg (2007)) und verfügt über ein wissenschaftlich sicher nicht belegbares aber als Grundkonsens doch bekanntes Aussehen, das dem gängigen Schönheitsideal (räumlich und zeitlich) entspricht. Bei der Stimmführung weiblicher Chatbotcharaktere hat sich offenbar eine tiefere Frauenstimme gegen eine höhere, femininere Stimme als erfolgreicher durchgesetzt (Braun, 2000, S. 63ff.). Im Hinblick auf die Zielgruppe ist eine junge erwachsene Person als Chatbotgrafik zu bevorzugen – so werden Jugendlichkeit und Erwachsenenkompetenz in dem Beratungsgespräch verbunden, wobei konzeptionell im Sinne der Gleichstellung und Gleichberechtigung dem Benutzer die Wahl zwischen einer männlichen und einer weiblichen Chatbotfigur gegeben werden sollte. Die Figur darf dabei nicht statisch sein, sondern muss möglichst synchron mit dem Gesprächsverlauf Mimik und Gestik zeigen und ändern (Braun, 2000). Diese Gestik und Mimik muss dabei im unten genannten Modell der Antwortgenerierung immer rudimentär bleiben, da eine syntaktische und satzsemantische Analyse nicht vorgesehen ist. Außersprachliche Ausdrucksformen zeigen damit im vorliegenden Entwurf nur eine Grundstimmung der Kommunikation wie dies auch

vielfach belegt und in der zugrundeliegenden Technologie begründet ist – mit Alice (Artificial Intelligence Foundation, 2007a) als prominentestem Vertreter.

Die Bindung des Nutzers an den Chatbot darf allerdings nicht zur Bildung neuer Identitäten führen, wie dies in normalen Chats teilweise möglich ist. Dort können – je nach System – Chatnutzer durch Figuren (Avatare) und ihr Chatverhalten teilweise oder völlig neue Identitäten im Verhältnis zum wirklichen Leben annehmen (Filinski, 1998, S. 107ff.; Willand, 2002, S. 79ff.). Diese Verhaltensweise wäre im Rahmen des Chatbot technisch realisierbar, aber sehr aufwendig und zudem nicht produktiv. Diese Identitätsbildung auf Nutzerseite ist letztlich nicht Teil eines reinen Informationsbedürfnisses und würde durch die vielfachen sozialen und psychologischen Elemente den Chatbot inhaltlich in seiner Funktion als Informationsassistent mit der vorhandenen Wissensbasis des Lexikons überfordern.

Der zu konzipierende Chatbot hat keinen Moderator, d.h. es handelt sich einerseits nicht um einen moderierten Chat und andererseits werden nicht beantwortbare oder beantwortete Fragen der Nutzer nicht automatisch an einen HelpDesk o.ä. weitergeleitet, der zeitnah antworten kann. Dieses Verfahren der zeitnahen Weiterleitung unbeantworteter Fragen an den Chatbot (z.B. produktbezogene Themen) ist aus der Wirtschaft bekannt und wünschenswert (Salimi, 2003, S. 136ff.; Thommes, 2001), für einen öffentlichen Träger ohne Nutzerfinanzierung oder direkte Kopplung eines Customer Relationship Managements an einen wirtschaftlichen Erfolg aber kaum durchführbar.

Die Lernfähigkeit des Chatbots ist erwünscht um die Akzeptanz des Chatbots bei den Nutzern zu fördern. Die Erstellung eines zu veröffentlichenden Transkriptes des Chat, das bei moderierten Chats u.a. in der Politik bekannt ist, ist an dieser Stelle nicht sinnvoll, da ein Nacharbeiten der Suchanfragen und Suchstrategien einzelner Nutzer durch andere Nutzer zeitaufwendig wäre. Dieser Zeitaufwand entspricht aber eben nicht dem Konzept der schnellen und benutzergesteuerten Information eines Chatbots. Andererseits ist der Chatverlauf in einem systeminternen Protokoll zu speichern, das dann für die Überprüfung der Funktionsfähigkeit, Nutzbarkeit und Anwendungsfelder durch redaktionelle Administratoren im Überblicksverfahren genutzt werden kann.

Zu beachten sind bei der Erstellung eines Chatbots in jedem Fall Kriterien der Benutzbarkeit von Webseiten und deren Dienste allgemein (z.B. Nielsen 2001). Hier ist zum einen eine Heuristik (Nielsen, 1993, S. 115ff.) für die Evaluation des Chatbots und seiner Charaktere zu entwickeln; zum anderen sind exemplarische Labortests mit einer Videokontrolle der Benutzerarbeit mit dem Chatbot anzufertigen (Nielsen, 1993, S. 200ff.). Insbesondere die Wartezeit des Benutzers auf eine Antwort ist bekannter Maßen ein wesentlicher Erfolgs- und Usabilityfaktor (Wirth, 2003, S. 125) und es gilt, diese durch eine entsprechende Konzeption des Programmflusses sowie durch die Programmierung und Gestaltung der Wissensbasis möglichst gering zu halten.

