

Projekt-Abschlussbericht

Digitale Assistenz

in der psychosozialen Beratung

09.2019 – 02.2021

Gefördert von



STAEDTLER
STIFTUNG

Durchführung:

Prof. Dr. Jens Albrecht (Fakultät Informatik)

in Kooperation mit

Prof. Dr. Robert Lehmann (Fakultät Sozialwissenschaften)

Sigrid Zauter (Institut für E-Beratung)

Studierende: Nigar Aufotu (MA Inf.), Markus Glas (MA Inf.), Florian Naumann (BA Inf.),
Jonas Noack (BA Inf.), Ronja Regler (BA SW), Mathias Rühl (BA Inf.), Runa Schulze (BA SW)

Technische Hochschule Nürnberg Georg Simon Ohm

Keßlerplatz 12, 90489 Nürnberg



Zusammenfassung

In vielen Unternehmen werden heutzutage Chatbots eingesetzt, um Standard-Anfragen von Kunden effizient zu beantworten. Auch im psychosozialen Beratungsumfeld können natürlichsprachliche digitale Assistenzen eine sinnvolle Ergänzung zu menschlichen Beratern sein, denn sie sind jederzeit verfügbare, anonyme Ansprechpartner. Die Hürden für den Einsatz in der psychosozialen Beratung sind allerdings hoch, da die Entwicklung und der Betrieb von Chatbots zunächst Kosten verursacht und systematische Studien zum Mehrwert derartiger Systeme in diesem Anwendungsfeld bisher nicht vorliegen.

Ziel des Projekts DIA (Digitale Assistenz) war die prototypische Entwicklung eines lernfähigen Chatbots für die psychosoziale Beratung. Dabei wurde insbesondere die Frage untersucht, wie aufwändig es ist, einen KI-basierten Chatbot ausschließlich mit Open-Source-Technologie zu entwickeln. Das von der Staedtler-Stiftung geförderte Projekt lief von September 2019 bis Februar 2021. Die Arbeiten wurden in einer interdisziplinären Kooperation vom Institut für E-Beratung mit der Fakultät Informatik der TH Nürnberg durchgeführt. Als Anwendungsdomäne diente die Bildungsberatung für den Zertifikatskurs Online-Beratung, die bisher hauptsächlich über E-Mail-Kommunikation erfolgt.

Inhalt

1	Einleitung	3
1.1	Ziele des Projektes	3
1.2	Ausgangslage.....	3
1.3	Forschungsfragen.....	5
2	Anforderungsanalyse	6
2.1	Anwendungsszenarien	6
2.1.1	Bildungsberatung für den Zertifikatskurs "Onlineberatung"	6
2.1.2	Zweite Domäne.....	6
2.2	Anforderungen an DIA.....	7
2.2.1	Anforderungen an das Conversational User Interface	7
2.2.2	Anforderungen für die Wartbarkeit der Wissensbasis.....	8
3	Projekttablauf	9
4	System-Überblick	10
4.1	Einbindung des Chatbots auf der Web-Seite.....	10
4.2	Architektur des Chatbots	10
4.3	Dialog-Design	12
4.3.1	Gestaltung der Wissensbasis für Faktenfragen zum Zertifikatskurs	12
4.3.2	Domänen-unabhängige Antworten	14

4.4	Semi-automatisierte Informationsextraktion für den Aufbau von Knowledge Graphen	15
4.5	Semi-automatisierte Erstellung und Pflege der unstrukturierten Wissensdatenbank für das Question-Answering-Modul	17
4.6	Probleme bei der Entwicklung und Bewertung der Architektur	19
5	Feldtest und Akzeptanzstudie	20
5.1	Hintergrund	20
5.2	Fragestellung	20
5.2.1	Allgemeine Funktionsweise des Chatbots	20
5.2.2	Akzeptanz des Chatbots bei der Zielgruppe	21
5.3	Methode	21
5.3.1	Vorgehen	21
5.3.2	Entwicklung der Erhebungsinstrumente	22
5.4	Ergebnisse	23
5.4.1	Deskriptive Daten zur Stichprobe	23
5.4.2	Interaktionsqualität	24
5.4.3	Charakteristika der Gesprächsverläufe	27
5.4.4	Darstellung der Antworten auf die offenen Fragen	28
5.4.5	Vergleich mit der Akzeptanzstudie von Denk (2019)	30
5.5	Diskussion der Ergebnisse aus sozialwissenschaftlicher Perspektive	31
5.6	Auswertungen mit PowerBI	32
5.7	Technische Verbesserungsbedarfe auf Basis der Untersuchung	34
6	Projektergebnisse	37
6.1	Studentische Arbeiten	37
6.2	Veröffentlichung	37
6.3	Weitere Forschungsprojekte	37
7	Zusammenfassung und Ausblick	39
	Literaturverzeichnis	41
	Anhang	43
	Anhang 1: Bildung der Untersuchungsskalen.	43
	Anhang 2: Mittelwertsvergleiche der Akzeptanzstudien	46

1 Einleitung

Digitale natürlichsprachliche Assistenten (Chatbots) werden heutzutage verstärkt in Unternehmen eingesetzt, um die Kundenkommunikation effizienter zu gestalten. Insbesondere im Service-Bereich werden digitale Assistenzen genutzt, um mit Hilfe einer zugrundeliegenden Wissensbasis spezifische Fragen von Kund*innen bzw. Nutzer*innen zu beantworten. Häufige Fragen lassen sich dadurch im besten Fall ohne menschliche Berater*in im 24x7-Modus abdecken. Genau das macht den Einsatz von Chatbots auch im psychosozialen Bereich interessant, wo der individuelle Beratungsbedarf oft hoch, die finanziellen Mittel für Beratung jedoch begrenzt sind.

1.1 Ziele des Projektes

Ziel des Vorhabens war die Entwicklung einer digitalen Assistenz in Form des lernfähigen Chatbots "DIA" in einer interdisziplinären Kooperation mit der Fakultät Informatik und dem Institut für E-Beratung der TH Nürnberg. Als konkrete Versuchsfelder waren die Bildungsberatung des Instituts im Kontext des Zertifikatskurses Online-Beratung sowie die Beratung pflegender Angehöriger geplant. DIA sollte die Aufgabe übernehmen, Fragen in natürlicher Sprache mit sinnvollen Antwortangeboten zu begegnen. Dabei sollte untersucht werden, inwiefern moderne Chatbot-Technologie unterstützt durch Methoden der künstlichen Intelligenz für die persönliche Beratung im genannten Umfeld sinnvoll eingesetzt werden kann. So sollte DIA im Gegensatz zu den meisten derzeit eingesetzten Systemen nicht nur auf vordefinierte Fragen antworten können, sondern selbständig Antworten aus der vorhandenen Dokumentation und den Erfahrungen anderer Ratsuchender erlernen. Auch mit ungenau spezifizierten Anfragen sollte DIA umgehen können. Dafür galt es, aktuelle Ansätze aus der Forschung wirkungsvoll zu kombinieren und für die konkrete Problemstellung zu adaptieren.

Da der Schutz der persönlichen Daten hierbei besonders wichtig ist, können die leistungsfähigen Cloud-basierten Systeme global agierender Konzerne (z.B. Google Dialogflow oder Amazon Lex) nicht ohne Weiteres eingesetzt werden. Eine wichtige Fragestellung und Teil der Untersuchung war daher, ob und wie das Projekt vollständig mit **Open-Source-Technologie** umgesetzt und als Blaupause für weitere Projekte verwendet werden kann.

Das System sollte möglichst generisch konzipiert werden, so dass es in unterschiedlichsten Beratungsszenarien eingesetzt werden kann. Im Rahmen dieses Projektes diente als erstes Szenario ein Zertifikatskurs für Studierende, ein zweites Szenario sollte folgen. In nächsten Schritten sollte das Konzept auch auf andere Szenarien ausgeweitet werden können, wie beispielsweise für Incomings oder Outgoings in der Studienberatung der TH Nürnberg.

1.2 Ausgangslage

Das Institut für E-Beratung unterstützt öffentliche Beratungsinstitute beim Aufbau digitaler Angebote für die psychosoziale Beratung. Dabei geht es vor allem um die Diagnostik und/oder Linderung psychosozialer Belastungen. Darüber hinaus gibt es ein eigenes Bildungsangebot im Kontext des Hochschulzertifikats „Onlineberatung“ (Gröning 2011).

Die Themen psychosozialer Beratung sind traditionell sehr sensibel und schützenswert gegenüber wirtschaftlichen Interessen. Davon unbesehen werden gerade auch Gesundheitsinformationen weitgehend im Internet gesucht (Medienpädagogischer Forschungsverbund Südwest (mpfs) 2017). Für fast die Hälfte aller Deutschen (44%) ist das Internet im Bedarfsfall bei psychosozialen Problemen eine Anlaufstelle (Eichenberg und Brähler 2013). Das Internet im Allgemeinen sowie psychosoziale Onlineberatung im Besonderen bietet Ratsuchenden demnach die Möglichkeit, sich im Schutze der Anonymität über sensible Themen auszutauschen. Dabei übersteigt die Nachfrage im Netz das durch seriöse Beratungsanbieter vorgehaltene Onlineberatungsangebot. Oftmals fehlt es im Netz zudem an Clearingstellen, durch die das Grundanliegen und die passende Hilfeform für Ratsuchende ermittelt werden könnte (Wenzel 2013). Häufig leiten Webseiten immer noch zuallererst mit Kontaktdaten wie Telefonnummern oder Öffnungszeiten auf persönliche Kontakte. Die Chancen der digitalen Beratung werden nicht ausgeschöpft.

Aus verschiedensten Forschungsergebnissen ist bekannt, dass sich mit der Etablierung des Internets sowie der Smartphone-Technologie die Lesegewohnheiten geändert haben (Maurer 2015). Die Bereitschaft lange Texte zu lesen ist gesunken (Loh und Kanai 2016). Die Aufbereitung wesentlicher Inhalte in kleinen Häppchen wird mehr und mehr üblich und schränkt die Aufnahmefähigkeit stark ein (DelVecchio et al. 2019). Bei vertieftem Interesse werden sog. Links angeboten, die zu ausführlicheren Erläuterungen führen. Dies passt sehr zielgenau zum Untersuchungsgegenstand, da die davorliegende Kommunikation mit einem Chatbot bereits für Eingrenzung sorgt.

Natürlichsprachliche digitale Assistenzen können gerade im psychosozialen Beratungsumfeld eine sehr sinnvolle Ergänzung zu menschlichen Beratern sein. Sie zeichnen sich vor allem dadurch aus, dass sie jederzeit, direkt und sofort antworten. Weiterhin haben sie das Potenzial, die Hemmschwelle für peinliche oder als dumm empfundene Fragen zu senken.

Allerdings können Chatbots den persönlichen Kontakt nicht ersetzen, sie haben kein Mitgefühl und können sich nicht in das Gegenüber hineinversetzen. Dennoch muss ein Chatbot mit einer Persönlichkeit bzw. einem Charakter ausgestattet werden, der auf die Beratungssituation abgestimmt ist und den Nutzenden ein angenehmes Gefühl gibt (Kaiser et al. 2019; McTear et al. 2016; Shevat 2017). Dabei muss jedoch jederzeit klar ersichtlich sein, dass kein Mensch, sondern eine Maschine antwortet.

Es geht also niemals darum, menschliche Beratung komplett durch einen Chatbot zu ersetzen, sondern vielmehr schnelle Antworten auf Standardfragen zu liefern. Ratsuchende bekommen so die Möglichkeit, Zeit-, Orts- und Personen-unabhängig Fragen zu stellen und schnelle Antworten zu erhalten. Gleichzeitig werden menschliche Berater entlastet und haben so mehr Zeit für ausführliche und individuelle Beratung.

Die hier untersuchte digitale Assistenz, die 24 Stunden jeden Tag der Woche für Ansprache zur Verfügung steht, erscheint zeitgemäß und bedarfsorientiert. Mit kurzgehaltenen Antworten, angereichert mit Vorschlägen für ein Weiterblättern, kann sie basierend auf unterschiedlichen Informationsquellen und Medien ansprechend Antworten und durch große, schwer zu durchschauende Wissensgebiete lotsen.

1.3 Forschungsfragen

Das Projekt dient der Untersuchung der Potenziale und Grenzen für den Einsatz von Chatbots in der psychosozialen Beratung. Konkret sollten folgende Forschungsfragen im Laufe des Projektes beantwortet werden:

Wie ist eine digitale Assistenz im Umfeld öffentlicher Beratungsaufträge zu entwickeln, damit sie Akzeptanz findet und von Ratsuchenden genutzt wird?

Wesentliche Kriterien für die Akzeptanz eines Chatbots gegenüber einer einfachen Web-Suche sind die leichte und intuitive Bedienbarkeit des Dialoges sowie ein angenehmer Charakter der digitalen Assistenz (Kaiser et al. 2019; McTear et al. 2016; Shevat 2017). Im Bereich dieser Forschungsfrage wurden Messinstrumente zusammengestellt und entwickelt, die helfen sollen, diese Fragen zu beantworten und eine ständige Optimierung im sich schnell entwickelnden digitalen Umfeld ermöglichen.

Wie aufwändig ist es, mit moderner Open-Source-Technologie einen leistungsfähigen Chatbot zu entwickeln?

Aus Gründen des Datenschutzes sollte der Einsatz Cloud-basierter Software in der psychosozialen Beratung vermieden werden. Daher wurde das Projekt als On-Premise-Lösung basierend auf dem Open-Source-Framework Rasa¹ entwickelt. Die Erfahrungen aus dem DIA-Projekt ermöglichen eine Abschätzung der Aufwände für zukünftige Projekte.

Wie kann basierend auf aktuellen Forschungsergebnissen aus den Bereichen Natural Language Processing und künstliche Intelligenz ein lernfähiges natürlichsprachliches Assistenzsystem im Beratungsumfeld gestaltet werden?

Die Verwaltung und Pflege der Wissensbasis bei Chatbots ist häufig mit hohem manuellem Aufwand verbunden. Daher wurde in dem Projekt untersucht, wie einfach oder aufwändig Wissen aus vorhandenen Datenbeständen extrahiert und in ein Frage-Antwort-System integriert werden kann. Dabei wurden aktuelle Verfahren und Sprach-Modelle für die Wissensextraktion und die Beantwortung von Fragen eingesetzt.

Wie können die Vertrauenswürdigkeit bzw. Zuverlässigkeit von Antworten bestimmt und mit ausgewiesen werden?

Insbesondere bei der persönlichen Beratung ist es wichtig, dass keine falschen oder irreführenden Informationen generiert werden. Daher sollten Wege gefunden werden, die Validität der Antworten sicherzustellen und gegebenenfalls Unwissen explizit zu äußern.

¹ <https://rasa.com>

2 Anforderungsanalyse

2.1 Anwendungsszenarien

2.1.1 Bildungsberatung für den Zertifikatskurs "Onlineberatung"

Der Chatbot soll die Aufgabe übernehmen, interessierte Studierende durch organisatorische Fragen bezüglich des an der TH Nürnberg angebotenen "Zertifikatskurs Onlineberatung" zu lotsen.

Die Beratung der Studierenden für den Zertifikatskurs findet aktuell überwiegend per E-Mail statt. Aus vielen Jahren der studentischen Beratung sind die am häufigsten gestellten Fragen bekannt. Die Antworten auf diese Standardfragen sind in Flyern und FAQ-Listen auf der Webseite veröffentlicht.

Die Struktur des Zertifikatskurses besteht aus 3 Modulen und einem Praxisteil, die aufeinander aufbauen und in zeitlicher Reihenfolge zu belegen sind. Der Zertifikatskurs ist somit komplex und birgt ein hohes Nachfragepotential. Typische Fragen sind:

- Welche Module gehören zum Zertifikatskurs Onlineberatung?
- Welche Anwesenheitspflicht habe ich für den Zertifikatskurs Onlineberatung?
- Kann ich den Zertifikatskurs Onlineberatung nach meinem Studienabschluss machen?
- Derzeit befinde ich mich im Praxissemester, wollte jedoch anfragen, ob es schon möglich wäre, dieses Semester den Workshop 'Formen und Methoden sozialpädagogischer Beratung' zu besuchen. (Also vor Beendigung des praktischen Studienseesters?)
- Wie läuft die Fallpraxis ab?
- Kann ich mir Module des Kurses auch für mein Studium anrechnen lassen?

Diese Fragen sind zwar klar zu beantworten, sie werden aber oft unklar formuliert und beinhalten neben impliziten (nur den Kurs betreffend) auch explizite (nicht den Kurs betreffende) Bezüge. Es soll im Zuge des Projektes untersucht werden, inwiefern mit Hilfe künstlicher Intelligenz dahinterliegende Fragestellungen erahnt und korrekt beantwortet werden können.

2.1.2 Zweite Domäne

Die ursprünglich vorgesehene Domäne "Beratung für pflegende Angehörige" am Institut für E-Beratung ließ sich nicht erreichen. Die Onlineberatung für pflegende Angehörige wurde wider Erwarten nicht durch einen Träger der Angehörigenberatung übernommen und weitergeführt und hörte damit auf zu existieren. Zum anderen erforderten einige Teilaufgaben im Projekt einen deutlich höheren Zeitaufwand als ursprünglich vorgesehen. Daher wurden die Prioritäten zu Gunsten der Lösung vielfältiger technischer Probleme beim Aufsetzen und Training des Chatbots verschoben, um hier einen lauffähigen Prototypen bereitstellen und testen zu können.

Ersatzweise entstand im Rahmen einer Masterarbeit eine konzeptionelle Planung eines Knowledge-Graphen für einen Chatbot, der durch die Informations- und Beratungsseite des Projektes TriN führen

kann. Daraus entstanden weitere Hintergrund-Erkenntnisse, die in die Gesamtbetrachtung dieses Projektes einfließen. Trialog im Netz (TriN) ist eine Onlineberatungsplattform an der Robert Kümmerl Akademie in Würzburg. Die Projektidee Trialog im Netz (TriN) will den Prozess der Implementierung des neuen Berufsbildes „Genesungsbegleiter“ unterstützen, mittels eines zu entwickelnden digital gestützten Informations- und Unterstützungsangebots mit interaktiven Beratungsangeboten, das deutschlandweit Unternehmen wie Arbeitnehmer*innen zur Verfügung steht. Genesungsbegleiter*innen sind Menschen, die schwere psychische Krisen überwunden haben und nach einer Qualifizierung für Unternehmen der psychiatrischen Versorgung als „Experten aus Erfahrung“ bei ihrer Arbeit unterstützen können.

