

GeCCo 1.0 – Erstellung eines öffentlichen Datensatzes für die KI-basierte Inhaltsanalyse in der Online-Beratung

Prof. Dr.-Ing. Jens Albrecht

Fakultät Informatik, Technische Hochschule Nürnberg

Prof. Dr. Robert Lehmann

Fakultät Sozialwissenschaften/Institut für E-Beratung, Technische Hochschule Nürnberg

Aleksandra Poltermann, M.A.

Fakultät Sozialwissenschaften/Institut für E-Beratung, Technische Hochschule Nürnberg

Zusammenfassung:

Ein Einsatzfeld Künstlicher Intelligenz im Bereich der Sozialen Arbeit ist die automatische Inhaltsanalyse von Konversationen aus der psychosozialen Onlineberatung. Mögliche Anwendungen sind die statistische Untersuchung von Wirkzusammenhängen sowie die Entwicklung KI-basierter Dialog- und Unterstützungssysteme. Die Grundlage für das Training entsprechender Modelle sind aussagekräftige und realistische Trainingsdaten.

Im Rahmen des Projektes GeCCo (German e-Counseling Conversation Dataset) wurde ein erster deutschsprachiger Datensatz für die Analyse psychosozialer Beratungskonversationen erstellt und veröffentlicht. Wesentlicher Bestandteil ist ein neu entwickeltes Kategoriensystem, das eine feingranulare inhaltliche Klassifikation in 40 Berater- und 28 Klienten-Kategorien ermöglicht. Basierend auf dem Datensatz konnten erfolgreich verschiedene Modelle mithilfe maschineller Lernverfahren trainiert und für die wissenschaftliche Nutzung bereitgestellt werden.

1. Projektdaten

Fördersumme	30.000 Euro
Laufzeit	Januar bis Dezember 2023
Fakultät/Einrichtung	Fakultät Informatik in Kooperation mit dem Institut für E-Beratung
Projektleitung	Prof. Dr. Jens Albrecht, Prof. Dr. Robert Lehmann
Projektkoordination	Aleksandra Poltermann, M.A.
Assoziierte Projektmitarbeitende	Jennifer Burghardt, M.A. Carolyn Haberkern, B.A. Mara Stieler, M.A. Studentische Hilfskräfte aus Sozialer Arbeit und Informatik
Kontakt Daten Projektleitung	E-Mail: jens.albrecht@th-nuernberg.de , robert.lehmann@th-nuernberg.de

2. Ausgangslage

Prognosen verdeutlichen, dass etwa 25 Prozent der Bevölkerung im Verlauf ihres Lebens auf soziale Beratung zurückgreifen (ZAUTER & LEHMANN, 2021). Insbesondere im Zuge der fortschreitenden Corona-Pandemie war eine verstärkte Internetrecherche im Bereich gesundheitsbezogener Beratungsthemen zu beobachten (HORCH, 2021), begleitet von einer vermehrten Nutzung von Onlineberatungsangeboten (STIELER u. a., 2022). Das Institut für E-Beratung (IEB) der TH Nürnberg hat es sich zur Aufgabe gemacht, innovative Ansätze in der Online-Sozialberatung zu erforschen. In Kooperation mit der Fakultät Informatik wurden bereits mehrere Projekte erfolgreich umgesetzt, die gezielt das Potential von Künstlicher Intelligenz in diesem Kontext untersuchten (vgl. dazu (ALBRECHT u. a., 2021; LEHMANN u. a., 2021)).

Das Ziel dieser Projekte ist es, mithilfe von technischer Unterstützung die Qualität der Online-Beratung zu verbessern. Dabei ist es für das IEB immer entscheidend, technische Möglichkeiten so nutzbar zu machen, dass sie möglichst direkt einen Nutzen für Ratsuchende bzw. Berater*innen entfalten. Ein Beispiel dafür ist das Projekt "KIA – KI-basierte Assistenz in der Online-Beratung", das durch das Bundesministerium für Familie, Senioren, Frauen und Jugend gefördert wird. In dem Projekt wird ein Assistenzsystem entwickelt, das Onlineberater*innen während einer Beratung interaktives Feedback gibt. Voraussetzung dafür ist eine inhaltliche Analyse der Beratungskonversation. In einem weiteren Forschungsprojekt, das durch die Stiftung Hochschullehre gefördert wird, wird ein Chatbot entwickelt, der eine Klientin oder einen Klienten mit psychosozialen Problemen simuliert. Dieser "virtuelle Klient" soll in der Ausbildung für Online-Berater*innen als Trainingssubjekt eingesetzt werden.

Voraussetzung für die Analyse von Beratungskonversationen und die Entwicklung von intelligenten Assistenten ist immer ein inhaltliches und methodisches Verständnis der Beratungskonversationen. Hierbei müssen die Konversationen mit einem vertieften Verständnis der fachlichen Anforderungen analysiert werden. Um Onlineberatung also belastbar wissenschaftlich zu analysieren, sind qualitative Forschungsmethoden, die mit einem relativ tiefen Textverständnis arbeiten notwendig. Bisher bedeutete das, dass Analysen von entsprechend geschulten menschlichen Forscher*innen erfolgen mussten. Qualitative Analysen zeichnen sich meist durch ein systematisches Vorgehen aus, bei dem Textdaten mithilfe von induktiver oder deduktiver Codierung strukturiert werden. Hierbei werden Kategorien entwickelt, um den Inhalt anschließend unter Verwendung klar definierter Regeln zu interpretieren. Das

limitierte die Menge der analysierten Texte deutlich. Die Entwicklungen im Bereich des maschinellen Lernens, speziell im Natural Language Processing, ermöglichen es heute, komplexe Eigenschaften aus entsprechenden Daten zu erlernen. Die technische Basis für aktuelle Systeme bilden Deep-Learning-Modelle mit Transformer-Architektur (VASWANI u. a., 2017). In der Praxis werden häufig vortrainierte Sprachmodelle wie BERT (DEVLIN u. a., 2019) eingesetzt, die mit relativ wenig Trainingsdaten auf ein konkretes Problem, z.B. das Erkennen von Emotionen oder anderen inhaltlichen Merkmalen, abgestimmt werden können ("Fine Tuning"). Im Vorlauforschungsprojekt CaSoTex konnte beispielsweise gezeigt werden, dass solche Modelle Kategorien wie "Mitgefühl" oder "Handlungsempfehlung" in Äußerungen von Ratsuchenden oder Beratenden mit vergleichbarer Präzision wie Menschen bestimmen konnten (GRANDEIT u. a., 2020). Dadurch wird es möglich, den Einsatz einer automatisierten Inhaltsanalyse ernsthaft zu erwägen. Allerdings ist es notwendig, vortrainierte Sprachmodelle auf passende Datensätze zur Onlineberatung abzustimmen.