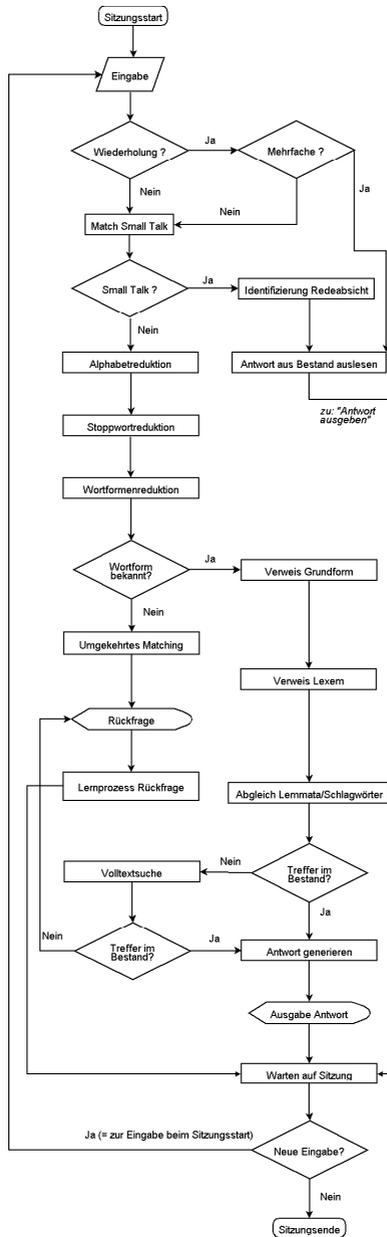


Abbildung 1: Schematischer Programmablauf eines Chatbots

Auch die Grundsätze der Barrierefreiheit der *Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz* (Behindertengleichstellungsgesetz, 2002) sind für fachlexikographische / terminologische Produkte aus ethischen Gründen – und im vorliegenden Fall auch z.T. aus rechtlichen Gründen – zu beachten, und so muss der Chatbot in reiner Textform steuerbar und verständlich sein. Auch muss aus diesem Grund nebenläufig zur Videokurzeinführung des Systems (s.o.) ein Textstream mit diesen Informationen angeboten werden. Der Programmablauf des Chatbots lässt sich wie in Abb. 1 dargestellt zusammenfassen.

Der Chatbot steht prinzipiell in unbegrenzt vielen Ausgaben zur Verfügung. Jede Kommunikation mit einem Benutzer von deren Beginn zu deren Ende wird als Sitzung bezeichnet und technisch gesehen auch als solche realisiert. Der Sitzungsbeginn entspricht der ersten Eingabe des Benutzers. Die Sitzungslänge und damit die Entscheidung, wann eine neue Sitzung beginnt und damit das Gedächtnis des Chatbots im Bezug auf diesen Benutzer gelöscht wird, muss getroffen werden, lässt sich aber jederzeit problemlos revidieren. Ein vorstellbares Vorgehen wäre nach 60 Sekunden Inaktivität die Sitzung serverseitig durch Löschen der Sitzungsdaten (= Gedächtnis des Chatbots) zu beenden. Ein Kurzzeitgedächtnis des Chatbots ist aber in jedem Fall erforderlich (Boguschewski, 2005).

Bei der ersten Eingabe tritt die darauf folgende Verzweigung mit Prüfung auf *Wiederholung* ist bei Sitzungsstart nicht aktuell. Die Eingabe wird daraufhin durch Pattern-matching auf *Small-Talk* untersucht wie dies durch Alice (Artificial Intelligence Foundation, 2007a) und z.B. ein deutsches Small-Talk-AIML-Set zu diesem Zweck belegt ist (Drossmann, 2001). Als Beispiel und Auszug aus dem genannten Set hier die übliche Frage nach der Identität des Bot, die dann noch in verschiedenen Schreibvarianten vorgelegt wird:

```
...
<category>
  <pattern>BIST DU EIN MENSCH</pattern>
  <template>Nein, ich bin eine Maschine.</template>
</category>
<category>
  <pattern>BIST DU EIN PROGRAMM</pattern>
  <template>Ja, ich bin ein Programm.</template>
</category>
...
<category>
  <pattern>DAS DENKE ICH *</pattern>
  <template>Aber Du bist Dir nicht sicher, oder?</template>
</category>
...
```

Alle Gesprächsversuche mit dem Chatbot, die nicht dem reinen Informationsbedürfnis dienen, werden hier durch hinterlegte Muster vergleichbar diesen rudimentären Beispielen aus AIML geprüft, wobei die Wirkung einer Erweiterung von AIML durch reguläre Ausdrücke und Stemming überprüft werden muss. Dies betrifft Fragen mit Inhalten wie „*Wie alt bist Du*“, „*Wer hat Dich programmiert*“ oder im Stil von „*Willst Du*

*Dich mit mir treffen*“. Diese Art der Kommunikation mit dem Chatbot ist durchaus üblich (Davidson, 2005; Reichle, 2006). Wenn es sich um Small-Talk handelt, wird dieser weiter auf die *Redeabsicht* untersucht und mit einer passenden Antwort aus dem Antwortbestand assoziiert. Handelt es sich um ernst gemeinte aber nicht relevante Fragen („Wie alt bist Du“ etc.) werden die Antworten ähnlich ernst gemeint sein, aber freundlich auf den eigentlichen Sinn des Chatbots hinweisen. Enthalten die Fragen Zweideutigkeiten, werden diese zurückgewiesen (Lindner 2003, 11; beim Chatbot *Stella*: Bachfeld et al. 2005, S. 209) und ggf. mit einem neutralen Satz beendet. Im Gegensatz zu einem normalen Chat (Willand, 2002, S. 64f.) kann der Bot den unerwünschten Nutzer auf Grund dessen völligen Anonymität aber nicht technisch ausblenden und muss immer wieder mit entsprechenden Eingaben rechnen. Es liegt in der Natur der Sache, dass nicht alle Small-Talk Fragen vorhergesehen werden können und deswegen u.U. weitere Prozesse der normalen Beantwortung trotz Small-Talk eingeleitet werden. Als Alternative könnte man die Definition des Small-Talk durch ein Patternmatching gegen die bekannten Informationsfragen definieren, würde dann aber einen Großteil der unbekannteren aber relevanten Anfragen als Small-Talk deklassifizieren. Zudem wäre die Lernmöglichkeit des Chatbots, die an späterer Stelle in der Bearbeitung normaler Anfragen vorliegt, wesentlich eingeschränkt. Bei der sprachlichen Analyse der Eingaben wäre eine prinzipielle Unterscheidung zwischen negativer und positiver Grundaussage hilfreich zur Steuerung des weiteren Gesprächsverlaufs (Turney, 2003). Dieser Analyse-schritt wird aber auf Grund des hierfür erforderlichen umfangreichen Textkorpus für dieses Vorhaben zunächst nicht in die Konzeption aufgenommen.