2.2 Anforderungen an DIA

Nachdem ein Chatbot letzten Endes ein User Interface für eine Wissensbasis ist, lassen sich die Anforderungen in zwei Teilbereiche aufteilen: Einerseits Anforderungen an die Kommunikationsschnittstelle und andererseits Anforderungen an die Wartbarkeit der Wissensbasis.

2.2.1 Anforderungen an das Conversational User Interface

Ein Chatbot muss mit verschiedenen Fragetypen umgehen können. Nachfolgend sind typische Beispiele für Fragetypen und die erwünschten Reaktionen von DIA aufgeführt:

- A.1) DIA soll Informations- und Faktenfragen als solche erkennen und möglichst mit dem entsprechenden Faktenwissen antworten:

Beispiel: Welche Anwesenheitspflicht habe ich? -> Es besteht in den Lehrveranstaltungen eine Anwesenheitspflicht von 75%

- A.2) DIA soll bei Verständnisfragen und unscharfen Formulierungen Informationen anbieten und gezielt nachfragen:

Beispiel: Wie viele Module muss ich belegen? -> Es gibt 4 Module, willst Du sie sehen?

Beispiel: Kann ich den Kurs auch nach dem Studium machen? -> Der Kursbeginn muss während des Studiums sein und der Kurs muss bis zu einem Jahr nach Ende des Studiums abgeschlossen werden.

Beispiel: Ich bin schwanger, schaffe ich den Kurs? ->

Bei Ausnahmeregelungen oder die persönliche Situation soll DIA an die (menschliche) Beratung verweisen.

- A.3) DIA soll durch den hinterlegten Wissensbereich im Beratungsfeld führen:

Im Beispiel Zertifikatskurs sind dies die Kursbeschreibung sowie FAQs. Sofern der Chatbot bei der Antwort nicht sicher ist, soll er nach Antwortalternativen fragen und ggf. auf die entsprechenden Informationen im Intra- und Internet verweisen.

- A.4) DIA soll Servicefunktionen übernehmen:
Für den Zertifikatskurs kann dies die Anmeldung, die Bekanntgabe von Räumen sowie Terminerinnerungen sein.
- A.5) DIA soll höflich und zuvorkommend sein.
DIA soll höflich und sachlich kommunizieren. Sie soll den oder die Kund*in freundlich begrüßen und verabschieden. DIA kann fragen, ob eine Serviceleistung gewünscht wird und diese dann auch übernehmen.
- A.6) DIA soll eine Feedback-Möglichkeit bieten.
Durch die Implementierung einer Feedback-Frage am Ende einer Beratungskonversation ("Wie zufrieden sind Sie mit der Beratung?") wird die systematische Analyse von Wissenslücken bzw. erfolgreichen Dialogmustern ermöglicht.

2.2.2 Anforderungen für die Wartbarkeit der Wissensbasis

Ein Chatbot wird nur dann akzeptiert, wenn die Antworten hilfreich, korrekt und zeitlich aktuell sind. Eine Wissensbasis wird benötigt, die gepflegt werden muss. Dafür soll das System Funktionen bereitstellen, welche die Verwaltung der vorhandenen Wissensbasis einerseits und das Erkennen von Wissenslücken andererseits ermöglichen.

- A.7) Der Aufbau der Wissensbasis soll durch das System unterstützt werden. Faktenwissen zu den wichtigsten Objekten der Anwendungsdomäne soll in einer Wissensbasis bereitgestellt werden. Beim Zertifikatskurs sind das bspw. Informationen zu den Modulen, zur Anmeldung usw. Das dafür notwendige Wissen soll system-unterstützt aus bestehenden Dokumenten (Webseiten, FAQs) extrahiert und strukturiert abgelegt werden.
- A.8) Das System soll Möglichkeiten zur kontinuierlichen Verbesserung der Inhalte bieten. Insbesondere soll eine Möglichkeit geschaffen werden, nicht-beantwortete Fragen zu erkennen und die entsprechenden Antworten einfach einzupflegen. Dafür sollen Antwortvorschläge zu unbeantworteten Fragen aus bestehenden Dokumenten (z.B. E-Mail-Kommunikation) generiert und den Redakteuren vorgelegt werden.

3 Projektablauf

In der ersten Projektphase war die technische Entwicklung des Dialogsystems mit einem Forschungsprojekt im Rahmen des Studiengangs Master of Applied Research verknüpft. Nach erfolgreichem Studienabschluss führten studentische Hilfskräfte mit unterschiedlichen Aufgaben und Abschlussarbeiten die begonnene Arbeit fort. Fachlich wurde das Projekt von den Professoren und Mitarbeiter*innen des Instituts für Online-Beratung begleitet. Sie erarbeiteten in enger Kooperation mit der Software-Entwicklung das fachliche Konzept, die Dialogstruktur, Akzeptanzkriterien und Testfälle. Darüber hinaus wurden Akzeptanztests mit Nutzer*innen durchgeführt. Das Projekt wurde in drei Phasen bearbeitet, die sich an den Semestern orientierten.

- **Projektphase 1: Grundlegende Konzeption, Entwicklung eines Prototyps für das erste Anwendungsszenario**

Dem Grundgedanken agiler Software folgend war das Ziel der ersten Projektphase die Entwicklung eines "Minimum Viable Product". Darunter wird eine erste, einsetzbare Version einer Software verstanden, die es ermöglichen soll, möglichst früh Feedback durch Anwender zu erhalten, um dieses in die Weiterentwicklung des Produktes einfließen zu lassen. In dieser Projektphase entstand basierend auf Open-Source-Technologie ein einfacher Chatbot, der auf (inhaltlich, aber nicht wörtlich) vordefinierte Fragestellungen festgelegte Antworten gibt. Grundlage sind bspw. die FAQ für den Zertifikatskurs Online-Beratung, aber auch anonymisierte Mailkommunikation sowie die Unterstützung der Sachbearbeiter*innen. Die Phase endete mit der Inbetriebnahme des ersten Prototyps.

- **Projektphase 2: Funktionale Erweiterung des Prototyps, Einbindung des zweiten Anwendungsszenarios**

Mit dem Prototyp erfolgten wie geplant Akzeptanzstudien und die Optimierung der Funktionsweise innerhalb der ersten Domäne sowie die technischen Dokumentation der Arbeitsschritte. Dabei wurde ein grundlegendes Verständnis von den Möglichkeiten und Grenzen von Chatbots für Fragestellungen aus dem Bereich der Psychosozialen Arbeit erlangt, das bisher so nicht vorlag. Der Transfer in andere Domänen blieb ein wichtiges Ziel, das im Rahmen der Projektlaufzeit allerdings nur konzeptuell bearbeitet wurde.

- **Projektphase 3: Akzeptanztest und Projektabschluss**

In der letzten Projektphase standen abschließende Akzeptanztests und deren Auswertung im Vordergrund.

4 System-Überblick

4.1 Einbindung des Chatbots auf der Web-Seite

Das Institut für E-Beratung unterhält in seinem Webauftritt einen Bereich "Für Studierende", auf dem diverse Informationen zum Zertifikatskurs abrufbar sind. Hier soll auch der Prototyp des Chatbots verfügbar gemacht werden.

Die Webseite ist mit dem Content Management System Wordpress implementiert. Bei der Einbindung des JavaScript-basierten Chatbots waren etliche technische Probleme zu lösen, die zu ungeplant hohen zeitlichen Aufwänden geführt haben. Aktuell ist der Chatbot unter der URL <http://www-e-beratungsinstitut.de/dia> erreichbar. Abbildung 1 zeigt, wie er sich auf der Webseite meldet.

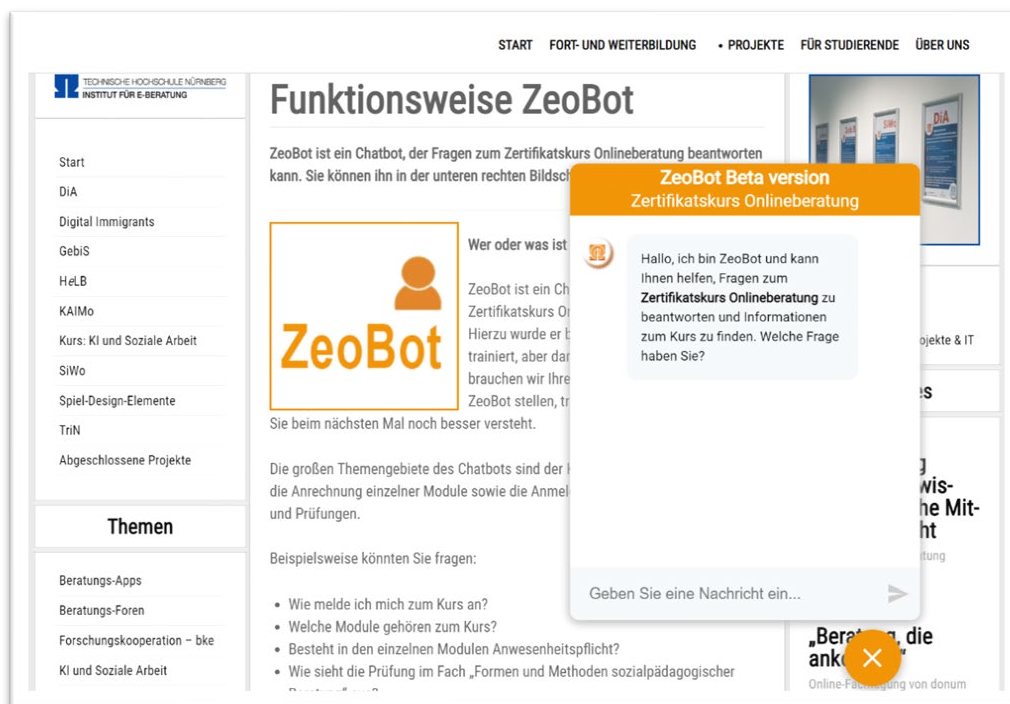


Abbildung 1 Einbindung des Chatbots auf der Webseite des Zertifikatskurses

4.2 Architektur des Chatbots

Ein wesentliches Design-Ziel beim Entwurf des Chatbots war die Realisierung eines möglichst guten Kompromisses zwischen folgenden Anforderungen:

1. Faktische Fragen zu wesentlichen Entitäten der Anwendungsdomäne sollen exakte Antworten liefern.
2. Die Wissensbasis des Chatbots soll sich mit möglichst wenig Aufwand pflegen lassen.

Für die erste Anforderung ist eine Wissensbasis erforderlich, in welcher das Faktenwissen abgelegt ist. Darüber hinaus müssen die Entitäten der Wissensbasis und deren Attribute präzise in den Nutzerfragen erkannt werden, was spezifische Trainingsdaten erfordert. Beides, die Wissensbasis und die Trainingsdaten, müssen manuell gepflegt und bei Änderungen aktualisiert werden. Um gleichzeitig die zweite Anforderung umzusetzen, wurde neben der Wissensbasis ein Question-Answering-Modul basierend auf einer Dokumentensuche implementiert. Dieses Modul sucht mithilfe semantischer Suchverfahren passende Dokumente. Abbildung 2 gibt einen Überblick über die Bestandteile des Chatbots.

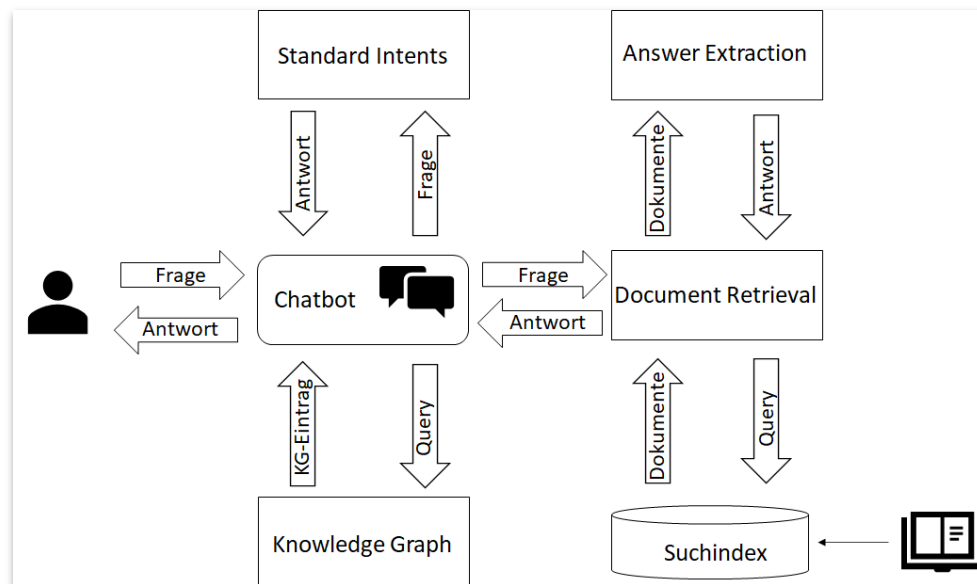


Abbildung 2 Architektur des Chatbots (Noack, 2021)

Wird vom Nutzenden eine Frage an das System gestellt, so wird sie zunächst von der NLU-Komponente (Natural Language Understanding) entgegengenommen. Aufgabe der NLU-Komponente ist das Erkennen der Nutzerabsicht (des Intents) und ggf. der beteiligten Entitäten und Attribute. Basierend darauf wird entweder eine Standard-Antwort zurückgegeben (z.B. bei einer Begrüßung) oder es wird eine Suche nach der Antwort ausgelöst.

Sofern eine Frage nach bekanntem Faktenwissen erkannt wird (z.B. "Wann beginnt der Kurs?"), werden die entsprechenden Informationen (Attribut "Beginn" zu Entität "Kurs") aus der Wissensbasis, hier als Knowledge Graph dargestellt, extrahiert und als Antwort zurückgegeben. Sofern die Antwort nicht eindeutig bestimmt werden kann, fragt der Chatbot nach weiteren Details (z.B. "Welcher Kurs?").

Werden keine Entitäten aus der Wissensbasis erkannt oder ist die Antwort für den Benutzer nicht zufriedenstellend, so wird eine Dokumentensuche (Document Retrieval) angestoßen. Mithilfe eines Volltext-Suchindex (konkret: ElasticSearch²) werden zunächst die für die Frage relevanten Dokumente identifiziert und gerankt. Aus diesen Dokumenten werden dann mithilfe eines entsprechenden Modells die Antwortpassagen extrahiert (Answer Extraction).

² <https://www.elastic.co/de/elasticsearch>

Für das Training des Modells zur Answer Extraction wurde ein Transfer-Learning-Ansatz basierend auf BERT (Devlin et al. 2018) verwendet. Als Basis diente ein multilinguales BERT-Modell, auf dem ein Fine Tuning für die Answer Extraction mit dem frei verfügbaren Stanford Question Answering Dataset (Rajpurkar et al. 2016) durchgeführt wurde.

4.3 Dialog-Design

Der Chatbot bestimmt die Antworten anhand der Nutzerabsichten (Intents), die mittels maschineller Lernverfahren aus der Frage erkannt werden müssen. Dafür ist es notwendig, die Intents zu definieren und Trainingsdaten mit möglichen Äußerungen zu hinterlegen. Bei den Intents wird zwischen Fragen zum Zertifikatskurs (z.B. "Was ist die Fallpraxis?") und allgemeinen Intents (Begrüßung etc.) unterschieden.

Für die Antworten wurde eine nüchterne, sachliche und höfliche Ausdrucksweise gewählt. Dabei wurde darauf verzichtet, den Chatbot mit markanten "Charaktereigenschaften" auszustatten.

4.3.1 Gestaltung der Wissensbasis für Faktenfragen zum Zertifikatskurs

Die Wissensdatenbank beinhaltet das Faktenwissen des Chatbots in strukturierter Form. Um Informationen flexibel zu verwalten, wurde eine Repräsentation als Knowledge Graph gewählt. Abbildung 3 zeigt einen Ausschnitt des Knowledge Graphen für den Zertifikatskurs. Er enthält beispielsweise die Informationen darüber, welche Module es gibt, wo und wie man sich anmelden muss.

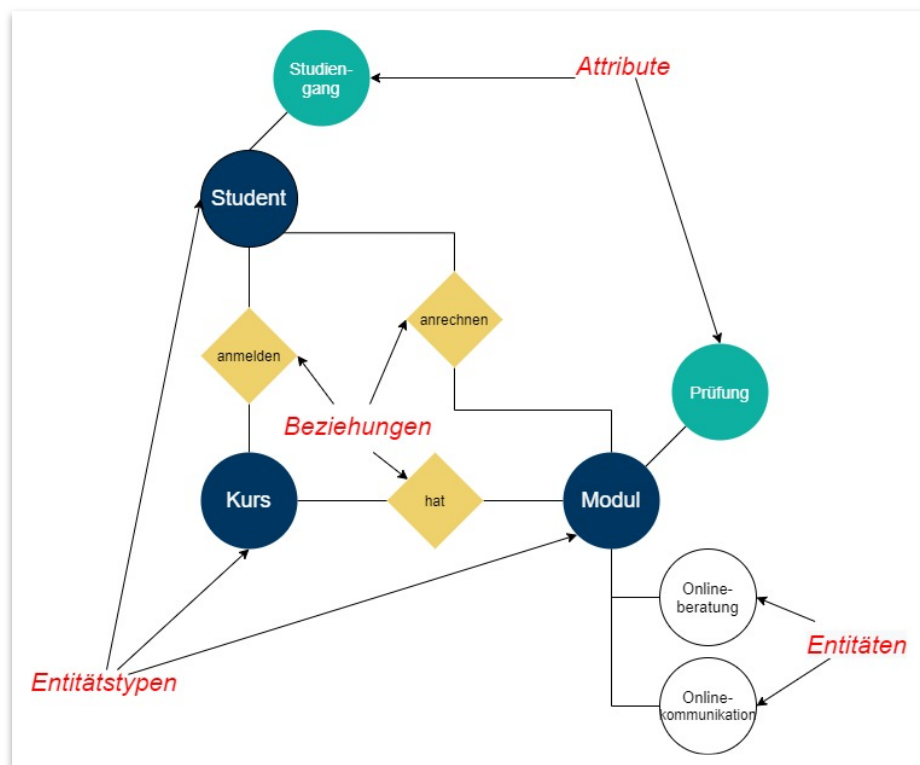


Abbildung 3 Illustration des Knowledge Graphen für den Zertifikatskurs

Für die Gestaltung des Knowledge Graphen ist sowohl technisches Knowhow als auch fachliches Wissen über das Themengebiet des Chatbots erforderlich. Im DIA-Team hat sich folgendes Vorgehen bewährt: Zunächst wurde von sozialwissenschaftlicher Seite eine grobe "Mindmap" zum Themengebiet erstellt, durch die die Entwickler*innen einen Überblick des Themas gewinnen konnten und anhand derer ein erster Knowledge-Graph entwickelt werden konnte. Auf diese Weise werden erste Entitäten, Beziehungen und Attribute festgelegt. Durch den Zwischenschritt der Mindmap, die in den Knowledge-Graphen überführt wird, ist dieser auch ohne technischen Background gut verständlich. In die erste Version des Knowledge-Graphen werden die formulierten Antworten einsortiert. Dabei war es mehrmals nötig, neue Entitäten, Beziehungen und Attribute anzulegen. In einem letzten Schritt musste kontrolliert werden, ob sich die Fragen für das Training sinnvoll labeln lassen können.