3. Ziele des Forschungsprojekts

Die automatisierte Inhaltsanalyse, wie sie auf großen Datenmengen erst durch den Einsatz maschineller Lernverfahren ermöglicht wird, ist eine wesentliche Voraussetzung für die statistische Untersuchung von Wirkzusammenhängen in der psychosozialen Beratung und die Entwicklung beratungsunterstützender, KI-basierter Dialogsysteme. Insgesamt gibt es nur wenige Datensätze, die sich explizit mit Beratung bzw. speziell mit der Online-Beratung beschäftigen, sodass entsprechendes Fine-Tuning bisher noch selten erfolgt ist. Ein Grund dafür ist, dass Daten aus psychosozialen Beratungen in der Regel nicht nur personenbezogene, sondern auch sehr persönliche und vertrauliche Informationen beinhalten. Zudem ist die manuelle Annotation der Daten sehr aufwändig, da sie Fachexpert*innen benötigt. Die wenigen vorhandenen Datensätze sind zudem fast ausschließlich auf Englisch verfügbar, im deutschsprachigen Raum gab es bisher keine frei verfügbaren Trainingsdaten, mit denen entsprechende Modelle erforscht werden könnten.

Ziel dieses Projektes war daher die Erstellung und Veröffentlichung eines ersten deutschsprachigen Datensatzes für die Inhaltsanalyse in der psychosozialen Beratung. Die Vertraulichkeit der Daten wird dadurch gewährleistet, dass entweder reale Daten manuell anonymisiert werden oder künstliche Beratungsdialoge durch Fachexpert*innen in Anlehnung an reale Daten kreiert werden. Ein wesentliches Teilziel des Projekts und Voraussetzung für die Annotation der Daten war die Definition der inhaltlichen Kategorien, die erkannt werden sollen (z. B. "Reflexion", "Anregung zur Ressourcenaktivierung"). Daneben sollte auch ein fertig trainiertes Klassifikationsmodell publiziert werden, das von anderen Forschenden direkt für die Textkategorisierung im Kontext sozialer Beratung eingesetzt und darüber hinaus als Benchmark für andere Modelle genutzt werden kann.

4. Herangehensweise und Forschungsergebnisse

Nach der Analyse des Standes der Forschung und einer eingehenden Sichtung der verfügbaren Datensätze und Kategoriensysteme erfolgte die Erstellung eines umfassenden Kategoriensystems zur Annotation des Datensatzes. Im Anschluss wurde ein Fine-Tuning von mehreren Basismodellen mit diesen Daten vorgenommen, um zu bewerten, ob bzw. wie gut die Kategorien automatisiert erkannt werden können.

4.1. Stand der Forschung

Mit der Inhaltsanalyse im Kontext digitaler Onlineberatung beschäftigen sich einige wenige Arbeiten im englischen Sprachraum (z.B. (ALTHOFF u. a., 2016; DEMASI u. a., 2020; LIU u. a., 2021)). Die vorhandenen Trainingsdatensätze werden jedoch häufig nicht publiziert und sind auch auf Nachfragen nicht verfügbar. So kann bspw. die Nutzung der Trainingsdaten aus (ALTHOFF u. a., 2016) und (DEMASI u. a., 2020) nur durch Forschende hochrangiger Institute aus den USA oder Kanada erfolgen.

Grundlegend für die Inhaltsanalyse ist ein Kategoriensystem (auch Codebuch oder Annotationsschema), welches die inhaltlichen Kategorien präzise definiert (LAMNEK & KRELL, 2016). Anhand dieses Kategoriensystems können dann bestehende Daten annotiert werden. Maschinelle Lernverfahren können

die so erstellten Trainingsdaten nutzen, um Modelle für die automatische Zuordnung von Kategorien in Beratungskonversationen zu berechnen.

Im Folgenden wird ein kurzer Überblick über verfügbare Datensätze und Kategoriensysteme gegeben, die teilweise konzeptuell in den GeCCo-Datensatz eingeflossen sind.

Kategoriensystem "Motivational Interviewing Skill Codes (MISC)" (MILLER u. a., 2008): Die motivierende Gesprächsführung (Motivational Interviewing, MI) ist ein klientenzentrierter Beratungsansatz, der ursprünglich für die Suchtberatung entwickelt wurde und mittlerweile unter anderem in der Psychotherapie Anwendung findet. MI richtet sich vor allem an Personen mit zunächst geringer oder ambivalenter Änderungsbereitschaft. Dabei soll die intrinsische Motivation der Patient*innen bzw. Klient*innen für eine Verhaltensänderung durch entsprechende Fragen aktiviert werden. Eine zentrale Technik, um Verständnis und Empathie zu zeigen sowie um den Dialog zwischen Berater*in und Klient*in zu vertiefen, ist beispielweise die Reflexion. Das Kategoriensystem MISC umfasst insgesamt 28 Kategorien, wobei besonders die Berater*innenkategorien interessant sind. Diese sind *Advise, Affirm, Confront, Direct, Emphasize Control, Facilitate, Filler, Giving Information, Question (Closed or Open), Raise Concern (with or without Permission), Reflection (Simple or Complex), Reframe, Support, Structure* und *Warn*. In (XIAO u. a., 2016) und (CAO u. a., 2019) werden maschinelle Lernverfahren zur Erkennung von MI-Codes vorgestellt. Dabei wurden allerdings die Berater*innen-Kategorien sehr grob in "MI-adherent" und "MI non-adherent" zusammengefasst. Die verwendeten Datensätze sind nicht öffentlich.

