



Die Verarbeitung natürlicher Sprache durch Maschinen (Natural Language Processing), wie Mensch-Chatbot-Interaktion, wird für universitäre Beratungsangebote immer wichtiger. Der Artikel beschäftigt sich entsprechend mit der Frage, inwiefern Chatbots Ratsuchende in schreibdidaktischen Beratungssituationen unterstützen können. Anhand eines ersten Gehversuchs, dem Praxisbeispiel „StefaN - Schreibtraining für alle Neugieren“, werden Anforderungen von Chatbots in der Schreibberatung exploriert und es wird ein Problemaufriss gegeben. Die Beschreibung des Chatbots FragBeLa®, der Beratungsschatbot des Zentrums für Lehrerbildung, zeigt dazu Lösungsansätze auf. Er bildet einen Ausgangspunkt für den schreibdidaktischen Nachfolge-Chatbot „String2“.

Schlagworte: JoSch; NLP; Mensch-Maschine-Interaktion; Chatbot; Schreibberatung; Studienberatung

Zitiervorschlag: Dohmen, Sören; Geisler, Andrea; Holste, Alexander (2022). Chatbots für die Schreibdidaktik: zwei Praxisbeispiele. JoSch, 1(22), 68-79. <https://doi.org/10.3278/JOS2201W068>

E-Journal Einzelbeitrag
von: Sören Dohmen, Andrea Geisler, Alexander Holste

Chatbots für die Schreibdidaktik: zwei Praxisbeispiele

aus: Ausgabe 23: Schreibzentrumsarbeit und Schreibdidaktik im Zeitalter der Digitalität (JOS2201W)
Erscheinungsjahr: 2022
Seiten: 68 - 79
DOI: 10.3278/JOS2201W068

Chatbots für die Schreibdidaktik: zwei Praxisbeispiele

Sören Dohmen, Andrea Geisler & Alexander Holste

Abstract

Die Verarbeitung natürlicher Sprache durch Maschinen (*Natural Language Processing*), wie Mensch-Chatbot-Interaktion, wird für universitäre Beratungsangebote immer wichtiger. Der Artikel beschäftigt sich entsprechend mit der Frage, inwiefern Chatbots Ratsuchende in schreibdidaktischen Beratungssituationen unterstützen können. Anhand eines ersten Gehversuchs, dem Praxisbeispiel „StefaN – Schreibtraining für alle Neugierigen“, werden Anforderungen von Chatbots in der Schreibberatung exploriert und es wird ein Problemaufriss gegeben. Die Beschreibung des Chatbots FragBeLa®, der Beratungschatbot des Zentrums für Lehrerbildung, zeigt dazu Lösungsansätze auf. Er bildet einen Ausgangspunkt für den schreibdidaktischen Nachfolge-Chatbot „Strin-g²“.

1 Der Lehramtsbot FragBeLa® – Vorbild für den Schreibdidaktikbot Strin-g²

Die Verarbeitung natürlicher Sprache durch Maschinen, sogenanntes *Natural Language Processing*, und Mensch-Maschine-Interaktion finden immer mehr Einsatz in Universitäten, beispielsweise um Klausurantworten von Studierenden inhaltlich auszuwerten – sogenanntes Automated Content Scoring (Horbach/Zesch 2019) –, um Studierende mithilfe eines Chatbots zu beraten (Dohmen/Geisler/Pohlmann 2020) oder um Fachtexte für Studierende mittels DeepL oder GoogleTranslate zu übersetzen – Letzteres mehr schlecht als recht (Büttner 2021: 97). Für Schreibberatungen und schreibdidaktische Settings stellt sich die Frage, inwiefern Chatbots Ratsuchende in der Beratung und Kompetenzvermittlung unterstützen können, indem sie deren Fragen zu Schreibprozess und -produkt beantworten. Dieser Praxisbeitrag geht dieser Frage anhand zweier Beispiele nach: zum einen anhand des Chatbots „StefaN – Schreibtraining für alle Neugierigen“, der einen ersten Gehversuch in der deutschsprachigen Schreibdidaktik und Schreibberatung darstellte. Er stand von Februar 2021 bis Oktober 2021 als frei verfügbares Tool im Netz; seine Architektur wurde mithilfe eines kommerziellen Plattformanbieters betrieben. Zum anderen FragBeLa® (Beratung Lehramt), der Beratungs-Chatbot des Zentrums für Lehrerbildung (ZLB), der seit der Orientierungswoche 2019 online ist und als interaktives, webbasiertes Dialogsystem für den Anwendungsfall „Standardfragen in der Studienberatung“ fungiert. Forschungsmethodisch stellt dieser Beitrag also keinen Anspruch auf Allgemeingültigkeit, sondern möchte Anforderungen an Chatbots in der Schreibberatung explorieren. Konkret

arbeitet die Darstellung eines Dialogs mit StefaN Probleme heraus, für die FragBeLa® Lösungsansätze aufzeigt.