Wenn normale Anfragen vorliegen, werden diese zunächst durch eine *Alphabetreduktion* geführt und nicht etwa mit einem extremen Patternmatching (von Wendt, 2003, S. 42ff.) bereits eine Antwort generiert. Bei der Alphabetreduktion werden alle nichtalphabetischen Zeichen entfernt mit Ausnahme von diakritischen Zeichen und Ligaturen sowie Bindestrichen, die ohne Zwischenraum zwischen zwei Buchstabenblöcken vorkommen. Dies könnten Komposita sein, die die Benutzer auf Grund der vermutlich weniger regeltreuen Schreibweise in Anlehnung an andere Sprachen mit Bindestrich schreiben. Zahlen (arabisch oder römisch) könnten ebenfalls für die Semantik der Anfrage interessant sein; daher wäre dieser Teil des Programmablauf so zu konstruieren, dass die Liste der gültigen Zeichen zu jedem Zeitpunkt einfach (z.B. ohne weiteres Kompilieren von Quellcode) erweitert werden kann; im genannten Fall der Zahlen wäre die Wirkung mit oder ohne Zahlen dann zu testen. Der Abschluss einer Benutzereingabe durch ein Fragezeichen wird in der Sitzung als besondere Information gespeichert; in diesem Fall kann davon ausgegangen werden, dass definitiv eine Frage gestellt wurde. Das Fragezeichen an sich wird aber aus der Eingabe gelöscht.

Gängige Wörter, die sehr häufig auftreten und im Retrieval von Informationen daher eine geringere Bedeutung für die Erfassung des Inhalts der Benutzereingabe haben, werden im nächsten Schritt durch eine *Stoppwortreduktion* entfernt; nicht gelöscht werden dürfen Wörter wie „Ja“ und „Nein“ etc., die bei Rückfragen Aufschluss über Sachzusammenhänge geben. Diese Stoppwortliste muss ebenfalls erweiterbar sein (z.B. Datenbank als Basis) und lässt sich durch einen entsprechenden Textkorpus z.B. auf

der Grundlage des Lexikons, für den der Chatbot geschrieben wurde, gewinnen und verfeinern. –Schließlich erfolgt eine *Wortformenreduktion*, in der die eingegeben Buchstabenblöcke durch Patternmatching auf eine Grundform, das Lexem, zurückgeführt werden. Die Datengrundlage für diesen Arbeitsschritt ist sehr aufwendig – und je aufwendiger sie gehalten wird, je besser wird das Ergebnis dieses Schritts sein. So müssen prinzipiell alle Wortformen der Lexeme, die Bestand des Lexikons als Lemmata oder auch in beschreibenden Texten sind und die nicht zur Stoppwortliste gehören sollen, hier vermerkt werden. Zusätzlich sind noch umgangssprachliche und gruppensprachlich relevante Schreibvarianten zu berücksichtigen. Alle Formen und Varianten zeigen über ihre Grundform auf ein Lexem (*Brille, Brillen, Nasenfahrrads, ...* → *Brille*). Somit wird das Thema der Suchanfrage eingegrenzt. Bei Suchanfragen wie „*brauche ich eine bildschirmbrille?*“ wird damit möglicherweise keine Wortform gefunden und bei der Verzweigung *Wortform bekannt?* wird der Programmfluss auf eine Rückfrage umgelenkt. Die Rückfrage des Chatbots sucht durch den Schritt *Umgekehrtes Matching* alle Lexeme heraus, die in der unbekanntem Wortform enthalten sind wie z.B. *Bildschirm* und *Brille*. Sollten sich daraus – wie im vorliegenden Fall – Komposita bilden lassen, die der nicht gefundenen Wortform ganz oder teilweise entsprechen, ist die Rückfrage mit Rückgriff auf die beiden bekannten Wortformen einfach: „*Bildschirmbrille kenne ich nicht. Meinst Du Bildschirm und Brille?*“