Datensatz "ESConv" (LIU u. a., 2021): Emotionale Unterstützung der Klient*innen ist ein wichtiges Element in der psychosozialen Beratung. (LIU u. a., 2021) stellen in ihrer Arbeit den Datensatz ESConv (Emotional Support Conversation) sowie ein entsprechendes Framework für die Modellierung von Dialogen und entsprechende Kategorien vor. Es wird davon ausgegangen, dass Konversationen zu emotionaler Unterstützung nach einem Phasenmodell mit folgenden Stufen (*Stages*) ablaufen: *Comforting, Exploration, Action*. Parallel dazu werden bestimmte *Support-Strategien*, die als inhaltliche Kategorien verstanden werden können, definiert. Unterschieden werden hier *Question, Restatement or Paraphrasing, Reflection of Feelings, Self-disclosure, Affirmation and Reassurance, Providing Suggestions, Information* und *Others*. Darüber hinaus wurden zusätzliche Merkmale wie die Art des Problems oder die Emotionalität erfasst. Der öffentlich verfügbare Datensatz umfasst 1.053 Konversationen bestehend aus über 31.000 Äußerungen.

Datensatz "Anno-MI" (WU u. a., 2022, 2023): Anno-MI ist ein öffentlicher englischsprachiger Datensatz, der transkribierte und von Therapeut*innen kommentierte Therapiedialoge zum Motivational Interviewing (MI) umfasst. Er besteht aus 133 per Video aufgezeichneten Gesprächen, die die wesentlichen Merkmale des MI wiedergeben. MI als psychologische Methode zielt auf Verhaltensänderung im therapeutischen Setting ab. Die Schlüsselemente des MI (OARS nach (MILLER & ROLLNICK, 2015)) sind aufgeteilt nach Klient*in und Berater*in. Die Berater*in-Kategorien (*Therapist Behaviour*) umfassen *Reflection (Simple, Complex), Question (Open, Closed), Input (Information, Advice, Giving Options und Negotiation)* und *Other*. Bei den Klient*innen (*Client Talk Type*) wird unterschieden in *Change, Neutral und Sustain*. Während der sogenannte *Change Talk* eindeutig direktiv ist und somit spezifisch auf das MI zurückgeht, sind die anderen MI-Strategien wie *Reflection* und *Question* in der nicht direktiv klientenzentrierten Gesprächsführung von Rogers verortet (MILLER u. a., 2008).

Datensatz "HOPE" (MALHOTRA u. a., 2022): Der öffentliche HOPE-Datensatz umfasst ca. 12.900 Äußerungen von 212 Therapiesitzungen aus dem Bereich der kognitiven Verhaltenstherapie (Cognitive Behavioral Therapy, CBT). Für die Annotation wurde ein Kategoriensystem bestehend aus 12 Kategorien, organisiert in drei Oberkategorien, entwickelt. Das Kategoriensystem deckt sowohl die Berater*innen- als auch die Klient*innen-Äußerungen ab und soll eine möglichst gute Inhaltsanalyse auf den Daten ermöglichen. Im Einzelnen sind folgende Kategorien definiert: *Speaker Initiative (Clarification Request, Opinion Request, Information Request, Yes/No Question), Speaker Responsive (Information Delivery, Clarification Delivery, Opinion Delivery, Positive Answer, Negative Answer)* und *General (Greeting, Acknowledgement, General Chat)*. Neben dem Datensatz wird in (MALHOTRA u. a., 2022) auch das Verfahren "Time-aware Attention" vorgestellt, welches die zeitliche Abfolge der Äußerungen bei der Klassifikation berücksichtigt.

Kategoriensystem von SafeZone.ch (SCHLAPBACH u. a., 2018) SafeZone.ch ist eine Online-Plattform für Suchtberatung. Die Beratung richtet sich an Betroffene, deren Angehörige und Nahestehende, an

Fachpersonen und Interessierte. Alle Beratungsangebote sind kostenlos und anonym. Die Plattform wurde ab 2014 tiefgehend evaluiert. Der für das Projekt relevante Part bezieht sich auf 100 Emailberatungen zur Anwendung von Wirkfaktoren in der Mailberatung von SafeZone.ch (Zeitraum 2014-2017). Die betrachteten Kategorien beziehen sich auf die sogenannten "Wirkfaktoren" (Problemaktivierung, motivationale Klärung, therapeutische Beziehung, Ressourcenaktivierung etc.) nach (GRAWE, 1998) und sechs weitere Wirkfaktoren entsprechend der spezifischen Rahmenbedingungen der Emailberatung (psychologische Anonymität, Asynchronität, Freiwilligkeit, Beratung als Ko-Produktion etc.). Ein Datensatz ist nicht verfügbar, sondern detaillierte Berichte (Evaluationsberichte, Abschlussberichte).

Kategoriensystem und Datensatz "CASoTex" (GRANDEIT u. a., 2020; LEHMANN u. a., 2021): In einem eigenen Vorprojekt wurde eine erste Version eines Kategoriensystems zur Erfassung relevanter Dimensionen der Onlineberatung definiert und evaluiert. Die Datenbasis bilden Nutzer*innenbeiträge aus acht großen öffentlichen psychosozialen Beratungsforen (Sucht- und Selbsthilfe, BKE Jugend- und Elternberatung, Beratungs4Kids, Dr. Azubi, DGB Jugend, Hilfe für Mädchen und Young Wings). Das Kategoriensystem ist mit 50 Einzelkategorien recht umfangreich. Es ist eingeteilt u.a. in die Gruppen "Formalitäten" (z.B. Anrede, Verabschiedung), die sich an den Spezifika der Onlineberatungsmethoden orientieren (ENGELHARDT, 2021; REINDL, 2018), "Grundhaltungen" wie Empathie und Sorge um andere (nach Rogers in: (PAULUS, 2016)), "Wirkfaktoren" wie Problemdarstellung, Handlungsempfehlung (vgl. (GRAWE, 1998)) und unterschieden nach Berater*in und Klient*in. Der Datensatz konnte aufgrund der enthaltenen personenbezogenen Daten nicht veröffentlicht werden, steht den Autor*innen jedoch zur Verfügung.