Ein Beispieldialog mit StefaN in Kapitel 3 zeigt Probleme und Fehler in der Struktur und Anwendung des Bots auf, die Ausgangspunkt und Motivation für die Neuauflage als „Strin-g² – Schreibtraining intelligent gestaltet und genutzt“ darstellen. Bei Strin-g² handelt es sich um einen Projektentwurf, der plant, einen Chatbot für eine Teilaufgabe in einem schreibdidaktischen Setting einzusetzen – für Fachlehrende (Schreibtraining intelligent gestaltet) und für Studierende (Schreibtraining intelligent genutzt). FragBeLa® wird nicht für schreibdidaktische Zwecke eingesetzt. Kapitel 4 beschreibt das Tool und dessen Einsatz. Diese Beschreibung gibt jedoch Hinweise, die sich Schreibzentrumsarbeit und Schreibdidaktik zunutze machen können. Daher fokussiert dieser Beitrag vor allem FragBeLa®. Vorab vermittelt Kapitel 2 Grundlagen zu *Natural Language Processing*, um die Beispiele besser einordnen zu können.

2 Chatbots und *Natural Language Processing*

Grundsätzlich lassen sich sprachverarbeitende Systeme bzw. Maschinen (*Natural Language Processing*) in zwei Arten unterteilen (Jurafsky/Martin 2020: 168 ff.): in regelbasierte Systeme und in datenbasierte. Letztere sind die Grundlage für sogenannte künstliche Intelligenzen (KI), in denen die Maschine durch selbstständiges Lernen Algorithmen erstellt. Grundlage für dieses maschinelle Lernen (*machine learning*) ist eine breite Basis an natürlichen Sprachrealisaten von Menschen, also ein Korpus. Allerdings unterliegen die sogenannten KIs etlichen Realisaten und funktionieren als Black Box. D.h., Maschinenentwickler*innen können nicht unmittelbar in die Maschine eingreifen, sondern nur sehr bedingt Einfluss darauf nehmen, wie und welche Ausgabe an Sprachrealisaten die Maschine produziert. Fehler in der Sprachausgabe einer Maschine können also nur schwer behoben werden. Irreführend ist auch die Metapher „Künstliche Intelligenz“, die evoziert, dass Maschinen über menschliche Intelligenz verfügten. Dem ist aber nicht so, weil Maschinen die Fähigkeit fehlt, komplexes Weltwissen in den Dialog mit dem Menschen einzubinden (Lotze 2016: 70 ff.) – man spricht daher von Mensch-Maschine-Interaktion, nicht von Kommunikation:

„Nüchtern betrachtet handelt es sich bei allem, was wir derzeit zur Künstlichen Intelligenz erleben, um sogenannte ‘schwache KI’ (weak AI). Die Maschinen und Systeme sind nicht in echt intelligent. Sie können nicht wirklich lernen, wir nennen das nur so. Sie denken nicht. Sie entscheiden nicht. Sie verstehen nicht, was wir sagen, und auch nicht, was sie selbst an Sprachausgaben produzieren.“ (Wendland 2020: 228)

Maschinenentwickler*innen können im Gegensatz zu den Arbeitsprozessen von KIs in die Prozesse regelbasierter Maschinen unmittelbar eingreifen: Bei diesen Systemen legen die Entwickler*innen die Algorithmen genau fest, nach denen die Maschine Sprache verarbeitet und ausgibt. Fehler lassen sich unmittelbar beheben. Maschinelles Lernen

kann in der Form stattfinden, dass die Entwickler*innen dokumentierte Nutzer*innen-Dialoge analysieren und in die Datenbank einarbeiten, auf die die Maschine zurückgreift. Ebenfalls können die Maschinenentwickler*innen auf der Grundlage von Nutzer*innen-Dialogen auch die bereits programmierten Algorithmen modifizieren, die festlegen, wie die Maschine auf die Datenbank zugreift. Die genaueren Arbeitsweisen der beiden Maschinenarten zeigen die Beschreibung von Beispiel 1 und die sieben Techniken von FragBeLa® auf: Dabei basiert die Maschine StefaN auf einer KI, während FragBeLa® regelbasiert arbeitet.