Hier wurde bereits über die Sitzung der Lernprozess gestartet, d.h. die initiale Frage des Benutzers wird gespeichert und auch die Antwort des Chatbots (= Rückfrage) wird gespeichert usw. Nach Beendigung der Sitzung werden diese Daten strukturiert zur redaktionellen Nachbearbeitung (= eventuelles Einpflegen von *Bildschirmbrille*) abgelegt. Nun wird auf eine neuerliche Eingabe des Benutzers gewartet. Wenn diese nach binnen der eingestellten Frist erscheint, wird die Sitzung fortgeführt und die neue Eingabe jetzt auf eine Wiederholung der ursprünglichen Benutzereingabe untersucht. Wird die ursprüngliche Eingabe mehrfach wiederholt (auch dieser Wert muss konfigurierbar sein), wird eine entsprechende Antwort aus dem festen Antwortbestand ausgelesen und der Timeoutzähler der Sitzung gestartet. Mit anderen Worten kann der Benutzer *n*-fach denselben Text eingeben und wird eine entsprechende Antwort im Stil von „*Das ist mir nicht neu*“ erhalten, wird dem Chatbot aber keine weitere Aktion entlocken können bzw. vermutlich durch den Timeout gestoppt. Der Timeout ist technisch gesehen nicht erforderlich, da der Bot durch die gleichbleibende Antwort nicht in seiner Funktion eingeschränkt wird. Da jedoch eine sinnvolle Kommunikation mit dem Bot gefördert werden soll, wäre der Timeout ein entsprechendes Mittel hier einzugreifen.

Wird die ursprüngliche Rückfrage des Chatbots mit „*Ja*“ beantwortet, ist einerseits der Lernprozess abgeschlossen (s.o.) und andererseits werden die beiden zuvor genannten Lexeme als relevant eingestuft. In diesem Fall – so wie auch wenn die Wortform bereits bei der ersten Anfrage auf ein Lexem reduziert werden konnte – wird über den Schritt *Verweis Grundform* die Grundform gesucht, die dann in *Verweis Lexem* auf ein bestimmtes Lexem zeigt und in der Datenbasis auch Informationen über das Verhältnis der Grundform zum Lexem enthält. So kann hier z.B. vermerkt werden, dass *Nasenfahrrad* Jugendslang für *Brille* ist. Durch diese Kennzeichnung von Bedeutungsbe-

ziehungen (Geeb, 1998, S. 164ff.) wird die Pflege der Datenbasis erheblich erleichtert, indem Bereiche wie Slang oder Abstraktionsbeziehungen kontrolliert gesichtet werden können.

Über die Kenntnis des Lexems kann der Chatbot dann im *Abgleich Lemmata/ Schlagwörter* die Lexikonartikel in der Sitzung ablegen, die über das Lexem referenziert sind. Das sind auf einer ersten Stufe Lexikonartikel, in denen das Lexem als Lemma auftritt. Wenn diese Suche nicht zum Erfolg führt erscheinen in einer zweiten Stufe alle Lexikonartikel relevant, in denen das Lexem als Schlagwort genannt ist, wobei eine intellektuelle Verschlagwortung der Lexikonartikel vorausgesetzt wird. Wird dann durch die Verzweigung *Treffer im Bestand?* ein passender Artikel gefunden, sind die folgenden Schritte *Antwort generieren* und *Ausgabe Antwort* problemlos. Wird aber unter *Treffer im Bestand?* kein Match gefunden, muss als dritte und damit relativ unscharfe Suche eine Suche des gefundenen Lemmas im Volltext des Lexikons gestartet werden. Führt dies über das folgende *Treffer im Bestand?* zu einem Ergebnis, kann eine Antwort angeboten werden, die aber semantisch an diese unscharfe Suche angepasst werden muss. Konnte kein Match gefunden werden, wird die *Rückfrage* angestoßen, die dann vom Benutzer eine Neuformulierung der Eingabe erbittet. Nach Ausgabe der Antwort als Link auf das entsprechende Lemma mit einsätziger kurzen Vorformulierung der semantischen Informationen/ Definition (Geeb, 1998, S. 145ff.) wird dann über den Timeout auf eine Neueingabe des Benutzers gewartet oder die Sitzung dementsprechend geschlossen.

Das gezeigte Modell führt allerdings in definierbaren Fällen, die bereits im jetzigen Konzeptstadium vorhersehbar sind, zu einer problematischen Vorgehensweise. So wird u.U. nur ein Lexem über das genannte Verfahren aus der Wissensbasis gefunden und als Antwort generiert; tatsächlich waren aber mehrere Lexeme relevant – aber dem System nicht bekannt. Als Beispiel sei hier genannt: *Ist Rippen von CDs verboten* wobei *CD* als Lexem in der Wissensbasis bekannt und damit antwortfähig ist während *Rippen* als Slang für *Auf den Rechner kopieren* nicht erkannt wurde. Die Antwort wäre damit allgemein zu Compactdiscs, nicht aber zur Frage der Legalität des Kopiervorgangs. Diese Problematik lässt sich in der gezeigten Annahme nur lösen, indem der Benutzer über eine Rückfrage erneut eindeutig nach *Rippen* nachfragt und damit einen protokollierten Lernprozess ausführt.

Ebenfalls problematisch ist das Auffinden mehrerer Lexeme gleicher Relevanz, d.h. durch die Reduktion werden mehrere Lexeme lokalisiert, die alle als Lemma vorhanden sind und damit höchste Relevanz besitzen. So fragt der Benutzer *Machen Alkopops süchtig* und erhält aus der Wissensbasis die beiden lemmatisierten Lexeme *Alkopops* und *Sucht*. Als Lösung kann ihm nur eine differenzierte Antwort angeboten werden die besagt, dass mehrere relevante Informationen vorliegen.