Datensatz "Counsel Chat" (BERTAGNOLLI, 2020): Der Datensatz enthält einen Abzug öffentlich verfügbarer Beratungskonversationen von der Internet-Plattform counselchat.com. Auf dieser Plattform bieten professionelle Berater- und Therapeut*innen Ratsuchenden Hilfe an. Die Daten sind nur auf Englisch verfügbar und reine Dialogdaten, d.h. es ist keinerlei Kategorisierung vorgenommen wurden.

Es wird deutlich, dass es eine sehr geringe Anzahl an Datensätzen gibt, die zur Analyse von Texten aus der Domäne der Onlineberatung nutzbar sind. Diese sind überwiegend nicht öffentlich verfügbar. Die oben erwähnten Projekte stammen nicht alle originär aus der Onlineberatung, sondern sind auf die Analyse von Konversationen aus der Präsenzberatung ausgerichtet. Daraus ergeben sich methodische Unterschiede. Zudem sind Beratung und Therapie jeweils auch nochmal voneinander zu differenzieren in ihrer unterschiedlichen methodischen Herangehensweise. Dennoch liegen viele Gemeinsamkeiten vor, die für die spätere heterogene Auswahl der Kategorien eine Rolle spielten. Während MI beispielsweise nur ein Beratungsansatz ist, zeichnet sich Onlineberatung in der Regel durch ein eklektizistisches Vorgehen aus, also die Nutzung mehrerer Beratungsansätze (vgl. (MODE-SCHEIBEL, 2014)).

Eine weitere Erschwernis besteht folglich in den unterschiedlichen Kategorisierungen der Inhalte. Diese sind zwar jeweils in anerkannten und weitverbreiteten Theorien aus dem Feld der psychosozialen Beratung fundiert, allerdings sind sie in der vorliegenden Form nicht kompatibel. Teilweise überlappen sich die Kategorien, teilweise weisen sie sehr unterschiedliche Erklärungsniveaus auf. Daher ist es erforderlich, einen großen öffentlichen Datensatz für diese Zwecke zu generieren. Weiterhin ist es notwendig, die Daten in dem Datensatz mit einem Kategoriensystem zu bearbeiten, das zu möglichst vielen Theorien und Konzepten der psychosozialen Beratung kompatibel ist und diese Bezüge auch explizit darlegt.

4.2. Erstellung des Kategoriensystems

Anforderungen und Rahmenbedingungen

Ethik und Datenschutz: Onlineberatungsverläufe und deren KI-gestützte Auswertung bringen ethische Implikationen mit sich hinsichtlich der Frage, wie persönliche Daten und Informationen erhoben, verarbeitet und gespeichert werden. Selbst bei vermeintlich öffentlich zugänglichem Material, wie Beratungen in öffentlichen Foren, ist unklar, inwieweit diese für die Analyse mit maschinellen Lernverfahren genutzt werden dürfen (GHANEM u. a., 2022). Dabei ist zu unterscheiden, ob es sich um *found*, also nicht primär für wissenschaftliche Studien erhobene, oder *designed data*, die für wissenschaftliche Zwecke erhoben wurden, handelt ((RAT FÜR SOZIAL- UND WIRTSCHAFTSDATEN (RATSWD), 2023)). Werden Trainingsdaten generiert, die für weitere Modelle genutzt werden können, schließen sich umfangreiche Fragen aus dem Bereich des Datenschutzes, Urheberrechts und der Forschungsethik an. Solche Fragen gehen dann mit

Beschränkungen einher und können eine Ursache dafür sein, dass lediglich Teile von Datensätzen veröffentlicht werden. Diese beschränkte Zugänglichkeit verhindert oftmals "[...] die Weitergabe der von Forschenden für eine Publikation verwendeten Daten und damit die Reproduzierbarkeit" (ebd., S. 12). Weiterhin muss der Prozess der Erfassung und Auswertung durch Transparenz und Nachvollziehbarkeit gekennzeichnet sein. Insbesondere im Feld der Sozialen Arbeit müssen Diskriminierungsrisiken vermieden werden, beispielsweise durch die Auswahl der Datensätze und Kennzahlen zur Berechnung und zum Training maschineller Verfahren. Bestimmte Personengruppen dürfen nicht unter- oder überrepräsentiert sein und keine Stereotypen abbilden. Zudem ist eine vollständige Anonymisierung Grundvoraussetzung für die Erstellung eines öffentlichen Datensatzes. Diese umfangreichen Anforderungen machen es nötig, dass entsprechende Datensätze fachlich kuratiert erstellt werden. Angesichts der Vielfalt der Anforderungen überrascht es nicht, dass bisher noch kein solcher Datensatz erstellt wurde.

Kategorien: An den Datensatz werden einerseits aus technischer Sicht Anforderungen für das Training von Modellen herangetragen. So sollen komplexe Sinneinheiten in den Beratungstexten automatisch erkannt und kategorisierbar werden. Die Kategorien können methodischer Natur sein ("Problemdarstellung" auf Klient*innenseite, "fachliche, förmliche, institutionelle Handlungsempfehlung" auf Berater*innenseite) oder andere Eigenschaften wie Emotionalität ("eigene Gefühlsdarstellung" auf Klient*innenseite, "Nachfrage nach emotionaler Lage" oder "Dank und Wertschätzung an Klient*in" auf Berater*innenseite) umfassen. Das verwendete System muss einerseits hinreichend viele Kategorien umfassen, um vielfältige Analysen zu ermöglichen. Andererseits müssen die Kategorien inhaltlich scharf genug voneinander abgrenzbar sein, damit ein Modell überhaupt in der Lage ist, Unterschiede zu erlernen. Diese Trennschärfe wird durch die jeweilige Definition bzw. Kategorienbeschreibung erreicht.

Angestrebt wird darüber hinaus eine weitgehende Vergleichbarkeit und theoretische Kompatibilität mit den Systemen aus den oben dargestellten Datensätzen.