3 Praxisbeispiel 1: Das Scheitern des Schreibdidaktik-Bots StefaN

Mit *ArgueTutor* (Wambsganss/Küng/Söllner/Leimeister 2021) wurde eine Innovation für die deutschsprachige Schreibdidaktik geschaffen: Der Chatbot gibt allgemeine Hinweise zum Argumentieren in Texten. Er behandelt allerdings keine Fragen zur Argumentation einer individuellen Haus- oder Abschlussarbeit. Deren roter Faden wird bekanntlich nicht nur vom jeweiligen Thema und der Fragestellung/Hypothese, sondern auch von der Fächertradition bestimmt. Unseres Wissens existiert im deutschsprachigen Raum noch kein funktionierender Chatbot für allgemeine Schreibberatungen. Ausgehend von einem semantischen Netz und einer Textbausteinsammlung (z. B. Witze), auf deren Basis die Spracherkennung und Sprachausgabe der Maschine StefaN lief, hätte die angemietete, kommerzielle Plattform die technische Möglichkeit geliefert, StefaN bei einer ausreichenden Anzahl an Nutzenden-Dialogen zu einer funktionierenden KI zu erweitern. Der Chatbot wurde als Testballon auf meiner Homepage unabhängig von einem unmittelbaren didaktischen Kontext wie einem Seminar betrieben. Ich (AH) habe daraufhin Personen um ihre Einschätzung gebeten, die im Bereich der Schreibberatung und -forschung, der Mediendidaktik und des NLP als Expert*innen gelten: Vorstandsmitglieder der gefsus, die Redaktion von eteaching.org und einen Lehrstuhlinhaber für maschinelle Sprachverarbeitung. Um neben diesen Einschätzungen von Expert*innen auch Rückmeldungen von realen Nutzer*innen zu erhalten, wurden auch Teilnehmende meiner schreibdidaktischen Seminare gebeten, das Tool zu nutzen. Die ausreichende Anzahl Nutzenden-Dialoge wurde in 260 Tagen mit exakt 20 Dialogen nicht annähernd erreicht und das folgende Beispiel soll zeigen, inwiefern das Projekt gescheitert ist und welche Implikationen daraus für das Folgeprojekt Strin-g2 entstehen (rechts die Frage des*der Nutzenden, links die Antworten des Chatbots):¹

1 Das Interface von StefaN wird hier nicht dargestellt, um die kommerziell betriebene Plattform zu anonymisieren.

Abbildung 1

Protokoll zum Chatdialog mit StefaN

Hier finden Sie das Chat-Protokoll von www.alexander-holste.de. (Stand: 05.10.2021 12:34)

Wie füge ich Seitennummern in ein Word-Dokument ein?
05.10.2021 12:31

Hier kann ich scheinbar nicht weiterhelfen. Bitte wende dich an
einen meiner menschlichen Kollegen.!

Was ist besser? a) Im Folgenden wird beschrieben... b) Im
Folgenden beschreibe ich...
05.10.2021 12:32

Freut mich.

Kann ich dir noch eine Frage zu unseren Produkten oder zu
unserem Service beantworten?

Warum freut sich eine Blondine so, wenn sie ein Puzzle nach 6
Monaten fertig hat?
"Weil auf der Packung steht: 2-4 Jahre."

Dieses Beispiel veranschaulicht mehrere Anforderungen und Probleme, die beim Einsatz eines Chatbots in der Schreibdidaktik eine Rolle spielen; wie angegeben, dient das Beispiel der Exploration des noch unbekannten Anwendungs- bzw. Forschungsfeldes (s. Abb.1):

1. Ein Chatbot kann bei Standardfragen weiterhelfen, die sich für Schreibende nicht bei der Suche nach wissenschaftlicher Erkenntnis ergeben. Dazu gehören Fragen zur Textoberfläche, also zu Rechtschreibung, Zeichensetzung, Grammatik, teils zu Zitation und Layout wie auch zur Bedienung von Software wie obige Frage zum Einfügen von Seitenzahlen. Menschliche Berater*innen hätten damit mehr Zeit, um Fragen zur individuellen Themenstruktur, Kohärenz und weiteren Elementen der textuellen Tiefenstruktur einer Haus-/Abschlussarbeit oder zu individuellen Schreibprozessen zu beantworten.
2. Maschinen können derart auf die Kooperation mit menschlichen Beratenden programmiert sein, dass Ratsuchende einen Dialog übernehmen, sobald die Maschine ihre Grenzen im Gesprächsverlauf erreicht.
3. Menschen können einer Maschine ihre Fragen zum Schreiben auf zwei Ebenen stellen, was für Themen zur Studienberatung unwahrscheinlich ist. So können Fragende auf der Objektebene um Auskunft zur Richtigkeit, Angemessenheit bzw. Konventionalität einer konkreten Formulierung bitten (*Was ist besser? Im Folgenden wird ...*). Menschen können aber auch Fragen stellen, die vom Einzelfall abstrahieren (*Wie füge ich Seitennummern in ein Word-Dokument ein?*). In den Chatprotokollen von StefaN sind auch Fragen dokumentiert wie: *Wann setzt man ein Komma beim Infinitiv?* Für die Architektur der Maschine stellt sich den Entwickler*innen damit die Frage, ob die Nutzenden

bereits über eine Metasprache zum Schreiben und zu Sprache (beispielsweise *Infinitiv*) verfügen. Denn daraus resultiert, wie die Antwortbausteine formuliert werden müssen (Ist Nutzenden z. B. bekannt, was eine Infinitivgruppe ist, oder muss dies veranschaulicht werden?). Konkrete Sprachbeispiele (*Im Folgenden wird ...*) zu erkennen und als richtig oder falsch zu bewerten, ist bereits durch Tools wie den Duden-Korrektor und teils durch Word möglich. Die textsortenspezifische Konventionalität von Formulierungen einzuschätzen, wird wohl vorerst (und vermutlich lange) menschlichen Berater*innen überlassen bleiben. Fragen zu konkreten Sprachrealisaten auf der Objektebene wurden auch in anderen Fällen gestellt.