Falls einzelne Fragen vom Chatbot nicht zuverlässig unterschieden werden konnten, war es nötig, Antworten auf verschiedene Fragen in einer Antwort zusammenzufassen. Dieses Zusammenfassen mehrerer Attribute zu einem soll an einem Beispiel erklärt werden: Zunächst wurde bei der Beziehung "Anrechnung" durch zwei Attribute unterschieden, ob sich jemand ein bereits absolviertes Fach aus dem Studium auch für den Kurs anrechnen lassen will oder ob jemand Fächer aus dem Kurs für das Studium geltend machen möchte. Da es jedoch nicht möglich gewesen wäre, dass der Chatbot die Frage eindeutig einem der beiden Attribute zuweisen kann, wurden die Attribute zusammengelegt und die Antwort enthält nun beide Themen.

Außerdem wurden die Antworten um thematisch verwandte Informationen ergänzt, auch wenn nach diesen nicht direkt gefragt wurde. Das soll verhindern, dass den Nutzer*innen wichtige Hinweise entgehen, die sie normalerweise beim Besuch der Webseite zufällig entdeckt hätten. Ein Beispiel ist der letzte Satz der Antwort zu den Zulassungsvoraussetzungen des Kurses (siehe Abbildung 4): Hier wird zusätzlich zur eigentlichen Antwort darauf aufmerksam gemacht, dass der Kurs nicht mit dem Ende des Studiums abgeschlossen werden sein muss, sondern erst ein Jahr später. Falls potenzielle Kursteilnehmer*innen nicht mit dieser Sondersituation rechnen, erkundigen sie sich nicht explizit danach und laufen so Gefahr, diese wichtige Information nicht zu erhalten.

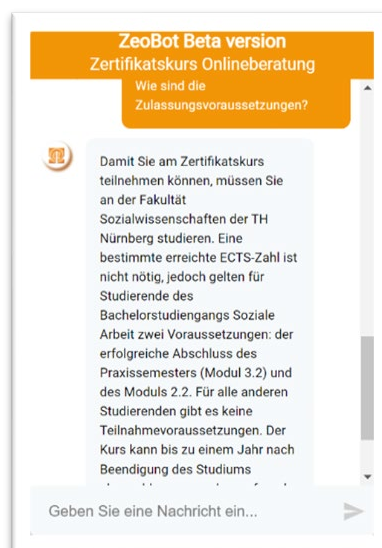


Abbildung 4 Wichtige Zusatzinformationen können zu langen Antworten führen

Sowohl das Zusammenfassen mehrerer Attribute zu einem als auch das Ergänzen von Antworten um Zusatzinformationen zum anderen haben allerdings auch einen Nachteil: Die Antworttexte werden durch dieses Vorgehen länger und enthalten teilweise Aspekte, die für die fragende Person irrelevant sind.

4.3.2 Domänen-unabhängige Antworten

Neben den Intents, die spezifisch auf den Kurs angepasst sind, hält DIA auch Antworten und Bemerkungen bereit, die auch in anderen Domänen direkt eingesetzt werden können. Einige davon sind in Abbildung 5 dargestellt.



Abbildung 5 Generische Intents des Chatbots

Beispielsweise soll der Chatbot erkennen und sich angemessen verhalten, wenn er begrüßt wird (Intent "greet"), nach seiner Funktionsweise gefragt wird (Intent "bot_function") oder Nutzer*innen ihre Unzufriedenheit mit dem Chatbot zum Ausdruck bringen (Intent "not_understood"). Mit dem Intent "get_started" wird ein neuer Dialog initiiert und der Bot selbst startet ungefragt mit einer Begrüßung und Vorstellung. Über den Intent "helpful" kann ein Benutzer-Feedback zur Zufriedenheit mit der Antwort verarbeitet werden.

An dieser Stelle wurde besonders bei Themen, die innerhalb eines Gesprächs mehrfach auftreten können, darauf geachtet, dass der Chatbot mehrere Antworttexte zur Verfügung hat und zufällig einen auswählt. Abbildung 6 zeigt die verschiedenen Reaktionen des Chatbots, wenn er eine Frage nicht verstanden hat.

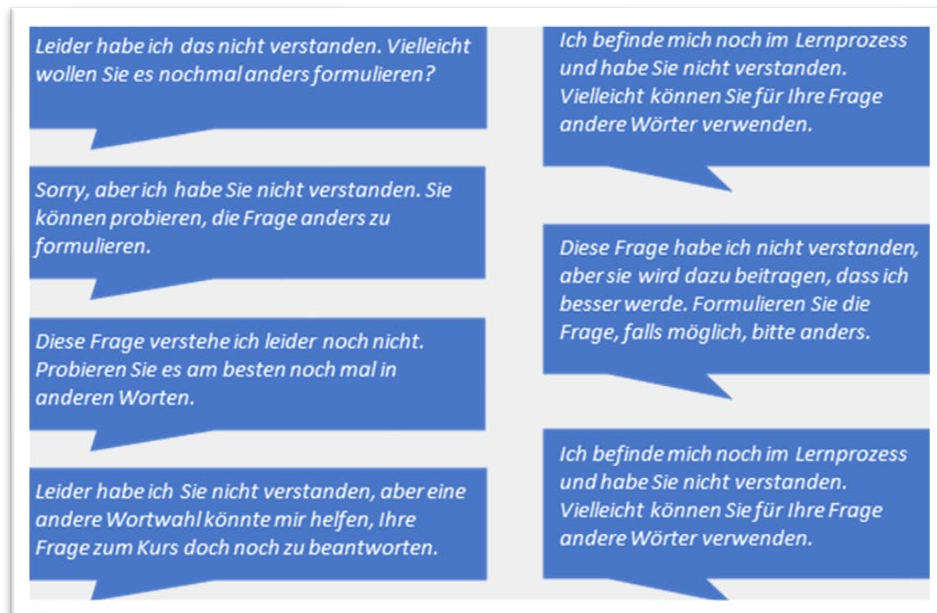


Abbildung 6 Reaktionen des Chatbots, wenn die Nutzereingabe nicht verstanden wurde

Besonders die Fragen und Anmerkungen der Nutzer*innen, die keinen Bezug zum Kurs aufweisen, können sehr unterschiedlich sein. Um diesem großen Spektrum gerecht zu werden und möglichst oft eine sinnvolle Reaktion zu zeigen, wurden die Antworten sehr unverbindlich gestaltet. Ein Beispiel ist das "intent agree". Der Chatbot erkennt, dass der*die Anwender*in eine zustimmende Reaktion zeigt. Er kann diese Zustimmung jedoch nicht in einen Kontext setzen und beispielsweise zwischen "Ja, das wollte ich wissen." und "Ja, ich habe noch eine Frage." unterscheiden. Als unverbindliche und dennoch oft passende Reaktion zeigt der Chatbot entweder einen lächelnden Smiley oder sagt etwas Freundliches, wie

- *Ich beantworte gerne Fragen zum Kurs. :)*
- *Ich freue mich über jede Frage zum Kurs. :)*
- *Ich freue mich, Ihre Fragen beantworten zu dürfen.*

4.4 Semi-automatisierte Informationsextraktion für den Aufbau von Knowledge Graphen

Für die Erstellung eines Knowledge Graphen muss zunächst eine domänenspezifische Ontologie entwickelt werden. Darin wird spezifiziert, welche Entitätstypen es geben soll, welche Attribute sie haben und wie sie über Beziehungen zusammenhängen. Wie bereits in Abschnitt 4.3.1 beschrieben, ist dieser Prozess sehr aufwändig, da Domänenexpert*innen die notwendigen Grundlagen der Technologien in der Regel nicht beherrschen und durch technische Expert*innen unterstützt werden müssen.

Vielfach ist das Domänenwissen aber bereits implizit in Form von unstrukturierten Texten vorhanden. Diese Texte können Informationen auf der Webseite sein, beschreibende Dokumente oder Emails mit Antworten auf bestimmte Fragen. Mithilfe von Verfahren zur Information Extraction (IE) können aus diesen unstrukturierten Texten Informationen zu Entitäten und Beziehungen extrahiert und in strukturierte Daten überführt werden. Da hier eine inhaltliche Strukturierung der Domäne vorgenommen werden muss, ist eine vollständige Automatisierung dieser Aufgabe nicht möglich. Durch einen semi-

automatischen Prozess können aber Empfehlungen generiert werden, die dann durch die Domänenexpert*innen angepasst werden.

Im Rahmen einer Masterarbeit (Aufotu, 2021) wurde ein Framework zur semi-automatischen Generierung eines Wissensgraphen entwickelt. Abbildung 7 zeigt den Prozess, der vollständig durch ein graphisches Modellierungswerkzeug unterstützt wird.

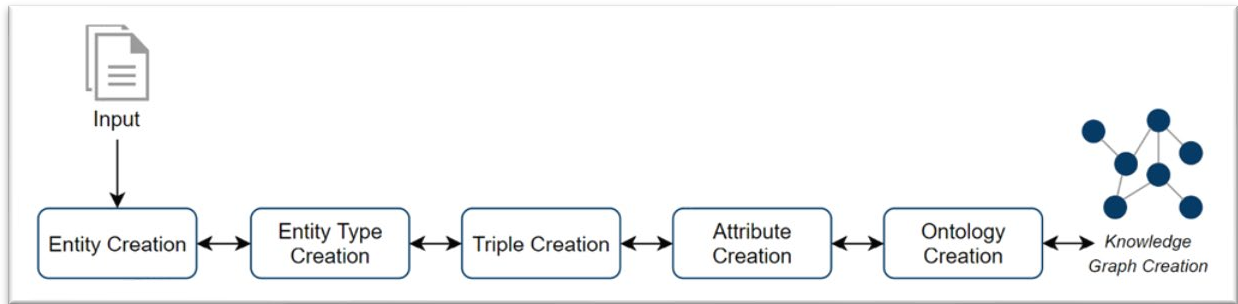


Abbildung 7 Prozess für die semi-automatische Generierung einer domänen-spezifischen Ontologie

In einem ersten Schritt werden durch eine Analyse häufig auftretender Terme Vorschläge für Entitäten ermittelt, die dann durch den Domänenexperten zu Entitätstypen verallgemeinert werden müssen. Mithilfe dieser Entitätstypen wird dann im Teilschritt "Triple Creation" nach Kandidaten für Beziehungstypen gesucht. Abbildung 8 zeigt die Visualisierung für die diesen Schritt. Dargestellt wird ein Graph, der eine Kante für jede Subjekt-Prädikat-Objekt-Kombination zwischen jeweils zwei Entitäten (als Subjekt und Objekt) enthält. Für die linguistische Analyse des Textes wird die Python-Bibliothek spaCy eingesetzt, die mittels vortrainierter Sprachmodelle den Zugriff auf die grammatikalische Satzstruktur ermöglicht.

Extract Triple
Backspace
Save Relations

Knowledge Graph Creation

- 1. Home
- 2. Entity Creation
- 3. Entity Type Creation
- 4. Triple Creation
- 5. Attribute Creation
- 6. Ontology Creation
- 7. Knowledge Graph Creation

<input type="checkbox"/>	term 1	relation	term 2	weight	rename relation	synonym
<input type="checkbox"/>	Student	vermitteln	Kurs	1		
<input type="checkbox"/>	Student	probieren	Modul	1		
<input type="checkbox"/>	Modul	empfehlen	Modul	2		
<input checked="" type="checkbox"/>	Student	absolvieren	Modul	7		
<input type="checkbox"/>	Student	absolvierenfen	Modul	1		
<input type="checkbox"/>	Student	anerkennen	Modul	3		
<input type="checkbox"/>	Student	schreiben	Prüfung	1		
<input type="checkbox"/>	Student	anschieben	Anmeldung	1		
<input type="checkbox"/>	Modul	geben	Anmeldung	1		
<input type="checkbox"/>	Student	Notenübersicht	Modul	1		
<input type="checkbox"/>	Anmeldung	erforderlich	Student	1		
<input type="checkbox"/>	Modul	benötigen	Platz	1		
<input type="checkbox"/>	Student	haben	Modul	2		
<input type="checkbox"/>	Modul	vergeben	Prüfung	1		
<input type="checkbox"/>	Student	schicken	Einschreibzeit	2		
<input type="checkbox"/>	Modul	benötigen	Prüfung	2		

Nach dem erfolgreichen Bestehen der drei Module , können Sie die Fallpraxis absolvieren .
 Sie können die Fallpraxis erst absolvieren , wenn Sie alle drei Module (Querschnitt , Workshop und Formen und Methoden) erfolgreich abgeschlossen haben .\n\nTermine\n\nTerminplanung Wintersemester 2019/20\n\nFormen und Methoden sozialpädagogischer Beratung\n\nFr 27.03/28.03 20 und Fr 24.04/25.04.20\n\nJeweils freitags von 14:00-19:00 Uhr und samstags von 9:00-17:00 Uhr . Raum : siehe Raumplan\n\nWorkshop Onlinekommunikation\n\nStart des Onlinekurses : 20.03.20;
Inhalt der Beziehung

Abbildung 8 Anwendung zur Definition von Knowledge Graphen im Teilschritt "Triple Creation"

Der oder die Anwender*in wird Schritt für Schritt durch den Prozess geführt. Dabei hat er oder sie stets die Kontrolle und den Überblick über die ermittelten Informationen. Das System generiert in jeder Phase Vorschläge durch eine Analyse der Texte, der Anwender strukturiert und abstrahiert. Am Ende des Prozesses steht das Modell eines Wissensgraphen, die Ontologie (siehe Abbildung 9).

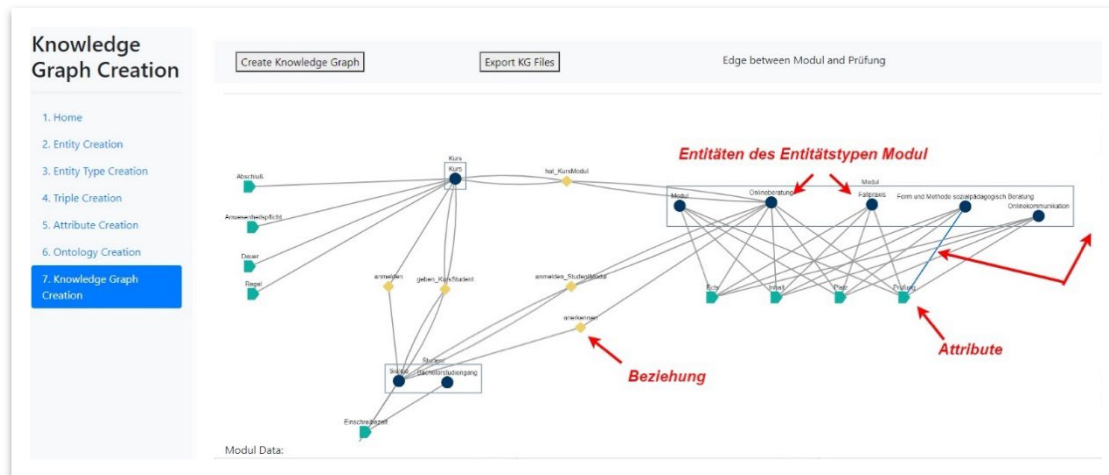


Abbildung 9 Beispiel für eine semi-automatisch erstellte Ontologie

4.5 Semi-automatisierte Erstellung und Pflege der unstrukturierten Wissensdatenbank für das Question-Answering-Modul

Wichtige Fakten zu zentralen Entitätstypen der Domäne werden im Wissensgraphen verwaltet. Darüber hinaus gibt es in der Beratung aber eine Vielzahl von Fragen zu individuellen Problemen, die nicht anhand des Wissensgraphen beantwortet werden können. Die Modellierung und strukturierte Erfassung aller möglichen Frage-Antwort-Informationen sind schlicht zu komplex und zu aufwändig.

Aus diesem Grund wird zunehmend versucht, Antworten mit Hilfe von Machine-Learning-Modellen aus unstrukturierten Daten zu ermitteln (Rajpurkar u.a., 2016; Devlin u.a. 2018). Grundlage hierfür ist die Datenqualität der darunterliegenden Dokumentenbasis. Die Erstellung und Erweiterung dieser Dokumentenbasis ist im Allgemeinen jedoch nur mit tiefgreifendem technischem Wissen möglich, da es keinen klar definierten Prozess und keine Benutzeroberfläche gibt.

Im Rahmen einer Bachelorarbeit (Noack, 2021) wurde ein Prozess für die semi-automatische Erstellung und Erweiterung dieser Dokumentenbasis entwickelt und in einer Webanwendung realisiert. Diese unterstützt die menschlichen Bearbeiter*innen, indem automatisch passende Antwortvorschläge ermittelt und durch wenig Aufwand in den Chatbot integriert werden können (vgl. Abbildung 10).