Zielsetzung

Aus der Perspektive der Sozialarbeitswissenschaft sollte der Datensatz mit einem Kategoriensystem bearbeitet sein, das relevante Kriterien aus der Beratungsforschung abbildet. So sind insbesondere im Bereich der Onlineberatung sowohl zur methodischen Gestaltung als auch zum Ausdruck der Emotionalität umfangreiche Wissensbestände vorhanden, zu denen das verwendete Kategoriensystem kompatibel sein sollte (ENGELHARDT, 2021; HILDEBRAND u. a., 2020; STIELER, 2022).

Um anschlussfähige Studienergebnisse generieren zu können, müssen die Kategorien im Datensatz mit diesem Fachdiskurs vergleichbar sein. Dabei ist weiter zu beachten, dass die Sozialarbeitswissenschaft allgemein und die Forschung zur Onlineberatung im Speziellen im deutschsprachigen Raum relativ wenig mit dem internationalen Diskurs verschränkt ist. Gleichzeitig liegen jedoch vereinzelt Trainingsdatensätze im internationalen Feld vor. Daher ist es erstrebenswert im Projekt GeCCo ein Kategoriensystem zu erstellen, das auch mit den internationalen Datensätzen kompatibel ist und so eine Brücke zwischen nationalem und internationalem Diskurs schlägt.

Vorgehensweise

Um diese Zielsetzung zu erreichen, musste ein mehrstufiges Verfahren angewandt werden. Als Grundlage wurde das Kategoriensystem aus dem Vorlaufforschungsprojekt "Computergestützte Analyse Sozialwissenschaftlicher Texte mit Hilfe maschineller Lernverfahren - CASoTex" (GRANDEIT u. a., 2020; LEHMANN u. a., 2021) verwendet. Bereits damals wurden Kategorien aus dem Fachdiskurs eingearbeitet. Allerdings war das Kategoriensystem auf den Anwendungsfall aus CASoTex zugeschnitten und wies einige Limitationen auf, die bereits im Abschlussbericht zu CASoTex thematisiert wurden (LEHMANN u. a., 2020). Daher wurde zunächst ein Expert*innenteam aus assoziierten Projektmitgliedern einberufen. Bei der Teamkonstellation wurde darauf geachtet, dass erfahrene Fachkräfte aus der Onlineberatung und Wissenschaftler*innen aus der qualitativen Onlineberatungsforschung kooperativ theoretische und forschungspraktische Anforderungen einbringen konnten.

Zu Beginn erfolgte eine umfangreiche Literaturrecherche zu bestehenden Kategoriensystemen und Datensätzen mit (Onlineberatungs-)Bezug. Das Projektteam betrachtete verschiedene Kategoriensysteme, die Aufbereitung der enthaltenen Datensätze und ihre vorangegangenen Forschungszwecke. Ziel dieses Vorgehens war es zunächst eine geeignete Schnittmenge an Kategorien für das Projekt zu identifizieren.

Neben den relevanten theoretischen Konstrukten ((GRAWÉ, 1998), (PAULUS, 2016) etc.), die für die Beurteilung von Onlineberatungsprozessen relevant sind, wurden also auch die bestehenden Klassifikationen dieser Modelle näher betrachtet, um Unterschiede, Ähnlichkeiten und gegebenenfalls auch Dopplungen in den Kategoriensystemen zu identifizieren.

Der Kern der Erstellung war ein interaktiver Workshop, in dem die verschiedenen bestehenden Ansätze buchstäblich nebeneinandergelegt wurden und so eine Synthese der verschiedenen Ansätze geschaffen wurde, die den Ansprüchen an ein wissenschaftlich fundiertes Kategoriensystem (Abbildung 1) genügt. Berücksichtigt wurden dabei die Unterscheidungen zwischen Klient*in und Berater*in und die verschiedenen Kategorie-Ebenen (von der Oberkategorie bis hin zur kleinsten Stufe der Unterkategorien).



Abbildung 1: Workshop zur Entwicklung des Kategoriensystems

Ergebnis

Das so erstellte Kategoriensystem weist folgende Eigenschaften auf:

- Als Basis wurden die eigens erstellten Kategoriensysteme des IEB zur Beschreibung fachlicher Elemente der Onlineberatung genutzt.
- Es bezieht sich auf die Beratenden- wie auch auf die Klient*innenseite.
- Es enthält Kategorien, diese sind maximal auf fünf Ebenen untergliedert, die Definition und Ankerbeispiele. Um die Kategorie genau zu bestimmen, wurden Kategoriekürzel angelegt.
Hier ein Beispiel:
Beratendenseite > Wirkfaktor > Analyse und Klärung der Probleme > Reflexion (Sachverhalt) > Nachfrage zur Lebenslage > soziale Beziehungen
Das entsprechende Kürzel lautet: B-WF-AK-RS-FLL-SB
- Beibehalten wurden Oberkategorien wie beispielsweise "Formalitäten am Anfang" und "Wirkfaktoren", die sich bewährt haben. Auf der vierten Kategorieebene wurden MI-spezifische Kategorien ergänzt.

Wie oben bereits erwähnt erfolgt Onlineberatung nach unterschiedlichen Beratungsansätzen, Stichwort Eklektizismus. Ein besonders häufig genutzter ist der des Psychologen Carl. R. Rogers. Er formuliert drei zentrale Linien in der Beratung. Diese sind das empathische Verstehen, die unbedingte Wertschätzung und Kongruenz (vgl. (WATSON, 2007)). Diese Grundhaltungen spielen neben der Präsenzberatung auch für Onlineberatung eine bedeutende Rolle. Allerdings sind aufgrund der medial vermittelten Kommunikationsform netzspezifische Charakteristika miteinzubeziehen, wie beispielsweise Aspekte der nonverbalen Kommunikation, die im Onlinesetting wegfallen. In der Onlineberatung können sich die obengenannten Haltungslinien eben nicht auf Gestik oder Mimik beziehen, sondern kommen in der Textanalyse von z. B. Chat- oder Mailberatung zum Tragen.