Eine ähnliche Konstruktion ergibt sich durch Homonymie und Polysemie, die letztlich nur durch eine Fachsystematik in Verbindung mit den Lemmata aufgelöst werden könnte. Diese Fachsystematik muss dann wieder auf die Lexeme nach dem Verweis aus den reduzierten Grundformen abgebildet werden und bei mehreren gefundenen relevanten Lexemen müsste ein Vergleich dieser damit gefundenen fachlichen Einträge und ggf. ein Entscheidungsprozess zur Relevanz eingeleitet werden. Dieses Verfah-

ren ist einerseits umfangreich im Programmfluss aber andererseits vor allen Dingen aufwendig in der redaktionellen Arbeit und wurde daher hier nicht berücksichtigt.

## 5 Ausblick

Die in der Literatur anzutreffende Differenzierung zwischen Fachlexikographie und Terminologie (z.B. Budin and Wright, 1997; Arntz et al., 2004; Bergenholtz, 1995; Wiegand, 1979) wurde hier übergangen und kann nicht Thema aus der Sicht des Chatbotkonzepts sein. Der Ansatz ist prinzipiell für beide Bereiche vergleichbar, denn die für den Chatbot erforderliche Wissensbasis wird auf ähnlicher Grundlage generiert. Durch die zahlreichen Nutzungsmöglichkeiten und Nutzungsänderungen im Bereich der neuen Medien wird die Grenze zwischen Online-Fachlexikographie und Online-Terminologie zukünftig noch mehr an Prägnanz verlieren und nur in Randbereichen wie z.B. Normung einerseits und etymologischen Fachwörterbüchern andererseits noch deutlich sichtbar sein. Ein Chatbot im vorgestellten Konzept könnte von dieser Differenzierung nur rudimentär profitieren.

In der Chatkommunikation wird auf der Nutzerseite oftmals eine Verbesserung der Chat-Kompetenz gefordert (Döring and Pöschl, 2005, S. 159). Die ebenfalls in diesem Zusammenhang oftmals geforderte Verbesserung der technischen Möglichkeiten betrifft in der Regel den Wunsch nach einer Vollduplex-Kommunikation. Auch wenn diese Entwicklung möglich wäre, löst sie aber nicht das Problem des „fehlenden Augenkontakts“ und damit des Sprecherwechsels in einem Gespräch im Chat (Zumbach and Spraul 2005, S. 379; Beißwenger 2005, S. 85) und somit auch mit dem Chatbot. Nicht unähnlich einer Telefonkonferenz ist eine nonverbale Absprache über die Sprecherreihenfolge im Netzwerk nicht möglich, wenn nicht zu einer strikten Redefolge gegriffen werden soll. Diese Problematik ist zwar für den Chatbot zunächst weniger wichtig, bei vollem Funktionsumfang könnte eine Vollduplexkommunikation mit dem Chatbot jedoch die Eröffnung eines Chatbot-Rooms ermöglichen.

Eine Lösung der Halbduplexproblematik könnte sich durch die Verwendung eines factchat (Harnoncourt et al., 2005) ergeben, in dem die Beiträge von Benutzern zeitlich und räumlich (zweidimensional) frei positioniert werden können. Die Eingabe von Texten ist für alle Teilnehmer fast zeitgleich sichtbar und gleichzeitig ist eine Ausgabe oder Eingabe von Beiträgen durch die Zeitschiene auch diachron möglich. Für einen Chatbot erscheint diese Lösung zunächst aber nicht erforderlich, da in der Regel auf die 1:1 Kommunikation gesetzt wird, in der eine synchrone Gesprächsführung nur sehr bedingt erforderlich ist. Zudem hat die räumliche Kommunikation mit mehreren Benutzern in diesem Fall auf Grund der voraussichtlich kurzen Gesprächsdauer und der stets besonders nutzerbezogenen Informationsbedürfnisse nur einen Mehrwert als Betrachtung historischer Daten. Schließlich erfordert die Realisierung der Antworten des Chatbots in einem mehrdimensionalen Raum auf Grund der z.T. erforderlichen Rückfragen weitere Überlegungen und Tests.

Der Lebenszyklus eines Chatbots in der angedachten Form ist vermutlich kürzer als der Lebenszyklus eines traditionellen fachlexikographischen Nachschlagewerks, denn

nicht zuletzt die Sprache und das sprachliche Verhalten werden sich im Takt der schnellen Technikentwicklung weiterentwickeln (Crystal, 2006, S. 57ff.). Am Horizont erwartet man bereits Chatbots mit Spracherkennung in eben dieser z.T. übereinfachen, wenig strukturgebundenen und fast immer informellen Sprache des Chats. Chats mit mikrogenerierenden, im 3D-Raum beweglichen Avataren sind bereits belegt und untersucht (Cassel, 2003; Drucker et al., 2000); auch das wäre eine Zukunftsperspektive für einen Chatbot, wobei das Lesen und Verstehen der Mimik und Gestik des Gegenüber (also des Nutzer-Avatars) eine besondere Herausforderung für die Entwicklung des Chatbots ist. Der Lebenszyklus eines Chatbots verlängert sich durch sein Training, d.h. durch die Pflege der Antworten auf Grund der gestellten Fragen – in inhaltlicher und sprachlicher Sicht. In ähnlicher Weise positiv verlängernd wirkt sich auf den Lebenszyklus die Erweiterung der fachlexikographischen / terminologischen Datenbasis aus.

Weitere Möglichkeiten ergeben sich durch eine Mehrsprachigkeit des Chatbots – mit all den Schwierigkeiten, die die oftmals fehlende volle Äquivalenz zwischen zwei Lexemen in L1 und L2 mit sich bringen (Svensén, 2004, S. 310ff.). Hinzu kommt der unverhältnismäßig steigende Pflegeaufwand für die Wissensbasis des Chatbots in mehreren Sprachen und evtl. mehreren Zeichensystemen.