4. Witze (eine Form von *Eastereggs*) wie der unpassende Blondinenwitz im Beispiel werden die Akzeptanz der Maschine durch die Ratsuchenden weniger erhöhen als vielmehr verringern. Dafür fanden sich im Nutzer*innen-Korpus mehrere Beispiele.
5. Das Beispiel verrät, dass der Bot eine relativ frequente Frage wie das Einfügen von Seitenzahlen nicht beantworten kann. Technisch wäre dies aber durchaus möglich. Grund für die fehlende Antwort könnte zum einen sein, dass in der Datenbank der Maschine kein passender Textbaustein hinterlegt ist, der diese Frage erfasst. Da es sich bei wissenschaftlichem Schreiben um einen recht komplexen Gesprächsgegenstand handelt, müssten für eine gut funktionierende Maschine sehr viele Textbausteine hinterlegt sein, um die häufigsten Fragen zu erfassen. StefaNs Textbausteinsammlung scheint einfach zu klein und auf zu wenige Themen bezogen zu sein. Zum anderen könnte im obigen Beispiel das Parsen und Prompten misslungen sein: Das Zerlegen einer natürlichsprachlichen Frage in Keywords wird als Parsing bezeichnet, deren Zuordnung zu Patterns durch Pattern-Matcher und die Ausgabe einer Antwort als Prompting (Lotze 2016: 32 ff.; weiteres s. Kapitel 4.2). Dieses Misslingen könnte auf zu wenige Nutzer*innen-Fragen – 20 Anfragen – zurückgeführt werden, um die Maschine auf deren Grundlage zu trainieren.

Auf der Grundlage des obigen Beispiels lassen sich zwar keine allgemeingültigen Aussagen über schreibdidaktische Chatbots treffen. Der Blick über den Tellerrand in den Bereich der Studienberatung für Lehramtsstudierende soll aber Hinweise darauf liefern, wie die in obigem Beispiel aufgezeigten Probleme möglicherweise gelöst werden könnten.

4 Praxisbeispiel 2: Einsatz des Chatbots FragBeLa® in der Studienberatung

Die Idee, einen Bot für Standardfragen im Lehramtsstudium einzusetzen, entstand im Jahr 2016. BeLa ist seit der Orientierungswoche im Oktober 2019 auf der Website des Zentrums für Lehrerbildung und seit Dezember 2020 ins LehramtsWiki eingebunden. Die Nutzer*innen stellten BeLa 7.085 verschiedene Fragen, häufig zu den Praxisphasen und Studienprojekten, zum Sprachassessment der Universität Duisburg-Essen SkaLa, zu Beratungen und aktuell zu Corona. Oft wird sie auch nach dem Wetter oder einem Witz,

also Easteregg, gefragt. Ein Viertel aller Fragen an FragBeLa® werden außerhalb der üblichen Bürozeiten gestellt. Die Maschine FragBeLa® ergänzt die menschliche Beratung und entlastet Berater*innen.

4.1 Anforderungen an FragBeLa®

Zu Beginn des Projektes wurde festgelegt, dass der Chatbot fünf Anforderungen erfüllen soll: 1. Seine Reaktionsgeschwindigkeit und die Qualität seiner Antworten sollen bei den Nutzenden das Gefühl eines echten Gesprächs ermöglichen. 2. Der Chatbot soll passende, produktive Rückfragen stellen können. 3. Die Maschine soll es ermöglichen, dass Entwickler*innen den Chatbot um neues Wissen erweitern und vorhandenes Wissen aktuell halten können. 4. Der Chatbot soll gute Ergebnisse für die Nutzersprache Deutsch erzielen und auch auf Fragen antworten, die Tippfehler oder eine fehlerhafte Grammatik enthalten oder die nur in Stichworten gestellt werden. 5. Die Maschine soll „Antworten auf Fragen, die Nutzer*innen Chatbots gerne aus Spaß stellen“, sogenannte Eastereggs, geben, um die Akzeptanz der Maschine und die ihrer Antworten bei den Nutzer*innen zu erhöhen (Campbell et al. 2009).

Weil die vorhandenen Beschreibungssprachen und Systemarchitekturen diese Anforderungen nicht erfüllten, fiel die Entscheidung, ein eigenes System zu entwickeln.