Eine besondere Lösung wurde für die Kategorie der Empathie gewählt. Angelehnt an Paulus (2016) wird Empathie als Grundhaltung des Mitgefühls und der Sorge verstanden. Es handelt sich hier um ein sehr relevantes Konstrukt, das sich in der schriftbasierte Onlineberatung allerdings selten in allein stehenden expliziten Aussagen widerspiegelt. Vielmehr tritt es als emotionale Färbung einer anderen, eher inhaltlichen Aussage auf. Daher wurde die Empathie als zusätzliche Dimension für die Beratenden-Seite in das Kategoriensystem integriert. Dabei wird zu jeder Aussage zusätzlich zu den inhaltlichen Kategorien ein

Intensitätsgrad der Empathie (geringe – moderate - ausgeprägte Empathie) kodiert. In der ersten, zum Zeitpunkt dieser Publikation veröffentlichten Version des Datensatzes wurde diese Dimension aus Zeitgründen noch nicht verwendet. Die Grundlage für eine zukünftige Nutzung ist allerdings angelegt. Der Dreiklang der Grundhaltungen in Anlehnung an Paulus (2016) besteht allerdings noch auf Klient*innenseite, da sich hier inhaltliche Bezüge zu anderen Beteiligten ergeben (Empathie für Dritte (PT-perspective taking, Mitgefühl mit anderen und Sorge um andere Person (EC-emotional concern) (vgl.(YU u. a., 2020) und (PAULUS, 2016)).

Eine weitere Erweiterung des Kategoriensystems ist die inhaltliche Kompatibilität zum ESCon-Datensatz (LIU u. a., 2021). Dabei soll ein möglichst analoges Phasenmodell in die Logik des GeCCo-Datensatzes integriert werden.

4.3. Erstellung des Datensatzes

Da für solch einen Datensatz ein hoher Qualitätsstandard angesetzt werden muss, wurde großer Aufwand auf verschiedenen Ebenen betrieben. Zum einen wurden die Annotator*innen ausführlich in der zugrundeliegenden Methodik unterwiesen. Das Team bestand aus fünf ausgebildeten Onlineberater*innen sowie fünf studentischen Hilfskräften aus dem Studiengang Soziale Arbeit und einer weiteren aus dem Studiengang Informatik. Fachlich angeleitet wurde diese Gruppe von der Projektkoordination und der Projektleitung.

Da die Projektleitung zusammen mit dem Institut für E-Beratung bereits seit etlichen Jahren sich eine große Expertise in diesem spezifischen Feld ausgebaut hat, konnte die Auswahl der studentischen Hilfskräfte auf Personen fallen, die explizit Erfahrungen im Umgang mit der Annotation von E-Beratungsdaten im psychosozialen Kontext aufweisen.

Die von den studentischen Hilfskräften zusammengetragenen Daten wurden von erfahrenen Onlineberater*innen hinsichtlich der Codier-Vorgaben überprüft. Grundlage für den Aufbau des Kategoriensystems bildet die Entwicklung und Anwendung eines Kategoriensystems in der qualitativen Inhaltsanalyse nach Mayring (MAYRING & FENZL, 2019). Mayring formuliert hier Anforderungen an Kategoriensysteme in der qualitativen Sozialforschung. Für den Zweck der Codierung eines Datensatzes im Machine Learning Umfeld wurde die Systematik daher vom Forschungsteam explorativ weiterentwickelt.

Alle generierten Textbeispiele wurden manuell gesichtet, gegebenenfalls modifiziert oder verworfen, wenn die Vorgaben nicht erfüllt waren. Sämtliche Daten sind synthetisch erstellt, sodass keine personenbezogenen Daten enthalten und dadurch die Regeln zum Datenschutz nicht relevant sind.

4.4. Beschreibung des Datensatzes

Das Kategoriensystem und der Datensatz wurden im Github-Repository der Technischen Hochschule Nürnberg¹ veröffentlicht. Die Details können dort eingesehen werden. Tabelle 1 gibt einen Überblick über das Kategoriensystem und die Anzahl der jeweils erstellten Beispieltexte.

Das finale Kategoriensystem umfasst 40 Berater*innen- und 28 Klient*innen-Kategorien, wobei der größte Teil jeweils auf die Wirkfaktoren entfällt. Das Kategoriensystem ist hierarchisch aufgebaut. Ein Beispiel für eine Kategorie auf Ebene vier ist die "Nachfrage zur Lebenslage", die im Ast "Wirkfaktoren → Analyse und Klärung der Probleme → Reflexion (Sachverhalt)" zu finden ist. Diese spezielle Kategorie ist darüber hinaus weiter nach der Art der Nachfrage (soziale Beziehungen, berufliche Situation, Gesundheit etc.) unterteilt. Die meisten Kategorien wurden aber auf Ebene vier ohne weitere Unterteilung definiert.

Die erste Version des GeCCo-Datensatzes umfasst insgesamt über 2800 Text-Beispiele für diese 68 Kategorien erzeugt. Die meisten Daten entfallen dabei auf die jeweilige Oberkategorie "Wirkfaktoren", da diese im Focus der Analyse steht und sehr fein unterteilt ist. Abbildung 2 zeigt die Datenverteilung der

¹ <https://github.com/th-nuernberg/gecco-dataset>

Kategorien im Bereich "Wirkfaktoren". Wie deutlich zu erkennen ist, sind die Daten nicht gleichmäßig auf die Kategorien verteilt. Das liegt hauptsächlich daran, dass die Äußerungen in den Beratungsdialogen in der Realität auch nicht gleichverteilt sind. Um durch ein Modell auch seltene Kategorien gut erkennen zu können, wären noch mehr Trainingsdaten wünschenswert. Nach entsprechenden Tests ist daher geplant, die Daten punktuell zu erweitern.