Umfangreich in der Realisierung wäre eine syntaktische und (satz-) semantische Analyse der Benutzereingaben an den Chatbot. Sprachliche und außersprachliche Anteile der Kommunikation des Chatbots lassen sich aber auf diese Weise voraussichtlich verbessern. Der Mehrwert dieses Ansatzes für den Benutzer im Verhältnis zur Wissensbasis (fachlexikographisches Nachschlagewerk) und im Verhältnis zum Implementierungsaufwand wäre dabei noch zu bewerten.

Die Frage der konkreten technischen serverseitigen Realisierung des Chatbots und damit die Wahl von Plattform, Programmiersprache und Datenhaltung z.B. in einer relationalen oder objektorientierten Form wurde hier nicht diskutiert und ist zunächst für die allgemeine Konzeption nicht erheblich sondern projektabhängig. Fest steht, dass die Wahl der Mittel auf der Seite des Klienten immer die Grundsätze von Usability und Barrierefreiheit beachten muss. Der naheliegende Einsatz von Technologien wie z.B. Flash oder Ajax auf der Clientseite ist daher genau in ihrer Konsequenz zu überprüfen. Der Bot muss im Betrieb insgesamt einem fortlaufenden Qualitätstest im Hinblick auf Lernfähigkeit, Wissensbasis, Kontexterkenkung und Gedächtnisfunktion unterzogen werden (Vetter, 2003). Nur so lässt sich die Qualität der Chatkommunikation mit dem fachlexikographisch/terminologisch unterstützten Chatbot stetig verbessern.

### Literatur

- Abu Shawar, B. and Atwell, E. (2002). A comparison between Alice and Elizabeth chatbot systems. Research Report 2002.19, University of Leeds School of Computing.
- Abu Shawar, B. and Atwell, E. (2005). Using corpora in machine-learning chatbot systems. *International Journal of Corpus Linguistics*, 10(4):489–516.
- Arntz, R., Picht, H., and Mayer, F. (2004). *Einführung in die Terminologiearbeit*. Studien zu Sprache und Technik. Olm, Hildesheim, 5. edition.

- Artificial Intelligence Foundation (2007a). The A. L. I. C. E. Artificial Intelligence Foundation. Online unter <http://www.alicebot.org>.
- Artificial Intelligence Foundation (2007b). Gespräch mit Alice. Online unter <http://www.pandorabots.com/pandora/talk?botid=f5d922d97e345aa1>.
- Bachfeld, S., Christensen, A., and Christof, J. (2005). Automatische Auskunftsgespräche? Anforderungen an den Einsatz von Chatbots in Bibliotheken. *Zeitschrift für Bibliothekswesen und Bibliographie*, 89:205–218.
- Behindertengleichstellungsgesetz (2002). Verordnung zur Schaffung barrierefreier Informationstechnik nach dem Behindertengleichstellungsgesetz. Online unter <http://bundesrecht.juris.de/bundesrecht/bitv/gesamt.pdf>.
- Beißwenger, M. (2000). *Kommunikation in virtuellen Welten: Sprache, Text und Wirklichkeit. Eine Untersuchung zu Konzeptionalität von Kommunikationsvollzügen und zur textuellen Konstruktion der Welt in synchroner Internet-Kommunikation, exemplifiziert am Beispiel eines Webchats*. Ibidem, Stuttgart.
- Beißwenger, M. (2005). Interaktionsmanagement in Chat und Diskurs. Technologiebedingte Besonderheiten bei der Aushandlung und Realisierung kommunikativer Züge in Chat-Umgebungen. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 63–87. Ibidem, Stuttgart.
- Beißwenger, M. and Storrer, A. (2005). Chat-Szenarien für Beruf, Bildung und Medien. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 9–25. Ibidem, Stuttgart.
- Bergenholtz, H. (1995). Wodurch unterscheidet sich Fachlexikographie von Terminographie? *Lexicographica*, 11:50–59.
- Boguschewski, F. (2005). Selbstredende Homepage. Pandorabots: Wenn Software antwortet. *c't - Magazin für Computertechnik*, 12:214.
- Bradner, E., Kellogg, W. A., and Erickson, T. (1999). The adoption and use of 'babble': A field study of chat in the workplace. In *Proceedings of the Sixth European conference on Computer supported cooperative work*, pages 139–157.
- Braun, A. (2000). *Chatbots in der Kundenkommunikation*. Springer, Berlin/ Heidelberg/ New York.
- Bremmer, C. (2005). Chats im eLearning. Rollenspiele und andere didaktische Elemente in der netzgestützten Hochschullehre. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 90–100. Ibidem, Stuttgart.
- BUB (2007). Auszeichnung. Hamburg holt mit Chatbot Stella den Innovationspreis. *BUB – Forum Bibliothek und Information*, 59:20.
- Budin, G. and Wright, S. E. (1997). *Handbook of Terminology Management*, volume 1 (Basic Aspects of Terminology Management). John Benjamins, Amsterdam.
- Cadiz, J., Burkhalter, B., and Smith, M. (2000). Conversation trees and threaded chats. Online unter [http://research.microsoft.com/research/pubs/view.aspx?tr\\_id=380](http://research.microsoft.com/research/pubs/view.aspx?tr_id=380).