Abbildung 2

Avatar von BeLa, der Eule (ZLB 2021: URL)



Die visuelle Erscheinung des Chatbots stellt neben der gelungenen Verarbeitung natürlicher Sprache ein weiteres Aufgabenfeld für Entwickler*innen dar, um die Akzeptanz der Maschine durch die Nutzer*innen zu erhöhen. So dient der Avatar – eine Eule (s. Abb. 2) – von BeLa als visueller Einstiegspunkt für Nutzer*innen auf der ZLB-Website. Denn ein Großteil der Nutzer*innen kennt die Eule BeLa bereits aus anderen Kontexten (Social-Media-Auftritte des ZLB und LehramtsWiki). Auch auf der LehramtsWiki-Website können Nutzer*innen mit BeLa Kontakt aufnehmen. Ausgangspunkt für diese Form der visuellen Kommunikation war die Erkenntnis, dass Nutzer*innen Avatare gut wiedererkennen, die auf Comicfiguren von Tieren basieren und eine sympathische und freundliche Ansprechperson verkörpern (Langner 2018). Für die Akzeptanz des Chatbots durch die Nut-

zer*innen ist es wichtig, dass der Bot eine Persönlichkeit erhält (Neupert 2018), zu der auch das Bild des Avatars gehört.

4.2 Die sieben Techniken von FragBeLa®

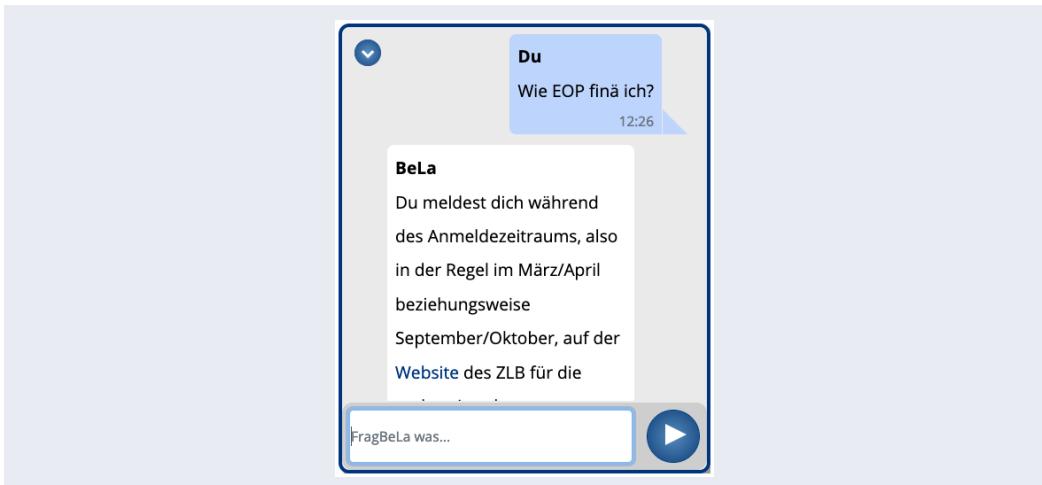
Der Einsatz von Chatbots in der Studienberatung wird immer weiter erforscht (Dohmen/Geisler 2021; Geisler/Dohmen/Pohlmann 2020; Hartmann 2021). Folgende sieben Techniken ermöglichen es FragBeLa®, auch „schlecht“ gestellte, also für Maschinen eher ungeeignete, Fragen zu beantworten. Sie liefern einen Lösungsansatz für die Probleme, die in obigem Beispieldialog mit StefaN auftraten (s. Kap. 2):

1. Das Verfahren, auf dessen Grundlage BeLa eine Antwort gibt, ist abgestuft. Zunächst wird nach einem Treffer der antizipierten Frage gesucht (Parsing). Dann kann BeLa bereits antworten (Prompting). Das ist die schnellste Möglichkeit, Nutzer*innen eine Antwort zu geben. Deshalb schlägt BeLa den Nutzer*innen während der Eingabe Fragen vor, die zu den ersten Wörtern ihrer Eingabe passen. BeLa kann auch Fragen beantworten, die sprachlich nicht so gestaltet sind, wie das Team sie erwartet hat. Dazu werden die folgenden Techniken angewandt.
2. FragBeLa® kann Vertipper erkennen und korrigieren (s. Abb. 3).
3. Mittels der Technik des Stemming [6] wird die Flexion der Wörter entfernt (z. B. wird *prüfen* zu *pruf* transformiert). Dies reduziert ein Wort nicht auf die grammatischen Stammform, ist bisher aber für das Chaterlebnis ausreichend. Der Algorithmus ist effektiver, als die korrekte Form in langen Listen nachzusehen, was geschehen müsste, um Wörter auf ihre grammatischen Stammform zu reduzieren.
4. Die Bestandteile von Komposita wie *Prüfungsanmeldung* (*Prüfung* und *Anmeldung*) werden erkannt. Denn es kann vorkommen, dass die Maschine zu diesem komplexen Suchbegriff nicht direkt antworten kann. Aber es ist möglich, dass zu einem Bestandteil des Kompositums (*Prüfung*) eine Antwort vorliegt, die ebenfalls für die Fragestellenden ausreichend ist.
5. Partikelverben mit trennbarem Präfix wie *anmelden* werden erkannt und wieder mit ihrem Präfix zusammengeführt. Dies ist relevant, weil beispielsweise *anmelden*, das auch als *melde ... an* in Fragen der Nutzer*innen geschrieben werden kann, eine andere Bedeutung hat als *melden* ohne die Partikel *an*.
6. Nutzer*innenfragen werden auf ihre charakterisierenden Teile reduziert, indem sogenannte Stopwörter getilgt werden und die verbleibenden Wörter (Feature) gefiltert werden. Auf diese Weise ist die Satzstellung der Wörter relativ und es können auch Fragen mit falschem Satzbau erkannt und beantwortet werden, wie beispielsweise die Frage *Was ich machen muss, wenn krank in EOP?* BeLa erkennt und beantwortet diese Frage korrekt.
7. Die Maschine kann frei gestellte Fragen besser beantworten, wenn dem System synonym verwendete Wörter für die Feature bekannt sind. BeLa beantwortet beispielweise auch die Frage *Was ist, wenn ich im EOP erkältet bin?*, ohne dass die Frage zuvor expli-