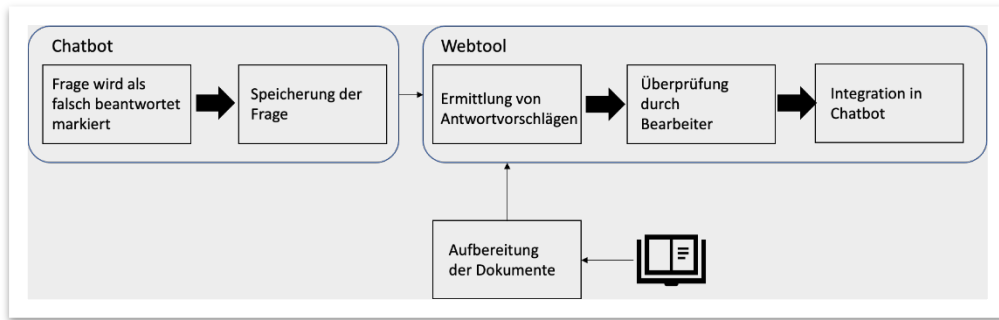


Abbildung 10 Prozess für die semi-automatische Erweiterung einer Dokumentenbasis

Der Prozess ist in zwei Teile untergliedert. Der erste Teil ist in den Chatbot integriert und dient der Ermittlung und Bereitstellung der zu beantworteten Fragen. Hierfür haben die Anwender*innen die Möglichkeit Antworten des Chatbots als „nicht hilfreich“ zu markieren, um so die entsprechenden Fragen in den Prozess zu integrieren.

Der zweite Teil des Prozesses umfasst die Extraktion von Antwortvorschlägen aus einer zuvor erstellten Dokumentenbasis. Diese Dokumentenbasis wurde aus einem großen bislang ungenutzten Bestand an E-Mail-Protokollen erzeugt. Die Protokolle stammen aus dem bisher durchgeführten E-Mail-Beratungsangebot. Vor der Integration in den Suchindex werden die Texte aus den E-Mails extrahiert, aufbereitet und anonymisiert.

Aus dieser Datengrundlage werden mögliche Antwortvorschläge für die zuvor gespeicherten Fragen ermittelt, welche den Anwender*innen als Grundlage für die Antwort auf die Frage vorgeschlagen werden. Durch die obligatorische Prüfung der Antworten durch Menschen wird verhindert, dass personenbezogene oder falsche Daten in den Chatbot gelangen. Im letzten Schritt wird die erstellte Antwort gespeichert und somit in die Datengrundlage des Chatbots integriert. Die Benutzeroberfläche ist in Abbildung 11 dargestellt.

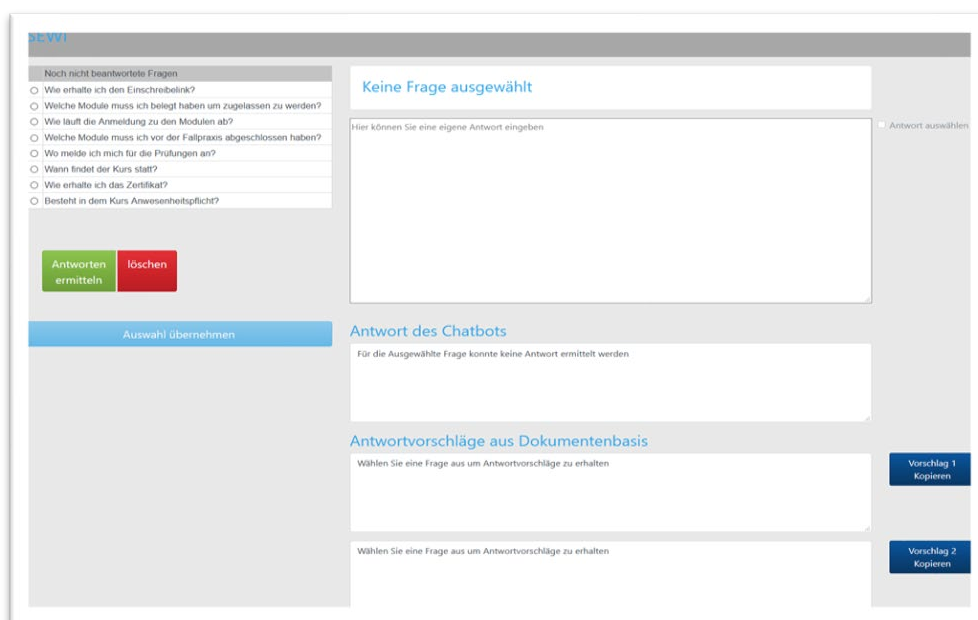


Abbildung 11 Web-Anwendung für die semi-automatische Erweiterung des Frage-Antwort-Systems

4.6 Probleme bei der Entwicklung und Bewertung der Architektur

Die Identifikation von Problemen und die Erarbeitung von Lösungsansätzen für nachfolgende Projekte war von Anfang an Teil der Forschungsaufgabe. So ging es bei der zweiten Forschungsfrage darum, wie aufwändig es ist, mit moderner Open-Source-Technologie einen leistungsfähigen Chatbot zu entwickeln. Für die Entwicklung von Chatbots und die Verwaltung der Wissensdatenbank gibt es inzwischen eine Vielzahl von Systemen. Im Projekt zum Einsatz kamen Rasa als Chatbot-Framework, spaCy für die Sprachverarbeitung, grakn.ai als Graph-Datenbank für den Knowledge-Graphen sowie SOLR bzw. später Elasticsearch für die Suche in unstrukturierten Texten. All diese Systeme sind sehr leistungsfähig und ermöglichen die professionelle Entwicklung eines Chatbots mit KI-basierter Intent-Erkennung und einer Wissensbasis für die Beantwortung der Fragen.

Es zeigte sich allerdings, dass die dynamische Entwicklung in den Bereichen maschinelles Lernen und Natural Language Processing zu häufigen und teilweise auch grundlegenden Änderungen führt. Insbesondere die Umstellung von Rasa 1.x auf Rasa 2.0 hat eine Reihe von aufwändiger Anpassungen erfordert. Hinzu kommt, dass die verwendeten Software-Pakete komplexe Abhängigkeiten untereinander haben. Die Kompatibilität der Versionen ist daher nicht immer einfach sicherzustellen. Versionsprobleme äußerten sich im Projektverlauf mehrfach durch unspezifische Fehler oder Systemabstürze, deren Analyse und Behebung ungeplant hohe Aufwände erforderte. Bei der Nutzung kommerzieller Cloud-Angebote von Google, Microsoft oder Amazon treten diese Probleme nicht in dieser Form auf, da sich hier der Service-Anbieter um das Versionsmanagement kümmert.

Die Entwicklung eines Chatbots mit Open-Source-Technologie ist daher auf jeden Fall etwas aufwändiger als mit kommerziellen Cloud-Lösungen. Gleichzeitig hat man, neben einer besseren Gewährleistung der Vertraulichkeit, deutlich mehr Freiheiten bei der Entwicklung. Die regelmäßige Aktualisierung der verwendeten Software-Frameworks muss aber auf jeden Fall bei der Planung der Projekte berücksichtigt werden.

Darüber hinaus zeigte sich, dass die Erstellung und Pflege des Knowledge Graphen und der zugehörigen Trainingsdaten unerwartet aufwändig sind. Die gewählte Architektur mit dem mehrstufigen Antwortverfahren lässt hier aber neben dem strukturierten Knowledge Graphen auch die Nutzung eines unstrukturierten Datenbestands mit Freitext-Fragen und -Antworten zu. Das im vorigen Abschnitt vorgestellte Verfahren zur systematischen Ergänzung dieser unstrukturierten Wissensbasis kann ein Weg sein, um den Aufwand für die Datenpflege deutlich zu verringern.

5 Feldtest und Akzeptanzstudie

5.1 Hintergrund

In einem Feldtest mit Studierenden der Sozialen Arbeit sollte die Funktionsfähigkeit des ZeoBots unter Bedingungen, die dem realen Einsatz ähneln, erprobt werden. Neben der Testung der technischen Funktionsfähigkeit sollte die Akzeptanz dieser für die Zielgruppe neuer Technologie untersucht werden. Die Akzeptanz technischer Innovationen ist ein inzwischen gut operationalisiertes Konstrukt, das im Technology Acceptance Model nach Venkatesh und Bala (2008) abgebildet ist. Dabei wird deutlich, dass die Akzeptanz sowohl von relativ objektiven Faktoren, als auch der subjektiven Wahrnehmung der nutzenden Personen abhängt. Da ZeoBot im Kontext des Studiengangs Soziale Arbeit eingesetzt werden soll, und Studierende und Praktiker*innen der Sozialen Arbeit beim Einsatz digitaler Technologien insgesamt eher zurückhaltend und skeptisch sind (Kreidenweis & Dilger 2012, Engelhardt et al. 2019), ist eine Untersuchung der Akzeptanz der Technologie auch für zukünftige Entwicklungen in diesem Feld von Interesse. Die Ergebnisse dieser Studie können weiterhin Hinweise liefern, wie die leichte und intuitive Bedienbarkeit des Dialoges sowie ein angenehmer Charakter der digitalen Assistenz (Kaiser et al. 2019; McTear et al. 2016; Shevat 2017) für diese Zielgruppe konkret umgesetzt werden können.

Da bereits aus verschiedenen Bereichen Studien zur Gestaltung von Chatbot-Interaktionen vorliegen, sollen die Ergebnisse dieser Studie mit den Ergebnissen von Denk (2019) verglichen werden, der einen sehr ähnlich strukturierten Chatbot bei Studierenden im Master-Studiengang „Kommunikations-, Multimedia- und Marktmanagement“ der Hochschule Düsseldorf getestet hat. So können die Ausprägungen der verschiedenen Dimensionen der Akzeptanz bei den Studierenden der Sozialen Arbeit in eine Relation zu anderen Studierenden gestellt werden.

5.2 Fragestellung

Die Fragestellungen der Akzeptanzstudie lassen sich in zwei Bereiche unterteilen: Funktionsweise und Akzeptanz.

5.2.1 Allgemeine Funktionsweise des Chatbots

Im ersten Fragenkomplex sollte untersucht werden, wie gut der Chatbot funktioniert. Dazu sollte insbesondere erhoben werden, welchen Anteil an Fragen der Chatbot richtig interpretiert und wie oft er eine korrekte Antwort formuliert. Weiterhin sollte untersucht werden, welche weiteren Merkmale bei der Interaktion zwischen Chatbot und Nutzer*innen beobachtet werden können. Für die weitere Entwicklung der Chatbot-Technologie wurde außerdem erhoben, welche Aspekte der Interaktion aus Nutzer*innen-Sicht beibehalten oder verändert werden sollten.

5.2.2 Akzeptanz des Chatbots bei der Zielgruppe

Im zweiten Fragekomplex stand die Akzeptanz des Chatbots im Fokus. Dabei war einerseits die Frage nach der Ausprägung der verschiedenen Akzeptanzdimensionen bei der Zielgruppe relevant, andererseits sollte diese Ausprägung in Relation zu Akzeptanzwerten bei Studierenden anderer Fachrichtungen gestellt werden. Damit kann eingeschätzt werden, inwiefern die Ausprägungen der Sozialarbeitsstudierenden mit bisherigen Studien übereinstimmen und wo feldspezifische Besonderheiten bestehen, die bei zukünftigen Anwendungen in diesem Feld zu berücksichtigen sind.

5.3 Methode

Die beiden Fragenkomplexe aus Feldtest und Akzeptanzstudie wurden in einem Studiendesign erhoben.

5.3.1 Vorgehen

Im Rahmen der vom Institut für E-Beratung durchgeführten Akzeptanzstudie wurden Bachelor-Studierende der Sozialen Arbeit befragt, die im Laufe ihres Studiums die Möglichkeit haben, den Zertifikatskurs Onlineberatung zu absolvieren. Dazu wurde in verschiedenen Lehrveranstaltungen zu einer freiwilligen Teilnahme an der Studie aufgerufen. Wie bei Denk (2019) wurden die Teilnehmer*innen aufgefordert, konkrete Aufgaben mithilfe von Zeobot zu lösen (siehe Abbildung 12). Dabei wurden sie darum gebeten, eigene Worte zu verwenden und das Gespräch so zu führen, wie sie die Kommunikation mit einem Chatbot als angenehm empfinden. Ausdrücklich wurde darauf hingewiesen, dass dazu Begrüßungen oder Reaktionen auf die erhaltenen Antworten gehören können.

Szenario 1	Finden Sie heraus, wie Sie sich für den Zertifikatskurs anmelden können.
Szenario 2	Erfragen Sie die Zulassungsvoraussetzungen für den Kurs.
Szenario 3	Informieren Sie sich darüber, welche Module für den Kurs absolviert werden müssen.
Szenario 4	Klären Sie, wie Sie sich für ein bestimmtes Modul, zum Beispiel das Modul "Onlinekommunikation", einschreiben können.
Szenario 5	Erkundigen Sie sich danach, ob im Kurs Anwesenheitspflicht besteht.

Abbildung 12 Aufgaben, die die Studienteilnehmer*innen mithilfe des Chatbots lösen sollten

Dieses Testsetting hatte im Vergleich zu völlig frei geführten Gesprächen ohne vorgegebene Anhaltspunkte einige Vorteile: Zunächst bekamen die Befragten, die sich größtenteils noch nie mit dem Zertifikatskurs befasst hatten, auf diese Weise Anregungen, was mögliche Fragen sein könnten. Schließlich konnte durch die thematische Reduktion eine größere Trennschärfe zwischen den zu erwarteten Fragen erzielt und damit die Qualität des Chatbots gesteigert werden. Durch die Wahl eines Laborsettings konnte die interne Validität der Studie gut sichergestellt werden. Aufgrund der mangelnden Feldnähe ist die externe Validität der Studienergebnisse notwendigerweise eingeschränkt.

Im Unterschied zu Denk (2019) wurde die Befragung nicht persönlich durchgeführt, sondern digital. Dies verringert zwar die Vergleichbarkeit der beiden Studien ein wenig, aufgrund der Pandemielage in Deutschland zum Erhebungszeitraum war jedoch eine persönliche Befragung nicht realisierbar. Während der Befragung wurde den Proband*innen freigestellt, ob sie während der Studienteilnahme über eine Videokonferenz mit den Versuchsleiter*innen in Kontakt stehen wollen oder nicht.

Die Akzeptanzstudie wurde online über einen Zeitraum von knapp vier Wochen (Herbst 2020) durchgeführt. Im Folgenden soll die Qualität des Chatbots zum Zeitpunkt der Akzeptanzstudie ausgewertet werden. Dafür wurden alle Gespräche analysiert, die sicher von Studierenden und nicht vom Entwicklungsteam geführt wurden. Dazu zählen alle Gespräche, die während der beiden Videokonferenzen, in denen Studierende live an der Studie teilgenommen haben, entstanden sind sowie alle Gespräche, deren Zeitstempel eindeutig einem abgesendeten Fragebogen zuzuordnen sind.

5.3.2 Entwicklung der Erhebungsinstrumente

Zur Beantwortung der Fragstellungen wurden verschiedene Erhebungsinstrumente eingesetzt.

Logdateien der Interaktion

Im Vorfeld der Studie wurden die Teilnehmer*innen darüber informiert, dass sämtliche Interaktionen mit dem ZeoBot aufgezeichnet werden. In einer Datenbank wurde die Interaktion mit einer eindeutigen anonymen ID für jede*n Nutzer*in hinterlegt.

Akzeptanzbefragung TAM

Dieses Erhebungsinstrument bestand aus Skalen zu den Akzeptanzdimensionen des TAM nach Venkatesh und Bala 2008. Hier wurde eine deutschsprachige Version des Instruments verwendet, die auch in der Vergleichsstudie von Denk (2019) zum Einsatz kam.

Der Wortlaut wurde nur bei denjenigen Fragen verändert, deren direkte Übernahme für den Zertifikatskurs und den dazugehörigen Chatbot nicht passend gewesen wäre.

Orientiert am TAM wurde versucht, die Skalen des Modells zu reproduzieren. Dazu wurden die theoretisch den jeweiligen Skalen zugeordneten Items einer Itemanalyse unterzogen. Dabei wurde berücksichtigt, dass ein Vergleich der Daten mit der Denk-Studie möglich sein soll, daher wurde versucht die Skalen analog zu bilden. In den Fällen, in denen die Itemanalyse eine ungenügende Reliabilität (Cronbachs Alpha < 0,7) aufwies, werden lediglich die einzelnen Items verglichen. Konkret entsprechen die Skalen-Werte dem Durchschnitt aller einzelnen Items. Wenn eine befragte Person nicht allen Items

einer Skala einen LIKERT-Wert zugeordnet hatte, wurden die Antworten dieser Person für die Skala nicht mitberücksichtigt. Die genauen statistischen Kennwerte der Skalenbildung sind im Anhang.

Offene Fragen zum Chatbot

Neben diesen likertskalierten Fragen beinhaltete der Fragebogen auch offene Fragen. Dabei bekamen die Befragten die Gelegenheit, Punkte anzusprechen, die ihnen bei der Nutzung des Bots aufgefallen sind und Anregungen für die weitere Entwicklung zu geben. Im Vergleich zu Denk (2019) wurde die Studie um zwei Fragen ergänzt: Durch die Selbsteinschätzung der eigenen Technikaffinität wird eine verbesserte Zielgruppenanalyse ermöglicht. Durch die Frage, ob der Chatbot in der Lage sein sollte, Smalltalk zu führen, entsteht ein genaueres Anforderungsprofil.

Zur Überprüfung des gesamten Erhebungsinstrumentes und des Testsettings wurde ein Pretest mit den Institutsmitarbeitenden durchgeführt.

5.4 Ergebnisse

5.4.1 Deskriptive Daten zur Stichprobe

Die überwiegende Mehrheit der Proband*innen war weiblich, was der Geschlechterverteilung im Studiengang Soziale Arbeit entspricht. Die Versuchsteilnehmer*innen waren zwischen 18 und 34 Jahre alt, wobei das Durchschnittsalter bei 23 Jahren lag. Die Hälfte der Befragten befanden sich im ersten oder zweiten Semester, aber auch Studierende im mittleren und im letzten Studienabschnitt waren vertreten.

Laut den Fragen zur Vorerfahrung wussten 59 Prozent der Befragten vor Beginn des Tests, was ein Chatbot sei. 12 Personen (44 Prozent) hatten wissentlich Kontakt mit einem Chatbot. Dazu im Widerspruch steht, dass 15 Studierende auf die Frage antworteten, wie dieser Kontakt empfunden wurde. Anscheinend wurde eine der beiden Fragen in drei Fällen missverstanden. 60 Prozent der Befragten bewerteten ihren bisherigen Chatbot-Kontakt als neutral, 40 Prozent als positiv und niemand als negativ. Die Beantwortung der Frage, ob sich die Befragten selbst als technikaffin einschätzen, entspricht einer Normalverteilung ($M: 2,96$).