Typ	Kategorie Ebene 1	Kategorie Ebene 2	Anzahl Kategorien	Anzahl Texte
Berater	Formalitäten am Anfang	Formalitäten am Anfang	2	213
	Wirkfaktoren	Analyse und Klärung der Probleme	13	879
		Analyse und Vereinbarung von Beratungszielen	2	56
		Erzeugung von Motivation	4	131
		Ressourcenaktivierung	5	117
		Hilfe, Problembewältigung	8	71
	Formales zum Abschluss		2	62
Andere Aussagen		4	19	
	Gesamt		40	1548
Klient	Formalitäten am Anfang	Formalitäten am Anfang	4	171
	Wirkfaktoren	Analyse und Klärung der Probleme	8	439
		Analyse und Vereinbarung von Beratungszielen	2	94
		Erzeugung von Motivation	2	76
		Ressourcenaktivierung	2	64
		Hilfe Problembewältigung	6	282
	Formales zum Abschluss		2	70
Andere Aussagen		2	59	
	Gesamt		28	1255

Tabelle 1: Oberste Ebenen des Kategoriensystems

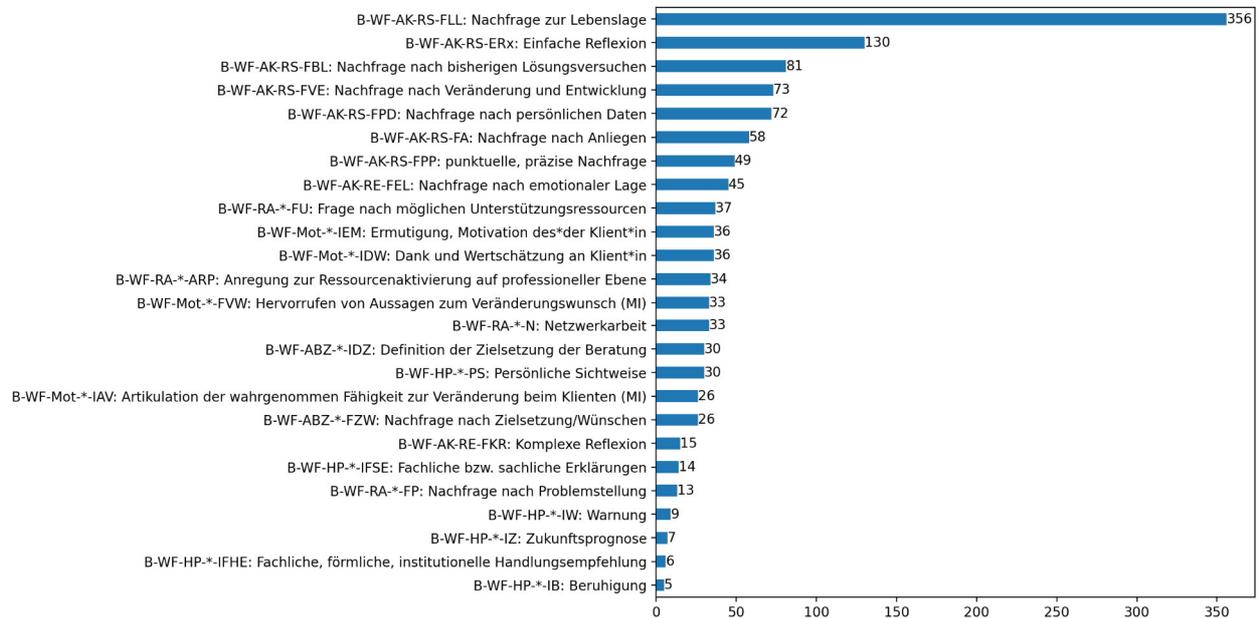


Abbildung 2: Datenverteilung in der Oberkategorie "Wirkfaktoren"

Um einen Eindruck davon zu bekommen, wie die Textbeispiele aussehen, seien hier zwei Beispiele aufgeführt:

- *"Es ist ganz normal, dass ein junger Mensch während der Pubertät beginnt, sexuelle Wünsche und Vorstellungen zu entwickeln."*

Die Berater*innen-Äußerung ist der Kategorie "Wirkfaktoren (WF); Hilfe, Problembewältigung (HP); Fachliche bzw. sachliche Erklärungen" zugeordnet, zu der bisher nur 14 Trainingsbeispiele bestehen.

- "Es geht mehr um meinen Bruder als um mich."

Diese Klient*innen-Äußerung ist ein Beispiel aus der Kategorie "Formalitäten am Anfang; Sorge um andere Person".

Die Texte sind meistens um die 100 Zeichen lang. Abbildung 3 zeigt Box-Plots der Werteverteilung der Textlängen pro Oberkategorie für Berater*innen und Klient*innen. Dabei wird deutlich, dass die Schwankungsbreite bei den Äußerungen aus dem Bereich "Wirkfaktoren" erwartungsgemäß am größten ist.

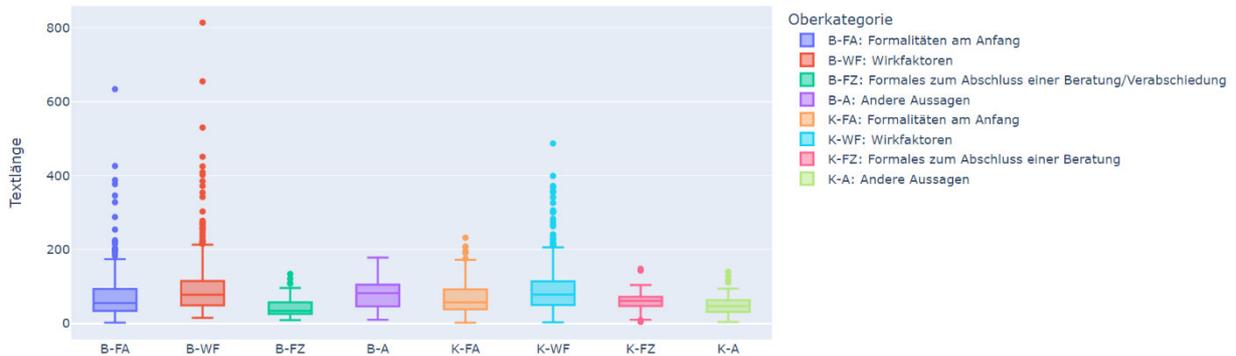


Abbildung 3: Längenverteilung der Beispieltex-te

4.5. Training mittels Transfer Learning

Ein wesentliches Projektziel war die Erstellung und Veröffentlichung eines mit diesen Daten trainierten Modells. In der Vergangenheit konnte bereits gezeigt werden, dass Transformer-basierte Modelle gut dafür geeignet sind (GRANDEIT u. a., 2020). Die Grundidee beim Training ist es, ein mit sehr vielen Textdaten vortrainiertes Modell als Basis zu nehmen und mit den eigenen Daten weiter zu trainieren. Dieser Prozess wird auch als Transfer Learning bezeichnet.