zit ins System eingegeben wurde. Die Frage wird über die Synonyme (*EOP, krank*) identifiziert und danach wird die passende Antwort geliefert.

Abbildung 3

Fiktiver Nutzer*innen-Dialog mit BeLa zur Veranschaulichung von fehlerhafter Spracherkennung (ZLB 2021: URL)



Die Fragen der Nutzer*innen werden ausgewertet und in BeLa eingepflegt, sofern sie nicht bereits vorhanden sind. Daneben werden FAQs, die auf Webseiten von Einrichtungen der Universität Duisburg-Essen vorhanden sind, redaktionell überarbeitet, Feature, Synonyme und Themengebiete identifiziert und Antworten klassifiziert. Diese „Füttervorgänge“ sind noch zeitintensiv. Die Datenhaltung findet in einer Excel-Datei statt, die in die Graphdatenbank importiert wird. Bei jeder neu hinzukommenden Frage muss die gesamte Excel-Datei importiert werden. Um diesen Prozess zu vereinfachen, zu beschleunigen und auch anderen in der Beratung tätigen Personen an der Universität die Möglichkeit zu geben, ihr Wissen mit BeLa zu teilen, wird eine Wissenseingabemaske entworfen, die Nutzer*innen Schritt für Schritt durch den Eingabeprozess leiten soll.

4.3 Evaluation von FragBeLa®

Bereits Ende 2018 fand eine erste Testphase statt, bei der Studierende und Kolleg*innen mittels eines standardisierten Fragebogens neben vorgegebenen Szenarien auch eigene Szenarien testen konnten. Die Resonanz auf den Prototypen war bereits positiv. Weitere Evaluierungen folgten, auch mit Berater*innen anderer Universitäten und Lehrer*innen. Die Testergebnisse fließen in die Weiterentwicklung des Systems ein. Alle Tests zeigen, dass zu den vorhandenen Themenfeldern noch viel mehr Fragen gestellt werden, als im Vorfeld antizipiert wurde. Weitere Evaluationen folgen, sobald neue Themenbereiche eingefügt sind.

Zur Qualitätsverbesserung werden die Fragen der Nutzer*innen während der Verwendung des Chatbots ohne Personenbezug mitgeschnitten. Das Fragenlog des Chatbots wird vom Team regelmäßig auf neue Fragen hin untersucht und das Team überprüft, ob BeLa richtig antwortet. Wenn BeLa „sinnvolle“ – also maschinell verarbeitbare – Fragen nicht beantworten kann, dann wird die Antwort ins System aufgenommen. Nur 3,2% der gestellten Fragen entsprechen den zuvor antizipierten Frageformulierungen vollständig. Anders formulierte Fragen kann FragBeLa® dennoch beantworten, ohne dass diese manuell eingepflegt werden müssen, weil die vorher beschriebenen sieben Techniken greifen. An den Mitschnitten lässt sich auch erkennen, dass das Frageverhalten der Nutzer*innen sehr unterschiedlich ist; es reicht vom Eingeben eines oder zweier Stichwörter bis hin zu ganzen Sätzen mit förmlicher Anrede. Die Auswertungen der dokumentierten Nutzer*innen-Fragen zeigen sehr häufig Tipp-/Schreibfehler und auch grammatischen Fehler.

Fehlerhaft eingegebene Fragen (s. Abb.3) kann das System wegen der getroffenen Vorüberlegungen und Techniken dennoch beantworten. Auf diese Weise kann (zumindest in Teilen) der Eindruck eines echten Gesprächs vor einem Nachschlagewerk für FAQ überwiegen. Es werden auch Fragen gestellt, auf die BeLa bislang noch nicht antworten kann wie Öffnungszeiten von Einrichtungen, Prüfungstermine etc.

Ein kleiner Ausblick: Der Chatbot FragBeLa® soll derart programmiert werden, dass er kontextbezogen antworten kann und nicht nur Einzelfragen, sondern auch Folgefragen vor dem Hintergrund der Ausgangsfrage beantworten kann. Dazu muss die Darstellung des Wissens in der Datenbank das natürliche Sprachverhalten von Menschen besser abbilden. Abschließend ist zu klären, wie die Erkenntnisse zu FragBeLa® für Schreibberatungen und die Schreibdidaktik fruchtbar gemacht werden können.