- Cassel, J. (2003). Mehr als nur ein nettes Gesicht: Embodied conversational interface agents. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 249–265. Springer, Berlin/ Heidelberg/ New York.
- Chen, D. and Sun, C. (2007). Realtime text chat via collaborative editing systems. Online unter <http://www.cit.gu.edu.au/david/CSCW02CEW/submissions/Chen.pdf>.
- Crystal, D. (2006). *Language and the Internet*. Cambridge University Press, Cambridge, 2nd edition.
- Cyranek, G. (1988). Menschliche Kommunikation und rechnerdialog. In Nullmeiner, E. and Rödiger, K.-H., editors, *Dialogsysteme in der Arbeitswelt*, pages 139–154. BI-Wissenschaftsverlag, Mannheim/ Wien/ Zürich.
- Darhower, M. (2002). Interactional features of synchronous computer-mediated communication in the intermediate L2 class: A sociocultural case study. *CALICO Journal*, 19(2):249–277.
- Davidson, P. (2005). This smartbot is a real smartass. *Wired*, 13(2).
- Diekmannshenke, H. (2005). Politische Kommunikation in Zeiten des Internet. Kommunikationswandel am Beispiel moderierter und unmoderierter Politik-Chats. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 119–144. Ibidem, Stuttgart.
- Döring, N. and Pöschl, S. (2005). Wissenskommunikation in themenbezogenen Online-Chats. Eine empirische Analyse von drei IRC-Channels zu Computertemen. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 145–160. Ibidem, Stuttgart.
- Drossmann, C. (2001). GermanAlice-D. Online unter <http://www.alicebot.org/downloads/aiml/GermanA1standalone.zip>.
- Drucker, S. M., Farnham, S. D., and Smith, M. A. (2000). The social life of small graphical chat spaces. In *Proceedings of CHI 2000*.
- Filinski, P. (1998). *Chatten in der CyberWorld*. VMI Buch AG, Bonn.
- Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg (2007). Zukunftsorientierte Waldwirtschaft. Online unter <http://www.zukunftswald.de/page.php>.
- Geeb, F. (1998). *Semantische und enzyklopädische Informationen in Fachwörterbüchern. Eine Untersuchung zu fachinformativen Informationstypen mit besonderer Berücksichtigung wortgebundener Darstellungsformen*. Wirtschaftsuniversität Aarhus, Århus.
- Geeb, F. (2003). Lexikographische Informationsstrukturierung mit XML. *Information Wissenschaft & Praxis*, 7:415–421.
- Geeb, F. (2006). lookedup4you – studerende, kunder og online-leksikografi. In Hansen, C., Lorientzen, H., and Trap-Jensen, L., editors, *Nordiske Studier i Leksikografi 8. Rapport fra Konference om Leksikografi i Norden*.
- Geeb, F. and Spree, U. (2005). lookedup4you – mikrostruktur und makrostruktur und all das. produktion eines online-nachschlagewerks als studentisches projekt. ein erfahrungsbericht. *Information – Wissenschaft & Praxis*, 3:133–142.

- Harnoncourt, M., Holzhauser, A., Seethaler, U., and Meinel, P. (2005). Referenzierbarkeit als Schlüssel zum effizienten Chat. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 160–179. Ibidem, Stuttgart.
- Hentschläger, U. and Wiener, Z. (2002). *Webdramaturgie. Das audio-visuelle Gesamterlebnis – 3D, Streaming, Flash*. Markt+Technik, München.
- Holmer, T. and Wessner, M. (2005). Gestaltung von Chat-Werkzeugen zur Verringerung der Inkohärenz. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 181–199. Ibidem, Stuttgart.
- Holzhauser, A. (2003). Chat-Technologien in der Arbeitswelt. Online unter <http://chat.themenplattform.com/126558/0/>.
- Husmann, H. (1998). *Chatten im Internet Relay Chat (IRC). Einführung und erste Analyse*. KoPäd, München.
- Jia, J. (2007). The study of the application of a keywords-based chatbot system on the teaching of foreign languages. Online unter <http://arxiv.org/abs/cs.CY/0310018/>.
- Kollock, P. and Smith, M. (1998). What do people do in virtual worlds, an analysis of v-chat log file data. Online unter <http://research.microsoft.com/scg/papers%5Ckollockv-chat.pdf>.
- Licklider, J. and Taylor, R. (1968). The computer as a communication device. *Science and Technology*, pages 21–31.
- Lindner, C. (2003). Wer braucht wofür Avatare? In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 5–24. Springer, Berlin/ Heidelberg/ New York.
- Medienpädagogischer Forschungsverbund Südwest (2007a). JIM-Studie 2006. Jugend, Information, (Multi-) Media. Online unter [http://www.mpfs.de/fileadmin/JIM-pdf06/JIM-Studie\\_2006.pdf](http://www.mpfs.de/fileadmin/JIM-pdf06/JIM-Studie_2006.pdf).
- Medienpädagogischer Forschungsverbund Südwest (2007b). KIM-Studie 2006. Kinder + Medien, Computer + Internet. Online unter <http://www.mpfs.de/fileadmin/KIM-pdf06/KIM2006.pdf>.
- Morphy, E. (2001). Jeeves exec explains self-service evolution. Online unter <http://www.crm-daily.com/perl/story/18064.html>.
- MySQL AB (2007). MySQL 5.1 Reference Manual. 12.8 Full-Text Search Functions. Online unter <http://dev.mysql.com/doc/refman/5.1/en/fulltext-search.html>.
- Nake, F. (1988). Dialogisieren mit dem computer - anmerkungen zu entwicklung, begriff und technik der dialogsysteme. In Nullmeiner, E. and Rödiger, K.-H., editors, *Dialogsysteme in der Arbeitswelt*, pages 16–46. BI-Wissenschaftsverlag, Mannheim/ Wien/ Zürich.
- Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann, San Francisco.
- Nielsen, J. (2001). *Designing Web Usability*. Markt+Technik, München.
- Niemann, F. (2003). Instant Messaging: Arbeitssitzung im Chat-Room. Online unter <http://www.computerwoche.de/index.cfm?pid=816&pk=535484>.