In der Dimension „Wahrgenommener Nutzen“ wurde die Variable „Qualität des Ergebnisses“ (Skala WN1, vgl. 5.3.2 und Anhang 1) als hoch eingestuft ($M = 4,2$; $SD = 0,9$), genauso wie die Skala WN2, die sich aus den beiden Variablen „Image des Informationssystems“ und „Systemrelevanz im Kontext der Aufgabe“ bildet ($M = 4,3$; $SD = 0,47$). Den niedrigsten Mittelwert innerhalb dieser Skala wies das Item „Die Verwendung von ZeoBot passt zu meinem Lebensstil“ mit $M = 3,52$ auf. Außerhalb der Skalen wurde die Variable „Bisherige Erfahrung mit der Technologie“ mit dem Item „Chatbots haben mir bereits vor diesem Test bei Fragen erfolgreich geholfen“ mittelmäßig bewertet ($M = 2,89$; $SD = 1,23$) und die Variable „Funktionalität“ mit dem Item „ZeoBot hat reibungslos funktioniert“ als eher hoch eingestuft ($M = 4,00$; $SD = 0,96$).

In der Dimension „Wahrgenommene, einfache Bedienbarkeit“ zeigt die Variable „Benutzerfreundlichkeit“, dass die Studienteilnehmer*innen die Bedienung Zeobots überwiegend als einfach empfanden

(M = 4,67; SD = 0,55). Die Variable „Spielerischer Umgang“ mit dem Item „Die Verwendung von ZeoBot macht mir Spaß“ wurde mittelmäßig bis positiv bewertet (M = 3,89; SD = 0,75). Innerhalb der Variable „Leichtigkeit der Adaption“ wurde es als sehr positiv erachtet, dass ZeoBot ohne Installation einer zusätzlichen Software wie bspw. einer App nutzbar ist (M = 4,96; SD = 0,2). Bei der Frage nach dem bevorzugten Kanal für Zeobot, war die Option „Webseite des Instituts für E-Beratung“ am beliebtesten. Weiterhin bestand die Möglichkeit, sich andere Plattformen für eine Chatbotnutzung zu wünschen. Diese war insgesamt weniger beliebt, als die Umsetzung auf der Webseite, wobei die Nutzung von WhatsApp positiver gesehen wurde als der Facebook-Messenger.

Die Skala der Dimension „Einstellung zur Nutzung“ wurde mittelmäßig bis positiv bewertet (M = 3,41; SD = 0,56). Die Dimension lässt sich in zwei Variablen unterteilen: Die Variable „Vertrauen in Bezug auf Datenschutz“ wurde sehr unterschiedlich und im Durchschnitt mittelmäßig beurteilt (Besorgnis über den Versand persönlicher Informationen: M = 3,0 und SD = 1,26); Befürchtung von Datenmissbrauch: M = 2,65 und SD = 1,02). Die Variable „Beziehungstreiber“ gibt Aufschluss darüber, wie die Persönlichkeit des Chatbots empfunden wurde: Am positivsten wurde die Art und Weise beurteilt, in der ZeoBot kommuniziert (M = 4,3). Tendenziell positiv wurde angesehen, dass ZeoBot eine Persönlichkeit besitzt (M = 3,7; SD = 1,15), wobei die Meinungen hier stark auseinanderliefen. ZeoBots Persönlichkeit wurde tendenziell als sympathisch eingestuft (M = 3,6). Unschlüssigkeit herrschte bei der Frage, ob ZeoBot die Nutzer*innen mit ihrem Namen ansprechen sollte (M = 3,0; SD = 1,21). Ebenfalls sehr unterschiedlich wurde auf die Frage geantwortet, ob ZeoBot das Gefühl vermitteln sollte, mit einer echten Person zu kommunizieren (M = 2,54; SD = 1,1).

Die Skala der Dimension „Intention zur Nutzung“ wurde positiv bewertet (M = 4,08; SD = 0,73): Die Studierenden gaben größtenteils an, dass sie einen Chatbot auf der Webseite des Instituts für E-Beratung nutzen würden (M = 4,19) und ihren Kommiliton*innen die Nutzung von Chatbots weitempfehlen würden (M = 3,96).

5.4.2 Interaktionsqualität

In den Fällen, in denen ZeoBot mit einer Auswahlmöglichkeit antwortet, durch die Nutzer*innen ihre Frage spezifizieren können, werden die Auswahlmöglichkeit und die damit zusammenhängende tatsächliche Antwort als eine Reaktion gezählt, die entweder falsch oder richtig ist. Abbildung 13 zeigt das Verhältnis zwischen richtigen (37%) und falschen (63%) Antworten, die der Bot gesendet hat.

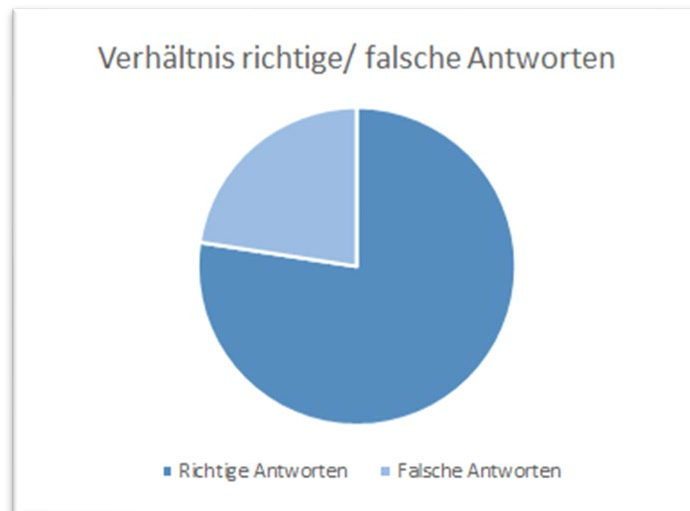


Abbildung 13 Klassifikation der Fehler der Bot-Antworten

Gar nicht in die Statistik eingegangen sind sehr strittige Antworten (knapp 20% aller Antworten), in denen der Chatbot zwar nicht optimal klassifiziert hat, aber die Antwort in dem Kontext trotzdem annehmbar war. Die Antworten, die nicht eindeutig als richtig klassifiziert werden konnten, wurden in einem zweiten Schritt zu unterschiedlichen Fehlertypen zugeordnet.

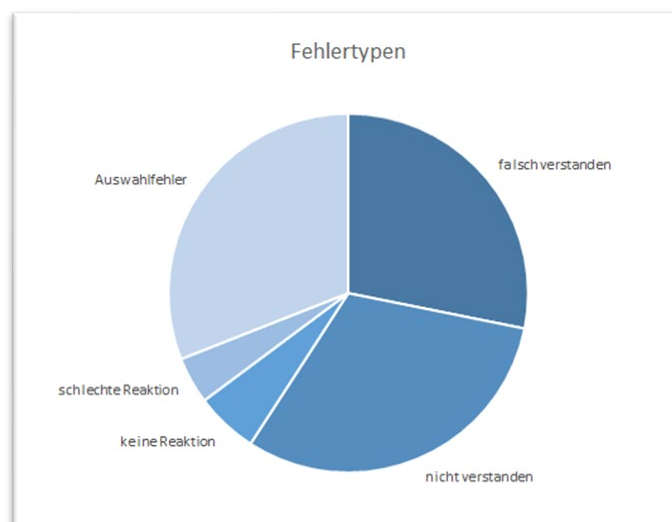


Abbildung 14 Klassifikation der Fehler der Bot-Antworten

Abbildung 14 zeigt eine Klassifikation der aufgetretenen Fehler. Als einer der beiden häufigsten Fehler konnte ZeoBot die Eingabe der Nutzer*innen nicht verstehen. Dabei handelte es sich bei allen nicht verstandenen Eingaben um konkrete Fragen zum Zertifikatskurs. In wenigen Fällen wurden Fragen wegen Rechtschreibfehlern nicht erkannt. Größtenteils jedoch lag das Nicht-Verstehen an mangelnden Trainingsdaten.

Genauso häufig kam es zu Auswahlfehlern. In dieser Fehler-Kategorie befinden sich alle Fehler, die dadurch entstanden sind, dass der Chatbot die Nutzer*innen bittet, ihre Frage zu spezifizieren, indem sie eine Option aus mehreren Möglichkeiten auswählen (vgl. Abbildung 15).

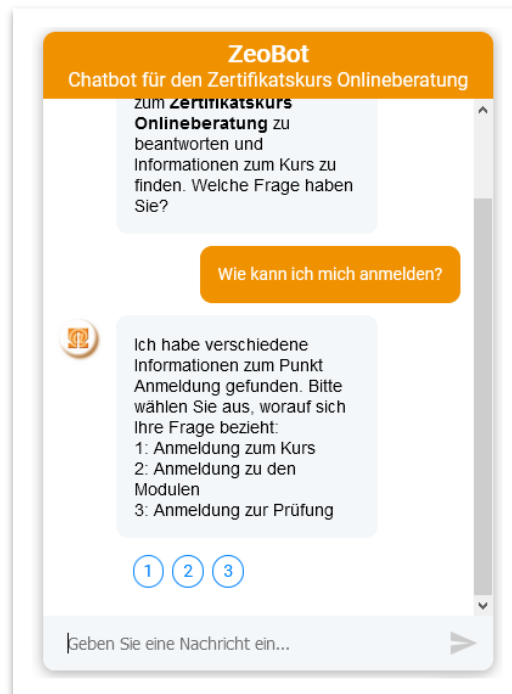


Abbildung 15 Auswahlmöglichkeiten, um mehrdeutige Fragen zu spezifizieren

Dabei versuchte ZeoBot teils für Begriffe eine Auswahl anzuzeigen, für die ihm keine zur Verfügung stand. Ein anderer Fehler im Zusammenhang mit der Auswahl war, dass Nutzer*innen zwar eine Zahl eingeben konnten, die ihrer ausgewählten Option entsprach, aber im Anschluss trotzdem eine Antwort bekamen, die nicht für das ausgewählte Attribut hinterlegt war. Dieses Verhalten folgt daraus, dass der Bot nicht unterscheiden kann, ob es sich bei der Eingabe um eine Beziehung oder ein Attribut handelt. Dies ist dem Trainingsverhalten geschuldet, da zu Beginn der Fokus mehr auf die Attribute gesetzt wurde und die Beziehungen im Training unterrepräsentiert sind. Um diesem Verhalten entgegenzuwirken müsste eine weitere Funktion vor der eigentlichen Beantwortung der Frage ausgeführt werden, die überprüft, ob es sich bei der Eingabe um ein Attribut oder eine Beziehung handelt.

Der dritthäufigste Fehler war, dass eine Eingabe des Users falsch interpretiert (Kategorie: „falsch verstanden“) wurde, so dass ZeoBot auf unpassend auf die Eingabe reagierte. In diesem Fall nahm der Bot also an, die Frage verstanden zu haben, antwortete jedoch unpassend. Als weitere Fehler traten auf, dass ZeoBot gar nicht reagierte oder ZeoBot die Eingabe richtig klassifizierte, aber die hinterlegte Antwort im Kontext unpassend war (Kategorie: „Schlechte Reaktion“).

Grund für etliche der falschen Reaktionen sind Mehrdeutigkeiten in den Trainingsdaten, die eine genaue Unterscheidung durch das Sprachmodell des Bots erschweren. Als Lösungsoptionen können hier einerseits mehr Trainingsdaten oder eine bessere Disambiguierung durch das Sprachmodell genannt werden.

5.4.3 Charakteristika der Gesprächsverläufe

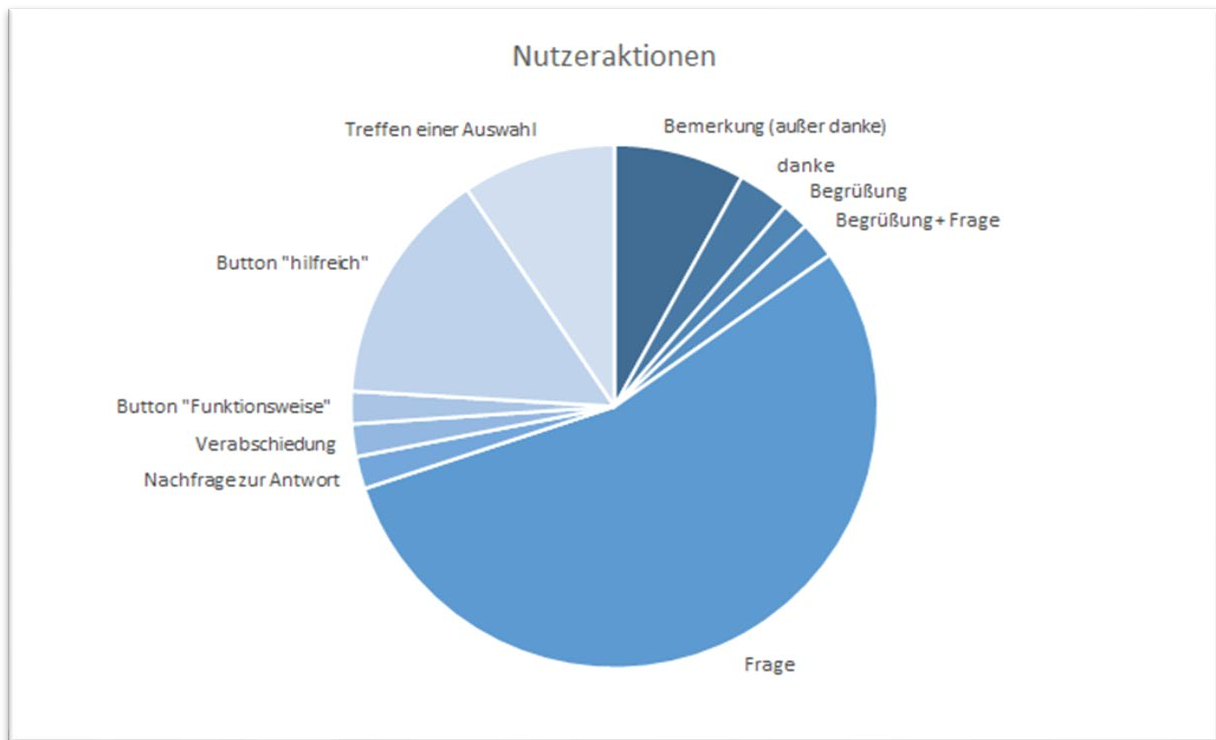


Abbildung 16 Klassifikation der Nutzeraktionen

Abbildung 16 ermöglicht Aussagen über das Verhalten der Nutzer*innen. Im Folgenden soll dies hinsichtlich der Frage untersucht werden, inwieweit die Dialoge mit dem Chatbot Gesprächen mit echten Menschen ähneln.

In etwa der Hälfte der Gespräche wurde ZeoBot begrüßt und in einem Viertel verabschiedet. Während die meisten Nutzeraktionen gezielte Fragen zum Zertifikatskurs waren, kam es darüber hinaus zu sonstigen Bemerkungen. Diese Bemerkungen wirkten oft so, als wollten die Nutzer*innen das Verständnis des Chatbots und seine Reaktion testen. Beispiele sind die Eingaben *"Hast du Hunger?"*, *"Ich liebe sie"* oder mehrmals *"wie geht's?"*.

Interessant ist folgende Beobachtung: Wenn der Chatbot eine Frage falsch beantwortet oder eine unpassende Reaktion zeigt, wird das nur selten rückgemeldet. In der Mehrheit der Fälle wird die Frage in anderen Worten wiederholt, etwas anderes gefragt oder das Gespräch abgebrochen. Auch wenn der Chatbot sagt, dass er eine Frage nicht beantworten kann, weil er sie nicht verstanden hat und noch am Lernen ist, wird darauf nie so reagiert, wie man sich einem Menschen gegenüber verhalten würde. Dahingegen bedanken sich viele User am Ende des Gesprächs und geben unaufgefordert die Rückmeldung, dass Ihnen geholfen wurde. (*"Danke, du warst hilfreich"*, o.ä.)

Betrachtet man alle Aktionen der Nutzer*innen, wie auch die des Bots, kann man sagen, dass nur ein bestimmter Anteil in Fragen und Antworten fließt. Ein knappes Viertel entfällt auf andere Aktivitäten wie Begrüßung, Verabschiedung sowie Bemerkungen.

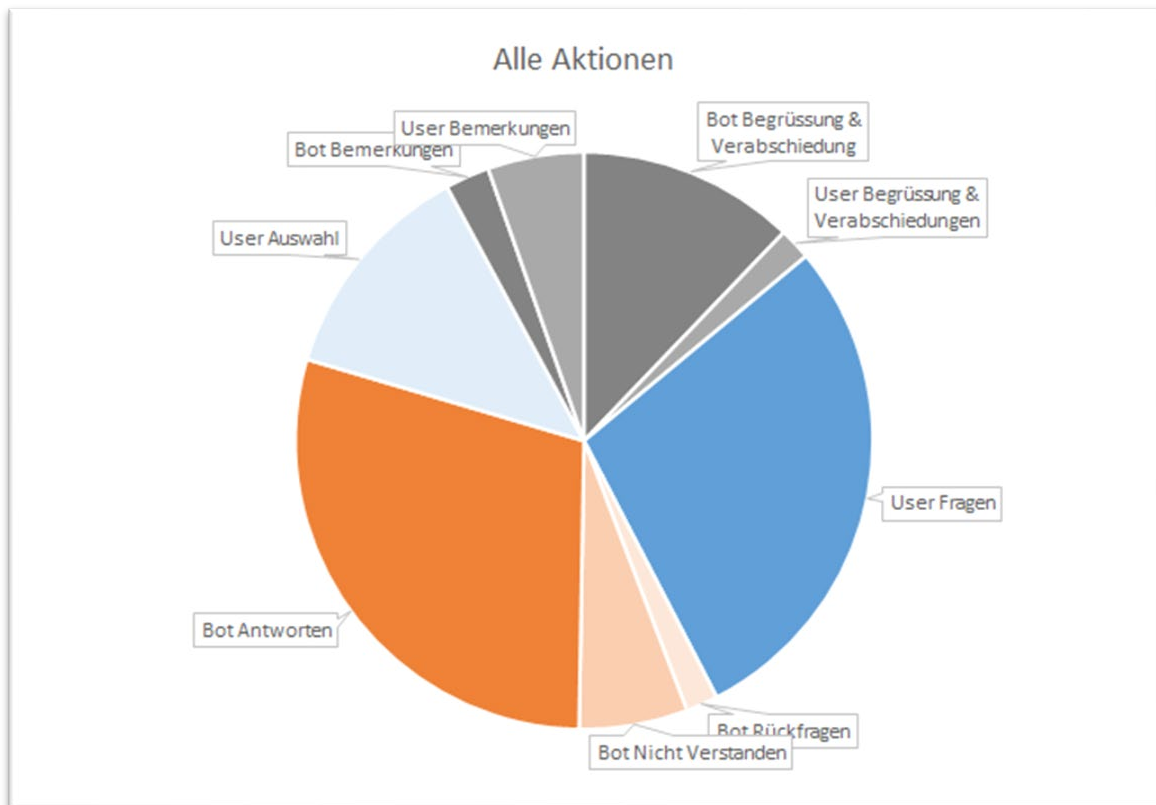


Abbildung 17 Klassifikation aller Aktionen

5.4.4 Darstellung der Antworten auf die offenen Fragen

Im Folgenden werden die Antworten auf die offenen Fragen verdichtet wiedergegeben. Antworten, die nicht in den Fragekontext passten, aber auch an anderer Stellen Erwähnung fanden, wurden in dieser Darstellung weggelassen.