5 Fazit

Der Artikel ist der Frage nachgegangen, inwiefern Chatbots Menschen in Schreibberatungen und bei der Kompetenzvermittlung in schreibdidaktischen Settings unterstützen können. Die Beschreibung des dokumentierten Beispieldialogs mit StefaN zeigt vorerst unabhängig von technischen Lösungsmöglichkeiten wie denen von FragBeLa®, dass der Einsatz eines Chatbots in der Schreibdidaktik losgelöst von einem didaktischen Setting kaum gelingen kann. Denn die Erwartungen an die Fähigkeiten der Maschine, ohne die Nutzer*innen auf deren Beschränkungen hinzuweisen, werden enorm sein. Enttäuschte Nutzer*innen-Erwartungen werden zwangsläufig zu mangelnder Akzeptanz des Bots führen, die aber eine wesentliche Voraussetzung für erfolgreiche Interaktion ist.

Als wesentliches Ergebnis dieses Artikels kann die Beschreibung der sieben Techniken von FragBeLa® gelten, um Probleme zu lösen, die in Dialogen mit StefaN auftreten; FragBeLa® definiert also Ansätze, die den technischen Anforderungen für den Aufbau eines schreibdidaktischen Chatbots entsprechen. So eignet sich eine regelbasierte Maschine wie FragBeLa®, während KI-basierte Systeme wie StefaN die Lehrenden bzw. Berar-

tenden vor das Problem stellen, kaum Einfluss auf die Ausgabe fehlerhafter oder unangemessener Antworten nehmen zu können – z. B. auf das Erzählen von Blondinenwitzen. Dieses Problem tritt auch bei KI-basierten Chatbots und Übersetzungen von Informationsangeboten staatlicher Institutionen auf (Holste 2021). Die Evaluation zu FragBeLa® durch dessen Nutzer*innen klärt, welche Erwartungen sie an einen Chatbot in einer Beratungssituation haben. Konkret lassen sich folgende Ergebnisse aus der Beschreibung von StefaN und von FragBeLa® ableiten. Sie verdeutlichen, unter welchen Bedingungen der Nachfolge-Chatbot Strin-g² sinnvoll eingesetzt werden kann:

1. innerhalb eines schreibdidaktischen Konzepts eine klar umrissene Teilaufgabe zu übernehmen, beispielsweise nur Fragen der Ratsuchenden zu Argumentationsstrukturen eines Textes zu beantworten;
2. die Erwartungen der Nutzer*innen vor der Benutzung darauf zu beschränken, vom Chatbot keine Antworten auf Fragen zur Richtigkeit konkreter Formulierungen wie *Was ist besser? Im Folgenden werden ...* (Objektebene) zu erhalten;
3. die Erwartungen der Nutzer*innen darauf zu beschränken, vom Chatbot nur Antworten auf Standardfragen zur Textoberfläche (mithilfe von Metasprache wie *Komma beim Infinitiv*) und zur Bedienung von Textverarbeitungsprogrammen zu erhalten;
4. Anfragen von Nutzenden an eine*n menschlichen Berater*in auf Wunsch weiterzugeben, sobald die Grenzen der Maschine erreicht sind oder Menschen außerhalb der Bürozeiten nicht erreichbar sind;
5. die Maschine in einem Setting einzusetzen, in dem ausreichend viele Nutzenden-Fragen erzeugt werden (also eher 7.085 anstatt 20 Anfragen), um die Maschine trainieren zu können;
6. dokumentierte Dialoge von Nutzer*innen durch Entwickler*innen auszuwerten und manuell in die Datenbank der Maschine einzuarbeiten;
7. die Persönlichkeit der Maschine auch visuell durch den Avatar auszugestalten, um einen Wiedererkennungswert an verschiedenen Stellen im Internet zu schaffen und um die Akzeptanz der Maschine zu erhöhen.

Die Autor*innen sehen, dass Chatbots einen Beitrag leisten können, Standardfragen zu beantworten. Gerade in schwierigen Situationen schätzen viele Ratsuchende ein vertrauliches Gespräch und einen anderen Menschen, der ihnen zuhört und individuell auf sie eingehet. Dies kann kein Chatbot so leisten wie ein Mensch. Der Chatbot FragBeLa® wie auch Strin-g² soll die menschliche Studien- bzw. Schreibberatung nicht ersetzen, sondern vielmehr sollen die Berater*innen entlastet werden, Standardfragen zu beantworten. Dadurch erhielten sie mehr Zeit für Anliegen von Studierenden, die über Standardfragen hinausgehen. Daher bieten Chatbots einen Mehrwert für Ratsuchende und auch für die Institution Universität. Dies zeigen in der Praxis auch Anfragen verschiedener NRW-Hochschulen, die darum bitten, FragBeLa® vorzustellen und/oder für ihre Beratungsanliegen nutzen zu dürfen.