- Puck, J. F. and Exter, A. (2005). Der Einsatz von Chats im Rahmen der Personalbeschaffung. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 289–301. Ibidem, Stuttgart.
- Reeves, B. and Nass, C. (1996). *The media equation: how people treat computers, television and new media like real people and places*. University of Chicago Press.
- Reichle, M. (2006). Entwicklung eines prototypischen Chatbots für die Universitätsbibliothek Hildesheim. In Womser-Hacker, C. and Mandl, T., editors, *Effektive Information Retrieval Verfahren in Theorie und Praxis: Proceedings Fünfter Hildesheimer Evaluierungs- und Retrieval (HIER) Workshop*, pages 43–46.
- Runkehl, J., Schlobinski, P., and Siever, T. (1998). *Sprache und Kommunikation im Internet*. Vandenhoeck & Ruprecht, Opladen.
- Salimi, M. (2003). Support-Chat und Avatare. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 131–138. Springer, Berlin/ Heidelberg/ New York.
- Samuelsen, P. (2003). Die Bedeutung von natürlichsprachlichen Dialogsystemen im Internet-Business. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 27–35. Springer, Berlin/ Heidelberg/ New York.
- Schmitt, A. A. (1983). *Dialogsysteme. Kommunikative Schnittstellen, Software-Ergonomie und Systemgestaltung*, volume 40 of *Informatik*. Bibliographisches Institut, Mannheim/ Wien/ Zürich.
- Shneidermann, B. (1997). Direct manipulation versus agents: Paths to predictable, controllable and comprehensive interfaces. In Bradshaw, J., editor, *Software Agents*, pages 97–106. MIT Press.
- Shuli, G. and Nielsen, J. (2002). Usability of websites for children: 70 design guidelines based on usability studies with kids. Usability report, Nielsen Norman Group, Freemont.
- Staats- und Universitätsbibliothek (2005). Chatbot Stella. Online unter <http://www.sub.uni-hamburg.de/informationen/projekte/infoass.html>.
- Stein, D. (2005). Distanz und Nähe in interaktiver Internetkommunikation. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 289–301. Ibidem, 339-347.
- Storrer, A. (2001). Sprachliche Besonderheiten getippter Gespräche: Sprecherwechsel und sprachliches Zeigen in der Chat-Kommunikation. In Beißwenger, M., editor, *Chat-Kommunikation. Sprache, Interaktion, Sozialität und Identität in synchroner computervermittelter Kommunikation*, pages 3–24. Ibidem, Stuttgart.
- Svensén, B. (2004). *Handbok i leksikografi*. Norstedts Akademiska Förlag, Stockholm.
- Thommes, J. (2001). Die Chatsetter. *HORIZONT*, 19:78.
- Trogemann, G. (2003). Mit Hand und Fuß. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 269–290. Springer, Berlin/ Heidelberg/ New York.
- Turney, P. D. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, 21(4):315–346.

- van Eckert, E. (2005). Termingebundene Chats one-to-one in der psychosozialen Beratung. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 349–359. Ibidem, Stuttgart.
- Vetter, M. (2003). Qualitätskriterien für Online-Tests von Bots. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 73–92. Springer, Berlin/ Heidelberg/ New York.
- von Wendt, K.-L. (2003). Technische Grundlagen von natürlichsprachlichen Dialogsystemen. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 39–47. Springer, Berlin/ Heidelberg/ New York.
- Vronay, D., Smith, M., and Drucker, S. (2004). Alternative interfaces for chat. chat as a streaming media data type. Online unter <http://research.microsoft.com/scg/papers/chat.htm>.
- Wallace, R. S. (2002). AIML overview. Online unter <http://www.pandorabots.com/pandora/pics/wallaceaimltutorial.html>.
- Weizenbaum, J. (1966). ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 10(8):36–45.
- Wiegand, H. E. (1979). Definition und Terminologienormung – Kritik und Vorschläge. In Felber, H., Lang, F., and Wersig, G., editors, *Terminologie als angewandte Sprachwissenschaft. Gedenkschrift für E. Wiüster*, pages 101–148. Saur, München/ New York/ London/ Paris.
- Willand, I. (2002). *Chatroom statt Marktplatz. Identität und Kommunikation zwischen Öffentlichkeit und Privatheit*. KoPäd, München.
- Wirth, T. (2003). Avatare und die Usability von Websites. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 123–128. Springer, Berlin/ Heidelberg/ New York.
- Wolff, P. (2003). Virtualisierung und Personalisierung – Technologietrends machen Avatare zur innovativen Mensch-Maschine-Schnittstelle. In Lindner, C., editor, *Avatare. Digitale Sprecher für Business und Marketing*, pages 293–299. Springer, Berlin/ Heidelberg/ New York.
- Zumbach, J. and Spraul, P. (2005). Tutoring in synchronen Lernumgebungen. In Beißwenger, M. and Storrer, A., editors, *Chat-Kommunikation in Beruf, Bildung und Medien. Konzepte, Werkzeuge, Anwendungsfelder*, pages 379–391. Ibidem, Stuttgart.