Wie könnte ZeoBot die Qualität der gesendeten Antworten verbessern?

Zur Verbesserung der Antworten des Chatbots wurde gefordert, diese zu verkürzen, bei wiederholter Nachfrage einen anderen Antworttext anzuzeigen sowie das Verständnis der Chatbots zu steigern. Außerdem sollte die Funktionalität des Auswahl-Treffens verbessert werden. Als unpassend wurde empfunden, dass der Chatbot mit einem traurigen Smiley darauf reagiert, dass man keine weiteren Fragen hat. Qualität der Antworten wurde in einigen Rückmeldungen ebenfalls gelobt. Als positiver Aspekt wurden hervorgehoben, dass Antworten Links enthalten.

Was war der Grund für eine erfolgreiche bzw. nicht erfolgreiche Nutzung des Chatbots?

Als Kriterium für eine erfolgreiche Chatbot-Nutzung wurden die Richtigkeit der Antwort, der schnelle Erhalt der Antwort sowie das Weitervermitteln an eine echte Person genannt. Als erfolglos wurde eine Nutzung empfunden, wenn der Chatbot Eingaben nicht bzw. schlecht verstanden hat. Prinzipiell hängen Erfolg bzw. Misserfolg mit der Qualität der Antworten sowie mit der Qualität der eingegebenen Frage zusammen.

Wie sollte ein Chatbot gestaltet sein, damit er zu Ihrem Lebensstil passt?

Damit ein Chatbot zum Lebensstil der Befragten passt, sollte er kurz und richtig antworten sowie jederzeit erreichbar, unkompliziert in der Verwendung und kostenlos sein. Im Zweifelsfall sollte der Chatbot an echte Ansprechpartner*innen weitervermitteln. Außerdem sollte er mehrsprachig sein, häufiger eingesetzt werden, bei Standardfragen weiterhelfen und Bilder und Grafiken mitliefern. Eine Person gab an, dass ein Chatbot nie zu ihrem Lebensstil passen würde.

Welche Aufgaben würden Sie einem Chatbot im Hochschulalltag (zusätzlich) zutrauen?

Die Befragten trauten einem Chatbot viele Aufgaben im Hochschulalltag zu. So könnte ein Chatbot Informationen zum Studienverlauf, zu Ansprechpartner*innen, zu Termine und Fristen, zum Stundenplan und zu Prüfungsmodalitäten liefern. Chatbots könnten außerdem bei der Einschreibung unterstützen und Studieninteressierten sowie Erstsemestern weiterhelfen. Sie könnten die Inhalte der Webseite vermitteln, Links zu wichtigen Dokumenten versenden, bei Standardfragen helfen und Informationen zu Praktikum und Auslandssemester bereitstellen. Weitere mögliche Themen für einen Chatbot könnten die Bibliothek, das Studium mit Familie und die Adresse der Hochschule sein. Er könnte allgemein leichte Aufgaben übernehmen, jedoch keine Tutor*innen ersetzen.

Wie könnte die Reaktion von ZeoBot auf Fragen, die er nicht versteht, verbessert werden?

Drei Studienteilnehmer*innen gaben an, dass ZeoBot bereits angemessen reagiert, wenn er etwas nicht versteht. Zwei Befragte würden sich wünschen, dass ZeoBot in diesem Fall nachfragt. Außerdem könnte er auf Rechtschreibfehler aufmerksam machen, reale Ansprechpartner*innen nennen oder einen Hinweis geben, wo die benötigten Informationen zu finden sind. Als schlechte Reaktion wurde empfunden, wenn ZeoBot anstatt konkreter Auswahlmöglichkeiten nur Zahlen anzeigt, falsche Antworten gibt, sich plötzlich verabschiedet, zu häufig traurige Smileys sendet, rhetorische Fragen nicht erkennen kann oder gar nicht reagiert.

Was könnte, Ihrer Meinung nach, die Bedienung von ZeoBot (weiter) vereinfachen?

Um die Bedienung ZeoBots zu vereinfachen wurde gefordert, dass die Auswahl-Funktionalität funktionieren sollte. Außerdem sollten die Antworten verkürzt und das Chatbot-Verständnis verbessert werden. Als zusätzliche Vereinfachung wurde vorgeschlagen, Bedienhinweise am Anfang des Gesprächs anzuzeigen und eine Vorlese- und Diktierfunktion einzubauen. Zudem sollte der Chatbot auch Eingaben mit Rechtschreibfehlern verstehen und auf der Webseite auffälliger platziert werden, da die kleine gelbe Sprechblase zu unauffällig ist.

Was fehlt ZeoBot für einen spielerischen Umgang?

Drei Personen gaben an, dass die Nutzung ZeoBots nicht spielerisch sein müsste bzw. sollte. Eine weitere Person gab an, dass dem Chatbot nichts für einen spielerischen Umgang fehlen würde. Als Vorschläge, wie ein spielerischer Umgang erzielt werden könnte, wurden schöne Farben, Motivationsprüche, Witz und Grußformeln genannt.

Was könnte Ihr Vertrauen in ZeoBot steigern?

Drei der Befragten gaben an, dass durch eine Datenschutzerklärung das Vertrauen in ZeoBot gesteigert werden kann. Zwei Personen sagten, dass sie bereits volles Vertrauen hätten. Außerdem wurde ge-

nannt, dass Daten nicht an Dritte weitergegeben werden dürfen, verschlüsselte Kommunikation verwendet werden muss und angegeben werden sollte, wer für die Inhalte der Antworten verantwortlich ist. Eine Person gab an, dem Internet gegenüber generell misstrauisch zu sein.

Wie sollte, Ihrer Meinung nach, die Persönlichkeit von ZeoBot sein?

Zwei der Befragten gaben an, dass ZeoBot keine menschlich anmutende Persönlichkeit haben sollte, da dies gruselig bzw. zeitverschwendend wäre. ZeoBot sollte freundlich, höflich, entgegenkommend, hilfsbereit, ehrgeizig und weder genervt noch aufdringlich sein. Außerdem sollte er sachlich und professionell wirken. Eine Person empfindet die Persönlichkeit von ZeoBot als locker und lobt die Verwendung von Umgangssprache.

Warum würden Sie einen Chatbot nutzen bzw. warum nicht? Bezogen auf die Vor- und Nachteile aus Sicht der Studierenden.

Für die Nutzung eines Chatbots sprechen der schnelle Erhalt der Antwort und die Richtigkeit bzw. die Präzision der Antworten. Außerdem würden Befragten einen Chatbot verwenden, wenn sich dadurch die Informationsbeschaffung vereinfacht oder wenn es sich um Standardanfragen handelt. Außerdem wurden als Gründe für eine potenzielle Nutzung die ständige Verfügbarkeit und die erfolgreiche Verarbeitung von ganzen Sätzen bzw. nur von einzelnen Schlagwörtern genannt. Außerdem eignen sich Chatbots bei schambehafteten Themen.

Die Befragten würden einen Chatbot nicht nutzen, falls dieser Fehler macht bzw. Eingaben schlecht versteht und wenn es sich um spezielle, individuelle oder komplexe Anfragen sowie um psychosoziale Themen handelt.

Zwei Personen gaben an, dass ihnen die Kommunikation mit einem Bot Unbehagen bereitet, während eine Person sagte, dass sie trotz ihrer bisherigen Skepsis gegenüber Chatbots durch ZeoBot positiv überrascht wurde.

Was möchten Sie außerdem noch anmerken?

Sechs der Befragten nutzen die Gelegenheit, um ein Kompliment für ZeoBot auszusprechen. Zwei Personen betonten an dieser Stelle noch mal, dass ZeoBot nicht menschlich wirken sollte bzw. der spielerische Charakter zweitrangig ist. Außerdem wurde angemerkt, dass Chatbots noch etwas Gruseliges an sich haben und keine psychosoziale Beratung ersetzen können. Konkret auf ZeoBot bezogen wurde die Auswahl-Funktionalität kritisiert und gefordert, dass ZeoBot Rückfragen stellt. Eine Person bemerkte, dass Chatbots anstelle eines Kontaktformulars eingesetzt und bei Schwierigkeiten ein echter Mensch konsultiert werden könnte.

5.4.5 Vergleich mit der Akzeptanzstudie von Denk (2019)

Im Folgenden sollen die erhobenen Daten mit der Akzeptanzstudie von Denk (2019) verglichen werden. Für den Vergleich wurde abhängig davon, ob eine Normalverteilung vorlag, ein T-Test bzw. ein Mann-Whitney-U-Test durchgeführt. Die Ergebnisse dieser Tests sowie die Mittelwerte und Standardabweichungen beider Gruppen sind in Anhang 2 dargestellt. In drei Fällen konnte ein signifikanter Unterschied festgestellt werden. Dabei handelte es sich um das Item „Wissen Sie was ein Chatbot ist“

(SW: $M=1,41$, $SD=0,50$; MKM: $M=1,14$, $SD=0,36$, $U=208,50$, $p < .05$), das Item „Chatbots haben mir bereits vor diesem Test bei Fragen erfolgreich geholfen“ (SW: $M=2,89$, $SD=1,23$; MKM: $M=3,86$, $SD=1,35$, $U=71,50$, $p < .05$), die Skala „Einstellung zur Nutzung“ (SW: $M=3,41$, $SD=0,56$; MKM: $M=4,35$, $SD=0,46$, $U=-5,96$, $p < .05$). Die Unterschiede bei allen anderen Items und Skalen erwiesen sich nicht als signifikant, sodass hier von keinen bedeutsamen Unterschieden zwischen den beiden Populationen ausgegangen werden kann.

5.5 Diskussion der Ergebnisse aus sozialwissenschaftlicher Perspektive

Insgesamt zeigt sich, dass ZeoBot in dem Untersuchungssetting gut in der Lage war, mit den Testpersonen zu interagieren. Die Menge an korrekten Antworten ist für einen Bot im vorgelegten Entwicklungsstadium als positiv zu bewerten. Die beobachteten Gesprächsverläufe und auch die Rückmeldungen der Testpersonen zeigen, dass diese Form der Interaktion und Informationsbeschaffung für Studierende der Sozialen Arbeit prinzipiell attraktiv ist und weiterentwickelt werden sollte.

Vergleicht man die Ergebnisse der Studie mit Denk (2019), wird zunächst deutlich, dass die soziodemographischen Daten der Befragten sehr ähnlich sind. Ein ähnliches Durchschnittsalter (SW: 23, MKM: 24,9) und ein ähnliches Geschlechterverhältnis (SW: 89% weiblich, MKM: 86% weiblich) stellen gute Voraussetzungen für einen Vergleich dar.

Ebenfalls eine gute Grundlage für den Vergleich ist die Tatsache, dass die Befragten die Qualität der beiden Chatbots ähnlich einschätzen: Der Skala zur Zufriedenheit mit den erhaltenen Informationen, allgemein und im Vergleich zur Suche auf der Webseite, wurde von beiden Gruppen im Durchschnitt gleich zugestimmt ($M=4,2$). Auch die Skala zum allgemeinen wahrgenommenen Nutzen (WN2), erhielt ähnliche Zustimmung (SW: $M=4,3$; MKM: $M=4,5$) und lässt wie die Skala zum szenario-bezogenen wahrgenommenen Nutzen (WN1) auf eine in beiden Studien annehmbare Qualität der Chatbots schließen.

Die beiden Gruppen unterscheiden sich in ihrer Vorerfahrung: So gaben 41 % der SW-Studierenden und nur 14 % der MKM-Studierenden an, nicht zu wissen, was ein Chatbot sei. Dieser Unterschied erwies sich als signifikant. Ebenfalls stimmten deutlich weniger SW-Studierende der Frage zu, ob ihnen vor dem Test schonmal ein Chatbot geholfen habe, als in der Vergleichsgruppe. Damit liegen die Ergebnisse innerhalb des Erwartungsbereichs, dass Studierende der Sozialen Arbeit insgesamt weniger Erfahrung mit Technik haben als Studierende anderer Studiengänge.

Die Skala „Einstellung zur Nutzung“ ergab einen signifikanten Unterschied (SW: $M=3,4$; MKM: $M=4,3$). Diese Skala besteht aus einigen Items, die sich direkt mit der emotionalen Einstellung gegenüber der Chatbot-Technologie beschäftigen, insbesondere zur Persönlichkeit des Chatbots. Dabei wird deutlich, dass die Studierenden der Sozialen Arbeit deutlich mehr Datenschutzbedenken haben und einem zu menschlichen Charakter des Chatbot deutlich kritischer gegenüberstehen als die Vergleichsgruppe. Dieser Eindruck deckt sich mit dem Antwortverhalten bei den offenen Fragen.

Die Studierenden der sozialen Arbeit scheinen die Imitation einer menschlichen Persönlichkeit durch Chatbots kritischer zu sehen. Allerdings zeigt die Analyse der Gesprächsverläufe, dass sie durchaus mit dem Chatbot ähnlich wie mit Menschen interagieren. Sie bedanken sich beispielsweise für Auskünfte,

zumindest, wenn sie mit ihnen zufrieden sind, sie tauschen Befindlichkeiten aus und reagieren auch einmal ärgerlich. Diese relative ungezwungene Interaktion deckt sich mit den Daten zur allgemeinen Interaktion von Sozialarbeiter*innen mit neuen Technologien, wie sie auch Engelhardt et al (2019) beschreiben. Für die Zielgruppe der Studierenden der Sozialen Arbeit ist die Interaktion mit digitalen Medien durchaus vertraut und auf einem unreflektierten Level erfolgt sie auch sehr ähnlich wie bei anderen Studierenden. Bei der Beurteilung aus einer professionellen Perspektive werden allerdings Vorbehalte und Sorgen formuliert. Eventuell handelt es sich hier um eine Form fachgebundener Sozialer Erwünschtheit, bzw. die Übernahme moralischer Argumentationsmuster, die im Fachkontext insbesondere von älteren Fachkräften der Psychosozialen Arbeit vertreten werden.

Für zukünftige Studien deuten die Ergebnisse an, dass auch im Feld der Psychosozialen Arbeit Chatbot-Technologien auf Akzeptanz stoßen können, wenn die genaue Gestaltung den größeren Vorbehalten Rechnung trägt. Ob sich die Vorbehalte im Laufe der Zeit und mit steigender Erfahrung im Umgang mit der Technologie verändern, sollte Gegenstand weiterer Studien werden.

5.6 Auswertungen mit PowerBI

Für die interaktive Visualisierung und Auswertung der Daten wurde die Software PowerBI Desktop verwendet. Mit dieser kann sich direkt mit der Datenbank des Chatbots verbunden werden, wodurch sich dieser Ansatz speziell für den laufenden Betrieb eignet. Für die folgenden Auswertungen wurde sich auf einen abgegrenzten Datensatz beschränkt, die Datenquelle kann jedoch mit wenig Aufwand auf eine Echtzeitabfrage umgebaut werden.

Anzahl der ausgetauschten Nachrichten

Der Fokus der Auswertung lag auf der Anzahl der ausgetauschten Nachrichten, sortiert nach den einzelnen Gesprächen. Es kann nach einem bestimmten Zeitraum sowie nach dem Testsetting wie beispielsweise der Akzeptanzstudie gefiltert werden. Es wird für den eingestellten Datensatz sowohl die minimale (2) als auch die maximale Anzahl (115) der ausgetauschten Nachrichten sowie der Mittelwert (47,6) angezeigt. Dies dient dem groben Überblick über den gesamten Datensatz.

In einem Kuchendiagramm wird das Verhältnis von Nutzer*innen und Bot Nachrichten dargestellt. Durch diese Auswertung können Rückschlüsse auf den Gesprächsverlauf gezogen werden. Eine ausgeglichene Aufteilung beispielsweise, lässt auf eine fortlaufende Konversation schließen.