Literatur

- Büttner, Gesa (2021): *Dolmetschervorbereitung digital. Professionelles Dolmetschen und DeepL*. TransÜD, Band II5. Berlin: Frank & Timme.
- Campbell, Robert H./Grimshaw, Mark/Green, Gill (2009): Relational agents: A critical review. In: *The Open Virtual Reality Journal*. Vol. 1. No. 1.
- Dohmen, Sören/Geisler, Andrea (2021): Entwicklung des Chatbots FragBeLa® für Lehramtsstudierende. In: *Wirtschaftsinformatik und Management (WuM)*. I3. 444–451.
- Geisler, Andrea (Hrsg.) (2020): *Jahresbericht 2019*. Unter Mitarbeit von DuEPublico: Duisburg-Essen Publications Online, University of Duisburg-Essen.
- Geisler, Andrea/Dohmen, Sören/Pohlmann, Ronja (2020): FragBeLa®: der Beratungs-Chatbot für Lehramtsstudierende; von der Idee bis zum ersten Einsatz. In: *Zeitschrift für Beratung und Studium*. Jg. 15. Nr. 2 + 3. 89–92.
- Hartmann, Daniela (2021): Künstliche Intelligenz im DaF-Unterricht? Disruptive Technologien als Herausforderung und Chance. In: *Info DaF: Informationen Deutsch als Fremdsprache*. Jg. 48. Nr. 6. 683–696.
- Holste, Alexander (2021): Wissensstrukturen in Kunden-Chatbot-Interaktion und in automatischer maschineller Übersetzung von Homepages. In: *doctima-Blog*. Online im WWW. URL: <https://www.doctima.de/2021/11/wissensstrukturen-in-kunden-chatbot-interaktion-und-in-automatischer-maschineller-uebersetzung-von-homepages/> (Zugriff: 21.02.2022).
- Holste, Alexander (2022): Knowledge Constructing in Human-bot Interaction. Language Based Knowledge Asymmetries in the Domain of e-Government. In: Engberg, Jan/Fage-Butler, Antoinette/Kastberg, Peter (eds.): *Perspectives on Knowledge Communication: Contexts and Settings*. London: Routledge (submitted).
- Horbach, Andrea/Zesch, Torsten (2019): The Influence of Variance in Learner Answers on Automatic Content Scoring. *frontiers in Education*. Vol. 28. S. 1–15. DOI 10.3389/feduc.2019.00028
- Jurafsky, Daniel/Martin, James H. (2020): Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Online im WWW. URL: <https://web.stanford.edu/%7Ejurafsky/slp3/ed3book.pdf> (Zugriff: 13.12.2021).
- Langner, Patrick (2018): *Entwicklung und Evaluierung eines Chatbots am Beispiel der Studienberatung der HAW Hamburg*. Online im WWW. URL: https://edoc.sub.uni-hamburg.de/haw/volltexte/2018/4364/pdf/PaLanger_BA.pdf (Zugriff: 08.07.2020).
- Lotze, Netaya (2016): *Chatbots*. Eine linguistische Analyse. Band 9. Frankfurt a. M.: Peter Lang.
- Neupert, Daniela (2018): Ein Chatbot zur Beratung von Studieninteressierten und Studierenden. In: *Bachelorarbeit im Studiengang Interaktive Medien*. Online im WWW. URL: <http://michaelkipp.de/student/Neupert2018.pdf> (Zugriff: 08.07.2020).

Wambsganss, Thiemo/Küng, Tobias/Söllner, Matthias/Leimeister, Jan Marco (2021): Argue-Tutor: An Adaptive Dialog-Based Learning System for Argumentation Skills. In: CHI '21: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13. Online im WWW. URL: <https://doi.org/10.1145/3411764.3445781> (Zugriff: 21.02.2022).

Wendland, Karsten (2020): Wird die Krone der Schöpfung auf ein neues Haupt gesetzt? Bewusste KI-Systeme im Fokus technischer Entwicklungen. In: Fürst, Ronny A. (Hrsg.): *Digitale Bildung und Künstliche Intelligenz in Deutschland*. Vol. AKAD University Edition. Wiesbaden: Springer. 221–240.

ZLB – Zentrum für Lehrerbildung (2021): *Der Beratungsschatbot FragBeLa® für Lehramtsstudierende*. Online im WWW. URL: <https://zlb.uni-due.de/fragbela/> (Zugriff: 16.10.2021).

Autor*innen

Sören Dohmen, B. Sc., leitet die IT-Koordination im Zentrum für Lehrkräftebildung (Universität Duisburg-Essen). Er ist technischer Projektleiter des Beratungs-Chatbots FragBela®.

Andrea Geisler, Dr. rer. nat., leitet das Ressort Studierendenservice und Öffentlichkeitsarbeit im Zentrum für Lehrkräftebildung (Universität Duisburg-Essen). Sie ist Projektleiterin des Beratungs-Chatbots FragBela®.

Alexander Holste, Dr. phil., ist Mitglied des gefsus-Vorstandes. Er lehrt und forscht als WMA zu Wissenskommunikation sowie zu Schreibprozessen und initiiert aktuell das Chatbot-Projekt „Strin-g²⁴“.