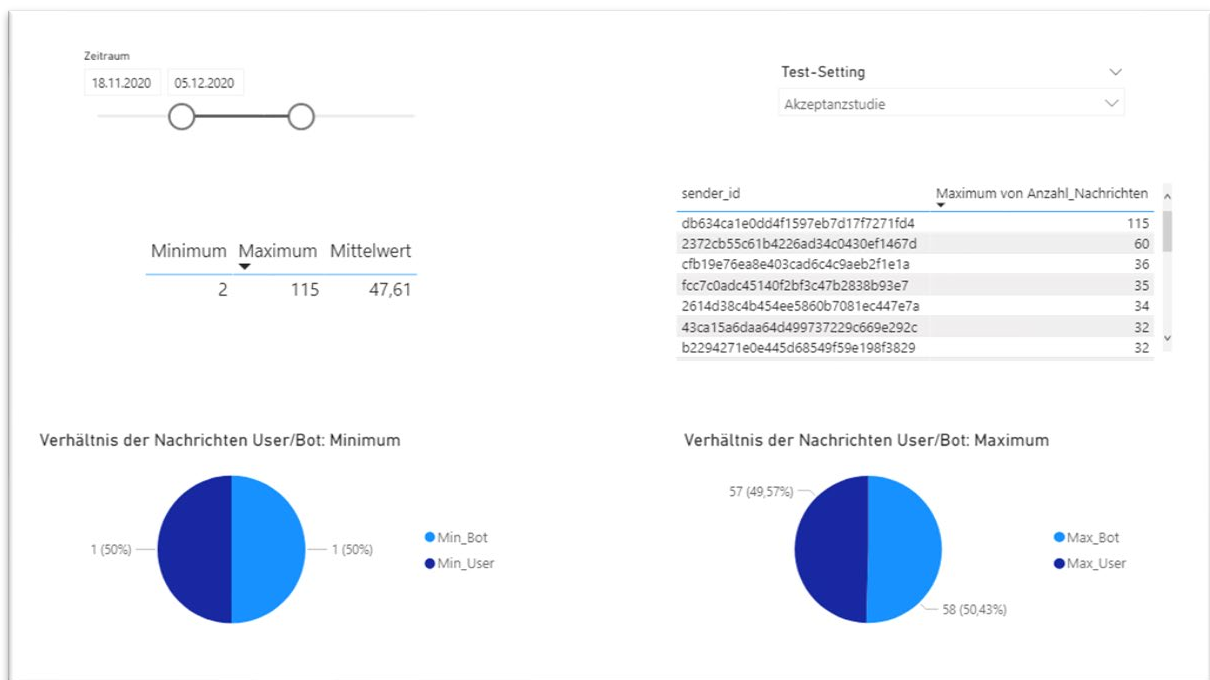


Abbildung 18 PowerBI - Anzahl der ausgetauschten Nachrichten

Eine weitere Komponente der Auswertung ist eine Liste der Anzahl der ausgetauschten Nachrichten, gruppiert nach einzelnen Konversationen. Diese kann sowohl absteigend als auch aufsteigend sortiert werden. Es besteht die Möglichkeit, einzelnen Einträge auszuwählen und eine tiefergehende Analyse durchzuführen, indem sich der gesamte Gesprächsverlauf der ausgewählten Konversation anzeigen lässt. So können auffällige Konversationen wie beispielsweise sehr lange oder kurze Gespräche genauer analysiert werden.

Confidence des Chatbot

Eine weitere Auswertungsmöglichkeit liegt darin, die sogenannte Confidence des Chatbots auszuwerten. Diese gibt an, wie sicher sich der Bot ist, die Nachrichten der Nutzer*innen richtig erkannt zu haben.

Dies dient der Filterung der einzelnen Gespräche, um beispielsweise nur diejenigen genauer zu analysieren, die eine besonders niedrige Confidence aufweisen. Auch hier können einzelne Gespräche ausgewählt und näher analysiert werden. Hierbei wird zu jeder Nachricht im Gespräch der jeweils vom Chatbot erkannte Intent sowie die dazugehörige Confidence angezeigt (siehe Abbildung 19). Diese Auswertung dient vor allem technischen Aspekten, da durch den Gesprächsverlauf Rückschlüsse auf eventuell falsch erkannte Intents gezogen werden können.

id	Absender	text	confidence	intent	sender_id
3777	bot	"Hallo, was m\u00f6chten Sie \u00fber den Zertifikatskurs Onlineberatung wissen?"			b0c98e17b4284174be92d06e721a154a
3824	user	"/get_started"	1,00000	get_started	b0c98e17b4284174be92d06e721a154a
3826	bot	"Hallo, ich bin ZeoBot und kann Ihnen helfen, Fragen zum \"Zertifikatskurs Onlineberatung\" zu beantworten und Informationen zum Kurs zu finden. Welche Frage haben Sie?"			b0c98e17b4284174be92d06e721a154a
3828	user	"hALLO"	0,99999	greet	b0c98e17b4284174be92d06e721a154a
3830	bot	"Hallo, haben Sie eine Frage zum Zertifikatskurs Onlineberatung?"			b0c98e17b4284174be92d06e721a154a
3835	user	"Wie schreibe ich mich ins Fach Onlinekommunikation ein?"	1,00000	query_relation_key_and_entity	b0c98e17b4284174be92d06e721a154a
3839	bot	"Sobald Sie sich bei uns angemeldet haben, bekommen Sie jedes Semester den Link zugeschickt, der f\u00fcr die Anmeldung zu den einzelnen Lehrveranstaltungen des Zertifikatskurses erforderlich ist. Dies geschieht ungef\u00e4hr zeitlich zur Einschreibzeit. Sie ben\u00fctigen diesen, um Zugriff auf die drei Module im VirtuOhm erhalten und um sich da\u00fcr einschreiben zu k\u00f6nnen."			b0c98e17b4284174be92d06e721a154a
3844	user	"Tschau"	1,00000	goodbye	b0c98e17b4284174be92d06e721a154a
3846	bot	"Auf Wiedersehen!"			b0c98e17b4284174be92d06e721a154a
3857	user	"Hallo ZeoBot"	0,99999	greet	b0c98e17b4284174be92d06e721a154a
3859	bot	"Guten Tag, was m\u00f6chten Sie \u00fber den Zertifikatskurs Onlineberatung wissen?"			b0c98e17b4284174be92d06e721a154a
3861	user	"Wie meldet man sich denn zu diesem Kurs an?"	0,49314	query_relation_key_and_entity	b0c98e17b4284174be92d06e721a154a
3867	user	"Wie meldet man sich zum Kurs an?"	1,00000	query_relation_key_and_entity	b0c98e17b4284174be92d06e721a154a
3872	bot	"Sobald Sie sich bei uns angemeldet haben, bekommen Sie jedes			b0c98e17b4284174be92d06e721a154a

Abbildung 19 PowerBI - Detailansicht der Confidence innerhalb einer Konversation

Anteil der als hilfreich markierten Antworten

Der Fokus dieser Auswertung liegt darin, zu ermitteln, in welchem Verhältnis die Anwender*innen die Antworten des Bots als „hilfreich“ und „nicht hilfreich“ bewertet haben. Durch diese Gegenüberstellung können Rückschlüsse auf die Qualität der Antworten und die daraus resultierende Zufriedenheit der Anwender*innen gezogen werden. Die Gegenüberstellung erfolgt sortiert nach den einzelnen Konversationen. Auch hier kann bei Bedarf der Gesprächsverlauf näher analysiert werden.

Zeichenanzahl der einzelnen Nachrichten

Durch die Auswertung der Zeichenanzahl der einzelnen Nachrichten können Informationen über die Komplexität und den Umfang der einzelnen Fragen gewonnen werden. Diese Informationen können genutzt werden, um weitere Analysen anzustoßen oder einzugrenzen. So kann sich beispielsweise nur auf Nachrichten mit einer besonders hohen Anzahl an Zeichen beschränkt werden, um zu ermitteln, ob die Frage besonders verschachtelt gestellt war und wie der Bot darauf reagiert hat.

5.7 Technische Verbesserungsbedarfe auf Basis der Untersuchung

Die Spracherkennungskomponente (siehe Rasa NLU, Abschnitt 4.2) nutzt maschinelles Lernen für die Erkennung der Intents. So sollen unterschiedliche Äußerungen, die inhaltlich die gleiche Frage ausdrücken, auf die gleiche Intension abgebildet werden. In der bisher eingesetzten Konfiguration weist die Sprachverarbeitungskomponente allerdings deutliche Schwächen auf. Zum einen waren unerwartet viele Trainingsdaten auch schon für relativ einfache Intents erforderlich.

Zum anderen gibt es aber auch Feinheiten in den Fragen, die für ein Sprachmodell sehr schwierig zu unterscheiden sind. Hier ein Beispiel:

- "Welche **Module** muss ich **absolvieren**, damit ich am **Kurs** teilnehmen kann?"
- "Welche **Module** muss ich **absolvieren**, wenn ich am **Kurs** teilnehme?"

Die beiden Fragen unterscheiden sich nur geringfügig in ihrer Wortabfolge und verlangen dennoch zwei unterschiedliche Antworten. Solche "Doppeldeutigkeiten" galt es bereits beim Konzipieren der Antworten vorzusehen und zu berücksichtigen: Entweder die Antwort wurde so gestaltet, dass sie auf alle zu erwartenden Fragen eingeht oder so, dass sie zumindest die am häufigsten zu erwartenden Fragen beantwortet. Auf die Möglichkeit einer spezifischen Rückfrage wurde verzichtet, da es zur Zielsetzung gehörte, den Chatbot möglichst generisch zu gestalten.

Ein weiteres Problem ist, dass nicht jede Frage Stichwörter enthält, die sich eindeutig auf Entitäten, Attribute oder Beziehungen beziehen. Nachfolgend einige Fragen, die nur die Entität "Kurs" aufweisen:

- *Was muss ich tun, damit ich am Kurs teilnehmen kann?*
- *Muss ich immer am Kurs teilnehmen?*
- *Wie erfahre ich, ob ich am Kurs teilnehmen kann?*
- *Was muss ich tun, wenn ich erfahren habe, dass ich nicht mit Sicherheit am Kurs teilnehmen kann?*

Wenn eine Eingabe nur eine Entität oder eine Beziehung enthält, für die keine Antwort hinterlegt ist, werden die entsprechenden Attribute zum Auswählen angezeigt. Passiert dies oft, wird der Gesprächsverlauf dadurch holprig.

Das vergleichsweise gute Chatbot-Verständnis während des Testdurchlaufs konnte nur dadurch erreicht werden, dass ein Testsetting konstruiert wurde, in dem vor allem mit den vorgegebenen Fragen zu rechnen war, die besonders gut trainiert worden waren. Außerdem wurden Antworten bewusst so gestaltet, dass sie zu mehreren verschiedenen Fragen passen. Dieses Vorgehen führte zu langen Chatbot-Antworten. In der Akzeptanzstudie zeigte sich jedoch, dass den Nutzer*innen kurze und prägnante Antworten wichtig sind.

Außerdem brachte die Akzeptanzstudie ein nicht bedachtes Problem zum Vorschein: Eigentlich stellt ZeoBot den Nutzer*innen keine Fragen. Einzige Ausnahme ist die Frage danach, ob die Nutzer*innen weitere Fragen haben bzw. welche Fragen sie haben. Das ist einerseits in der Begrüßung der Fall, aber auch gelegentlich als Reaktion auf ein "danke" oder ein positives Feedback.

Dabei erwiesen sich alle Formulierungen als nachteilhaft, die mit "ja" oder "nein" zu beantworten sind, z.B. „Haben Sie noch weitere Fragen?“. In der aktuellen Version des Chatbots kann dieser nicht unterscheiden, ob Nutzer*innen auf diese Frage reagieren oder ob mit dem "ja" bzw. "nein" Zustimmung bzw. Ablehnung bezüglich einer Chatbot-Antwort signalisiert wird. Der Versuch, Antworten so zu formulieren, dass sie in beiden Fällen stimmen, ist teilweise misslungen. Die Antwort auf ein "ja" mit Sätzen wie „Ich freue mich über jede Frage zum Kurs!“ passte in allen Fällen, während die Antwort auf ein Nein mit "Schade!" oder einem traurigen Smiley öfter als unpassend empfunden. Ein Beispiel ist die folgende Konversation:

- *"Interessiert Sie noch etwas zum Zertifikatskurs?"*
- *"Nein, danke für die Infos!"*
- *":("*

Zu beachten gilt also, dass der aktuelle Gesprächskontext an Bedeutung gewinnt, wenn der Chatbot sich mit Fragen an die Nutzer*innen wendet. Das Verständnis des Chatbots in diese Richtung zu erweitern, hätte zusätzlich den Vorteil, dass auch Rückfragen der Nutzer*innen besser verarbeitet werden könnten. Ein "Kontext-Verständnis" bedarf jedoch ein ausgesprochen gutes Sprachverständnis.

6 Projektergebnisse

6.1 Studentische Arbeiten

Im Rahmen des Projekts wurden Abschlussarbeiten erstellt, die von den beteiligten Professoren (Informatik: Prof. Jens Albrecht, Institut für E-Beratung: Prof. Robert Lehmann) betreut wurden:

- Markus Glas: *Entwicklung eines wissensbasierten Chatbots für die interaktive Beratung zu sozial-wissenschaftlichen Lehrangeboten* (Master-Arbeit 2019, Wirtschaftsinformatik:)
- Nigar Aufotu: *Entwicklung eines Frameworks zur semi-automatisierten Informationsextraktion für Knowledge Graphen* (Master-Arbeit 2020, Informatik)
- Mathias Rühl: *Prototypische Implementierung eines Chatbots zur Administration einer Wissensdatenbank* (Bachelor-Arbeit 2020/21, Informatik)
- Jonas Noack: *Semi-automatisierte Erstellung und Erweiterung einer unstrukturierten Wissensdatenbank für Question-Answering-Systeme* (Bachelor-Arbeit 2021, Informatik)
- Malte Bendl: *Optimierung der Intent-Erkennung eines Chatbots durch Evaluierung verschiedener Sprachmodelle* (Bachelor-Arbeit 2021, Wirtschaftsinformatik, aktuell laufend)

6.2 Veröffentlichung

Die gewonnen Erkenntnisse, sowohl zur technischen Realisierung, als auch zu den Akzeptanztests werden zur wissenschaftlichen Darstellung in Fachpublikationen und Fachtagungen aufbereitet.

6.3 Weitere Forschungsprojekte

Die Staedtler-Stiftung ermöglichte mit der Förderung dieses Projekts den Aufbau von Know-how, das direkt in die Beantragung weiterer Forschungsprojekte fließt und so maßgeblich an der Etablierung der beteiligten Professoren in diesem Forschungsfeld beiträgt.

So wurde die Projektskizze "Unkafka – Behördliche Schreiben einfach und verständlich" im März durch das Bundesministerium für Arbeit und Soziales ausgezeichnet. Bei dem Projekt geht es darum, mithilfe von künstlicher Intelligenz die sprachliche Kommunikation zwischen Behörden und ähnlichen Organisationen mit den Bürgern zu verbessern. Aktuell wird hierzu ein Antrag vorbereitet.

Bereits während der Projektlaufzeit flossen die Erkenntnisse in den erfolgreichen Antrag: "Interactive Artificial Intelligence – IAI" ein. In diesem Projekt soll für Jugendliche prototypisch ein Chatbot aufgebaut werden, der mit seinen Nutzer*innen zu schambesetzten Themen niederschwellig interagiert. Die Eignung des Modells für die Beratungsunterstützung bei schambesetzten Themen wird parallel evaluiert.

Die Expertise aus diesem Projekt fließt weiterhin in dem Aufbau des Schwerpunkts „KI und Soziale Arbeit“ ein, den das Institut für E-Beratung in Kooperation mit dem Fachbereich Informatik ausbaut. Die Aktivitäten zielen darauf, KI zu verstehen (Kurs KI, KAIMO), außerhalb kommerzieller Interessen

mit zu gestalten (CaSoTex, IAI) und sie als Entlastungsangebot (DiA) in die Soziale Arbeit und Bildungsberatung (Empowering Learning) zu integrieren.

- Kurs KI und Soziale Arbeit: Das Institut für E-Beratung entwickelt einen Kurs im Rahmen der Open vhb für alle interessierten Praktiker*innen und Studierende der Sozialen Arbeit. Die Teilnahme ist kostenfrei und ab 2021 an der Virtuellen Hochschule Bayern (vhb) möglich.
- KAIMo: Kann ein Algorithmus im Konflikt moralisch kalkulieren? Dieser Frage geht ein interdisziplinäres Forschungsteam aus Blickwinkeln der Philosophie, Informatik und Sozialwissenschaften nach.
- CASoTex: Im Projekt Computergestützte Analyse Sozialwissenschaftlicher Texte mit Hilfe maschineller Lernverfahren – wurden mit computerlinguistischer Methoden und maschineller Lernverfahren zur automatisierten Analyse sozialwissenschaftlicher Texte große Textmengen unter verschiedenen Fragestellungen untersucht und ein Modell geschaffen, welches weitere Untersuchungen ermöglicht (Lehmann u.a., 2020, Grandeit u.a., 2020)
- IAI: Im Projekt “Interactive Artificial Intelligence” soll für Jugendlichen prototypisch ein Chatbot aufgebaut werden, der mit seinen Nutzer*innen zu schambesetzten Themen niederschwellig interagiert. Die Eignung des Modells für die Beratungsunterstützung bei schambesetzten Themen wird parallel evaluiert.
- Empowering Learning: Das Projekt ermöglicht adaptiven Lernens in der beruflichen Weiterbildung durch den Einsatz und die Erprobung von KI.

All diese Aktivitäten stärken nicht zuletzt den Aufbau des neu gegründeten Zentrums für Künstliche Intelligenz (KIZ) an der Technischen Hochschule Nürnberg. Das KIZ bündelt die interdisziplinär vorhandenen Kompetenzen im Bereich KI ist damit Katalysator für neue Forschungsprojekte.

7 Zusammenfassung und Ausblick

Im Rahmen des Projektes DiA wurde erfolgreich eine digitale Assistenz für die psychosoziale Beratung entwickelt. Konkret entstanden ist ein Chatbot für die Studierenden-Beratung zu einem Zertifikatskurs. Die meisten der gesetzten Projektziele konnten im Projektverlauf wie geplant umgesetzt werden. Erwartungsgemäß war die Entwicklung aber auch von einigen Fehlschlägen und ernüchternden Erfahrungen begleitet, die zu wichtigen Erkenntnissen für zukünftige Projekte geführt haben.

Die Umsetzung des Chatbots hat es möglich gemacht, mit Studierenden der Sozialen Arbeit in einem realen Szenario den Einsatz der Chatbot-Technologie zu testen. Angesichts der sehr kritischen Rezeption digitaler Technologien im Feld der psychosozialen Arbeit war dabei vor allen Dingen die Frage relevant, inwiefern die Chatbot-Technologie bei dieser Zielgruppe auf Akzeptanz stößt. Die Akzeptanzstudie konnte zeigen, dass eine prinzipielle Akzeptanz von Chatbots auch in der psychosozialen Arbeit erreichbar ist. Der Vergleich zu Denk (2019) zeigte hier relativ viele Übereinstimmungen mit Studierenden anderer Fachrichtungen und auch die Unterschiede in einzelnen Bereichen können nicht als prinzipielle Ablehnung dieser Technologie interpretiert werden. Vielmehr sind diese Unterschiede dahingehend zu interpretieren, dass die jeweiligen Zielgruppen an das Interaktionsdesign eines Chatbots unterschiedliche Anforderungen stellen, die bei der Konzeption eines Bots frühzeitig berücksichtigt werden müssen.

Die zweite Forschungsfrage befasste sich damit, wie aufwändig die Entwicklung eines Chatbots mit Open-Source-Technologie ist. Hierbei wurde deutlich, dass die eingesetzten Technologien aufgrund der Entwicklungsdynamik einen nicht unerheblichen Aufwand für die technische Integration und für Aktualisierungen erfordern. Im Projektverlauf gab es für zwei der verwendeten Bibliotheken (Rasa und spaCy) Release-Wechsel, die umfassende und zeitraubende Nacharbeiten im Code erforderten. Hier ist auch bei zukünftigen Projekten mit entsprechenden Aufwänden zu planen.

Zudem wurde deutlich, dass die Repräsentation der Wissensbestände, auf die der Chatbot Zugriff hat, inhaltlich und zeitlich sehr aufwändig ist. Weiterhin muss das Interaktionsdesign aufwändig an jedes geplante Anwendungsszenario angepasst werden. Dieser Sachverhalt steht einer schnellen und breiten Nutzung der Chatbot-Technologie gerade in den Feldern der psychosozialen Arbeit aktuell im Weg. Hier sind die Zielgruppen und Szenarien einerseits sehr spezifisch, die finanziellen Mittel für Innovationen dagegen sehr begrenzt. Der Befund spricht daher nicht prinzipiell gegen eine weitere Nutzung, allerdings muss er bei der Kosten-Nutzen Kalkulation neuer Projekte kritisch reflektiert werden. Gleichzeitig ist hier ein großes Potenzial für zukünftige Studien zu sehen, die einfachere Administrationswege für die Wissensbestände von Chatbots entwickeln.

Für die dritte Forschungsfrage, die klären sollte, wie ein lernfähiges Assistenzsystem im Beratungsumfeld gestaltet werden kann, konnten basierend auf den aufgeführten Erkenntnissen etliche innovative Ideen entwickelt werden. Diese werden in Folgeprojekten weiter untersucht. So wird derzeit evaluiert, wie aktuelle Modelle zur Sprachverarbeitung eingesetzt werden können, um mit weniger Trainingsdaten gleichwertige oder sogar bessere Ergebnisse bei der Identifikation der Nutzer-Absichten erreicht werden können. Ein weiterer Ansatz ist die Verwendung KI-basierter Frage-Antwort-Modelle, um mehr

unstrukturierte Texte anstelle einer aufwändig zu erstellenden strukturierten Wissensbasis für die Beantwortung der Anwenderfragen zu erschließen. Dadurch könnte der Aufwand für die Implementierung neuer Chatbots massiv verringert werden.

Insgesamt kann festgehalten werden, dass das Thema Chatbot aus der Sicht der psychosozialen Arbeit über ein großes Potenzial verfügt. Zukünftige Studien sollten sich vor allem auf die Vereinfachung der Pflege des Chatbots konzentrieren und damit ein langes Leben des Chatbots mit stetiger Verbesserung der Wissensbasis sowie der Antwortperformance anzustreben. Weiterhin ist das Wissen um die gegenseitige Bedingtheit von Akzeptanz und Zielgruppe bisher noch nicht ausreichend elaboriert. Hier sollte insbesondere die Sozialwissenschaft weitere Studien planen, die es ermöglichen, Präferenzen im Interaktionsdesign systematisch prognostizieren zu können. Wenn dieses Wissen in technische Modelle für die Implementation der Chatbots integriert werden kann, steht dem breiteren Einsatz dieser Technologie in durchdachten und passenden Settings der psychosozialen Beratung nichts mehr im Wege.

Literaturverzeichnis

Aufotu, Nigar (2021): Entwicklung eines Frameworks zur semi-automatisierten Informationsextraktion für Knowledge Graphen. Master-Arbeit, Technische Hochschule Nürnberg.

DelVecchio, Devon S.; Jae, Haeran & Ferguson, Jodie L. (2019): Consumer aliteracy. *Psychology & Marketing*, 36 (2), S. 89–101.

Denk, Maximilian; Khabyuk Olexiy (2019): Wie relevant sind Chatbots als Kommunikations- und Marketinginstrument für Hochschulen? Forschungsbericht, Fachbereich Wirtschaft der FH Düsseldorf

Devlin, Jacob; Chang, Ming-Wei; Lee, Kenton & Toutanova, Kristina (2018): Bert: Pre-training of deep bidirectional transformers for language understanding. *CoRR Bd. abs/1810.0*.

Engelhardt, Emily; Henrich, Marlene, Reindl, Richard; Weinhardt, Marc; Zauter, Sigrid & Dietrich, Christina (2019). Beratungsbedingte Internetnutzung. Welche Dienste und Angebote nutzen Beratungsfachkräfte privat und beruflich? *e-beratungsjournal* (15/1).

Eichenberg, Christiane & Brähler, Elmar (2013): Das Internet als Ratgeber bei psychischen Problemen. Bevölkerungsrepräsentative Befragung in Deutschland. *Psychotherapeut*, 58 (1), S. 63–72.

Grandeit, Philipp; Haberkern, Carolyn; Lang, Maximiliane; Albrecht, Jens; Lehmann, Robert (2020): Using BERT for Qualitative Content Analysis in Psychosocial Online Counseling. Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science, <https://doi.org/10.18653/v1/2020.nlpccs-1.2>.

Gröning, Katharina (2011): *Pädagogische Beratung: Konzepte und Positionen*: Springer-Verlag.

Kaiser, Markus; Buttkeireit, Aline-Florence & Hagenauer, Johanna (2019): *Journalistische Praxis: Chatbots. Automatisierte Kommunikation im Journalismus und in der Public Relation (essentials)*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-25494-0>.

Kreidenweis, Helmut & Dilger, Corvin (2012). IT von unten—Eine Rekonstruktion des Blicks von Fachkräften auf die Informationstechnologie in sozialen Organisationen. *Sozialmagazin: die Zeitschrift für soziale Arbeit*, (1), 18–22.

Lehmann, Robert; Albrecht, Jens; Zauter, Sigrid (2020): CaSoTex: Computerunterstützte Analyse Sozialwissenschaftlicher Texte. Forschungsbericht, Technische Hochschule Nürnberg. <https://doi.org/10.34646/thn/ohmdok-619>.

Loh, Kep Kee & Kanai, Ryota (2016): How has the Internet reshaped human cognition? *The Neuroscientist*, 22 (5), S. 506–520.

Maurer, Hermann (2015): Does the internet make us stupid? *Communications of the ACM*, 58 (1), S. 48–51.

McTear, Michael; Griol, David & Callejas, Zoraida (2016): *The conversational interface*. [Cham]: Springer International Publishing.

Medienpädagogischer Forschungsverbund Südwest (Hrsg.) (2017): *JIM 2017, Jugend, Information, (Multi-) Media Basisstudie zum Medienumgang 12- bis 19-Jähriger in Deutschland*. HYps://www.mpfs.de/fileadmin/files/Studien/JIM/2017/JIM_2017.pdf (abgerufen am 20.02.2019.).

Noack, Jonas (2021): Semi-automatische Erstellung von Wissensdatenbanken für Question-Answering-Systeme. Bachelor-Arbeit, Technische Hochschule Nürnberg.

Rajpurkar, Pranav; Zhang, Jian; Lopyrev, Konstantin; Liang, Percy (2016): SQuAD: 100,000+ Questions for Machine Comprehension of Text. arXiv:1606.05250

Shevat, Amir (2017): *Designing Bots: Creating Conversational Experiences*: O'Reilly Media, Inc.

Venkatesh, V. & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39(2), 273–315. <https://doi.org/https://doi.org/10.1111/j.1540-5915.2008.00192.x>

Venkatesh, Viswanath; Thong, James Y. L. & Xu, Xin (2012): Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36 (1), S. 157–178.

Wenzel, Joachim (2013): Wandel der Beratung durch neue Medien. Göttingen: V&R unipress.

Anhang

Anhang 1: Bildung der Untersuchungsskalen.

Variable/ Skala			
Pretest-Fragen zur Demographie			
Was ist Ihr Geschlecht?	Männlich: 3 Weiblich: 24 Divers: 0		Gültig: 27
Wie alt sind Sie?	M = 23	SD = 3,94	Gültig: 26
In welchem Fachsemester befinden Sie sich?	1. – 2. Semester: 14 3. – 4. Semester: 2 5. – 6. Semester: 4 7. – 8. Semester: 5 9. Semester oder höher: 2		Gültig: 27
Pretest-Fragen zur Vorerfahrung			
Wissen Sie, was ein Chatbot ist?	Ja: 16 Nein: 11		Gültig: 27
Hatten Sie bereits wissentlich Kontakt mit einem Chatbot?	Ja: 12 Nein: 15		Gültig: 27
Falls ja, würden Sie diesen Kontakt als positiv, neutral oder negativ beschreiben?	Positiv: 6 Neutral: 9 Negativ: 0		Gültig: 15
Ich schätze mich selbst als technikaffin ein	M = 2,96	SD = 0,98	Gültig: 27
Wahrgenommener Nutzen			
Skala WN1*	M = 4,20	SD = 0,90	Gültig: 24
Skala WN2*	M = 4,30	SD = 0,47	Gültig: 24

Chatbots haben mir bereits vor diesem Test bei Fragen erfolgreich geholfen.	M = 2,89	SD = 1,23	Gültig: 18
ZeoBot hat reibungslos funktioniert.	M = 4,00	SD = 0,96	Gültig: 27
ZeoBot reagiert in angemessener Weise, wenn er Fragen nicht versteht.	M = 4,00	SD = 0,96	Gültig: 27
Wahrgenommene, einfache Bedienbarkeit			
Ich empfinde die Bedienung von ZeoBot als einfach.	M = 4,67	SD = 0,55	Gültig: 27
Die Verwendung von ZeoBot macht mir Spaß.	M = 3,89	SD = 0,75	Gültig: 27
Ich empfinde es als positiv, dass ich ZeoBot ohne Installation einer zusätzlichen Software wie bspw. einer App nutzen kann.	M = 4,96	SD = 0,20	Gültig: 26
Welchen Kanal zum Abruf von ZeoBot würden Sie präferieren? Bitte ordnen Sie den Facebook-Messenger, WhatsApp und die Webseite des Instituts für E-Beratung nach ihrer Präferenz.			
Webseite des Instituts für E-Beratung	M = 1,35	SD = 0,56	Gültig: 26
Kanäle: WhatsApp	M = 1,92	SD = 0,74	Gültig: 26
Kanäle: Facebook-Messenger	M = 2,68	SD = 0,56	Gültig: 25
Einstellung zur Nutzung			
Skala EN*	M = 3,41	SD = 0,56	Gültig: 21
Persönlichkeit: Ich begrüße, dass ZeoBot mich siezt, anstatt mich zu duzen.	2,85	1,29	Gültig: 27
Intention zur Nutzung			

Skala IN*

M = 4,08

SD = 0,73

Gültig: 24

*Legende der verendeten Skalen:

Skala WN1 – szenariobezogener wahrgenommener Nutzen - Verwendung folgender Items:

- Ich bin zufrieden mit den von ZeoBot bereitgestellten Informationen zur Beantwortung meiner Frage in Bezug auf Szenario 1-3/1-5.
- Ich denke, die Verwendung von ZeoBot bietet ausführlichere Antworten bei Standardanfragen als die Suche auf der Webseite des Instituts für Onlineberatung in Bezug auf Szenario 1-3/1-5.

Skala WN2 – allgemeiner wahrgenommener Nutzen - Verwendung folgender Items:

- Ich empfinde die Nutzung von ZeoBot als modern.
- Die Verwendung von ZeoBot passt zu meinem Lebensstil.
- Ich denke, ZeoBot kann mir nachgefragte Informationen zum Zertifikatskurs zusenden.
- Ich denke, die Verwendung von ZeoBot bei Standardanfragen ist schneller als der Kontakt zur Ansprechpartnerin des Zertifikatskurs.
- Ich denke, die Verwendung von ZeoBot bei Standardanfragen ist schneller als die Suche auf der Webseite des Instituts für Onlineberatung.

Skala EN - Verwendung folgender Items:

- Invertiert: Ich bin besorgt über das Versenden von persönlichen Informationen im Gespräch mit ZeoBot.
- Invertiert: Ich befürchte, dass die von mir per ZeoBot übermittelten Nachrichten missbraucht werden könnten.
- Persönlichkeit: Ich empfinde die Persönlichkeit von ZeoBot als sympathisch.
- Persönlichkeit: Ich empfinde es als positiv, dass ZeoBot eine Persönlichkeit besitzt.
- Persönlichkeit: Ich fände es schön, wenn mich ZeoBot mit meinem Namen ansprechen würde. Bzw.: Ich empfinde die personalisierte Ansprache des MKM-Chatbots als positiv.
- Persönlichkeit: ZeoBot hat auf eine angemessene Art und Weise mit mir kommuniziert.
- Persönlichkeit: Ich hatte das Gefühl mit einer realen Person zu kommunizieren.

Skala IN - Verwendung folgender Items:

- Wenn ein Chatbot auf der Webseite des Instituts für E-Beratung zur Verfügung stehen würde, dann würde ich ihn benutzen.
- Ich würde Kommiliton*innen die Verwendung von Chatbots zur Beantwortung von Studierenden-Fragen empfehlen.

Anhang 2: Mittelwertsvergleiche der Akzeptanzstudien

Variable/ Skala	SW	MKM	M-W-U	p
Pretest-Fragen				
Wissen Sie, was ein Chatbot ist?	M = 1,41 SD = 0,50 Gültig: 27	M = 1,14 SD = 0,36 Gültig: 21	208,50	<0,05
Hatten Sie bereits wesentlich Kontakt mit einem Chatbot?	M = 1,56 SD = 0,51 Gültig: 27	M = 1,38 SD = 0,50 Gültig: 21	234,00	n.s.
Falls ja, würden Sie diesen Kontakt als positiv, neutral oder negativ beschreiben?	M = 1,60 SD = 0,51 Gültig: 15	M = 1,50 SD = 0,85 Gültig: 14	85,50	n.s.
Wahrgenommener Nutzen				
Skala WN1*	M = 4,20 SD = 0,90 Gültig: 24	M = 4,21 SD = 0,57 Gültig: 21	215,00	n.s.
Skala WN2*	M = 4,30 SD = 0,47 Gültig: 24	M = 4,50 SD = 0,44 Gültig: 21	-1,50	n.s.
Chatbots haben mir bereits vor diesem Test bei Fragen erfolgreich geholfen.	M = 2,89 SD = 1,23 Gültig: 18	M = 3,86 SD = 1,35 Gültig: 14	71,50	<0,05
ZeoBot hat reibungslos funktioniert.	M = 4,00 SD = 0,96 Gültig: 27	M = 4,43 SD = 0,87 Gültig: 21	204,00	n.s.
Wahrgenommene, einfache Bedienbarkeit				
Ich empfinde die Bedienung von ZeoBot als einfach.	M = 4,67 SD = 0,55 Gültig: 27	M = 4,67 SD = 0,73 Gültig: 21	269,00	n.s.
Die Verwendung von ZeoBot macht mir Spaß.	M = 3,89 SD = 0,75 Gültig: 27	M = 4,14 SD = 1,01 Gültig: 21	223,50	n.s.
Ich empfinde es als positiv, dass ich ZeoBot ohne Installation einer zusätzlichen Software wie bspw. einer App nutzen kann.	M = 4,96 SD = 0,20 Gültig: 26	M = 4,95 SD = 0,22 Gültig: 21	270,50	n.s.

Einstellung zur Nutzung				
Skala EN*	M = 3,41	M = 4,35		
	SD = 0,56	SD = 0,46	-5,96	<0,05
	Gültig: 21	Gültig: 21		
Intention zur Nutzung				
Skala IN*	M = 4,08	M = 4,48		
	SD = 0,73	SD = 0,64	174,50	n.s.
	Gültig: 24	Gültig: 21		

*Legende der verwendeten Skalen:

Skala WN1 - Verwendung folgender Items:

- Ich bin zufrieden mit den von ZeoBot bereitgestellten Informationen zur Beantwortung meiner Frage in Bezug auf Szenario 1-3/1-5.
- Ich denke, die Verwendung von ZeoBot bietet ausführlichere Antworten bei Standardanfragen als die Suche auf der Webseite des Instituts für Onlineberatung in Bezug auf Szenario 1-3/1-5.

Skala WN2 - Verwendung folgender Items:

- Ich empfinde die Nutzung von ZeoBot als modern.
- Die Verwendung von ZeoBot passt zu meinem Lebensstil.
- Ich denke, ZeoBot kann mir nachgefragte Informationen zum Zertifikatskurs zusenden.
- Ich denke, die Verwendung von ZeoBot bei Standardanfragen ist schneller als der Kontakt zur Ansprechpartnerin des Zertifikatskurs.
- Ich denke, die Verwendung von ZeoBot bei Standardanfragen ist schneller als die Suche auf der Webseite des Instituts für Onlineberatung.

Skala EN - Verwendung folgender Items:

- Invertiert: Ich bin besorgt über das Versenden von persönlichen Informationen im Gespräch mit ZeoBot.
- Invertiert: Ich befürchte, dass die von mir per ZeoBot übermittelten Nachrichten missbraucht werden könnten.
- Persönlichkeit: Ich empfinde die Persönlichkeit von ZeoBot als sympathisch.
- Persönlichkeit: Ich empfinde es als positiv, dass ZeoBot eine Persönlichkeit besitzt.
- Persönlichkeit: Ich fände es schön, wenn mich ZeoBot mit meinem Namen ansprechen würde.
Bzw.: Ich empfinde die personalisierte Ansprache des MKM-Chatbots als positiv.
- Persönlichkeit: ZeoBot hat auf eine angemessene Art und Weise mit mir kommuniziert.
- Persönlichkeit: Ich hatte das Gefühl mit einer realen Person zu kommunizieren.

Skala IN - Verwendung folgender Items:

- Wenn ein Chatbot auf der Webseite des Instituts für E-Beratung zur Verfügung stehen würde, dann würde ich ihn benutzen.
- Ich würde Kommiliton*innen die Verwendung von Chatbots zur Beantwortung von Studierenden-Fragen empfehlen.