Als Basis-Modelle wurden verschiedene Modelle professioneller Institutionen ausgewählt, die mit deutschsprachigen Texten trainiert wurden. Konkret wurden ein Basismodell des Münchner Digitalisierungszentrums (MDZ) der Bayerischen Staatsbibliothek sowie drei Modelle der Firma deepset.ai für das Fine-Tuning mit den GeCCo-Daten verwendet.

Basismodell	GeCCo-Modell	Parameter	Accuracy	F1
dbmdz/bert-base-german-uncased	bert-base-german-counseling-gecco	110Mio.	0,70	0,45
deepset/gelectra-large	gelectra-large-german-counseling-gecco	336 Mio.	0,73	0,46
deepset/gbert-large	gbert-large-german-counseling-gecco	336 Mio.	0,78	0,66
deepset/gbert-base	gbert-base-german-counseling-gecco	110 Mio.	0,72	0,51

Tabelle 2: Ergebnisse des Fine-Tunings mit den GeCCo-Daten

Tabelle 2 zeigt eine Übersicht über die verglichenen Modelle und die damit erreichten Ergebnisse. Das beste Ergebnis wurde mit gbert-large, einem mit deutschen Texten trainierten Modell basierend auf dem BERT-Modell von Google (DEVLIN u. a., 2019), erreicht. Die Accuracy, also der Anteil der korrekt erkannten Kategorien auf den 20% Testdaten, liegt hier bei sehr guten 75%. Bei schiefen Verteilungen ist die Accuracy aber nur bedingt aussagekräftig, da die Klassen gewichtet mit der Anzahl der Beispiele in die Kennzahl eingehen und somit die schlechte Erkennung von Klassen mit wenig Beispieldaten kaum ins Gewicht fällt. Eine bessere Aussage liefert in diesem Fall der F1-Score, der auch den Recall (Trefferquote) mitberücksichtigt. Mit 0,66 ist dieser Wert bei gleicher Gewichtung aller Klassen (F1 Macro) trotz der schiefen Datenverteilung überraschend gut. Tatsächlich werden auch Kategorien mit sehr wenigen Trainingsdaten relativ gut erkannt, wie Tabelle 3 zu entnehmen ist.

Kategorie	Precision	Recall	F1	Support
B-A-*. *-UB-*)Unpassende Bemerkung	0.50	0.33	0.40	3
B-FA-*. *-G-*)Gesprächseröffnung	0.92	0.92	0.92	25
B-FA-*. *-Mod-*)Moderation	0.89	0.89	0.89	18
B-FZ-*. *-APR-*)Angebot zur Nutzung der prof. Ressourcen des*der Berater*in	0.67	1.00	0.80	2
B-WF-ABZ-*. -FZW-*)Nachfrage nach Zielsetzung/Wünschen	0.25	0.40	0.31	5
B-WF-ABZ-*. -IDZ-*)Definition der Zielsetzung der Beratung	1.00	1.00	1.00	6
B-WF-AK-RE-FEL-*)Nachfrage nach emotionaler Lage	0.73	0.89	0.80	9
B-WF-AK-RE-FKR-*)Komplexe Reflexion	0.50	0.33	0.40	3
B-WF-AK-RS-ERx-*)Einfache Reflexion	0.91	0.81	0.86	26
B-WF-HP-*. -IFSE-*)Fachliche bzw. sachliche Erklärungen	1.00	1.00	1.00	3
B-WF-RA-*. -FU-*)Frage nach möglichen Unterstützungsressourcen	0.71	0.71	0.71	7
B-WF-RA-*. -N-ARFa)Netzwerkarbeit Ressourcenaktivierung auf Familienebene	1.00	1.00	1.00	5
B-WF-RA-*. -N-ARFr)Netzwerkarbeit Ressourcenaktivierung auf Freundesebene	1.00	1.00	1.00	1

Tabelle 3: Auszug aus den Detail-Ergebnissen (Classification Report) der einzelnen Kategorien auf den Testdaten.

Alle genannten Modelle wurden auf den Huggingface Model-Hub hochgeladen. Huggingface ist ein amerikanisches Unternehmen, das zum einen Software für maschinelles Lernen entwickelt und zum anderen einen stark frequentierten kostenlosen Marktplatz für Modelle und Daten bereitstellt. Die GeCCo-Modelle stehen unter <https://huggingface.co/th-nuernberg> zum Abruf und zum Test zur Verfügung.

⚡ Inference API ⓘ

🔗 Text Classification Examples ▾

Es ist ganz normal, dass ein junger Mensch während der Pubertät beginnt, sexuelle Wünsche und Vorstellungen zu entwickeln.

Compute

Computation time on Intel Xeon 3rd Gen Scalable cpu: cached

B-WF-HP-*. -IFSE-*)Fachliche bzw. sachliche Erklärungen	0.789
B-WF-AK-RS-FPP-*)punktuelle, präzise Nachfrage	0.016
B-WF-AK-RS-FPD-*)Nachfrage nach persönlichen Daten	0.010

Abbildung 4: Inference API zum Testen von Modellen auf huggingface.co

Das beste Modell kann auf <https://huggingface.co/th-nuernberg/gbert-large-german-counseling-gecco> direkt ausprobiert werden. Abbildung 4 zeigt die Vorhersage für Berater*innen-Äußerung aus dem vorherigen Abschnitt. Das Modell klassifiziert die Äußerung mit hoher Wahrscheinlichkeit korrekt in die Kategorie "Fachliche bzw. sachliche Erklärungen".

4.6. Zusammenfassung der Ergebnisse und weitere Schritte

Das Kategoriensystem, das im Rahmen dieses Projektes verschiedene bestehende Systeme und Beratungsansätze integriert, ist für die sozialwissenschaftliche Onlineberatungsforschung eine große Bereicherung. Neben der Codierung des GeCCo-Datensatzes ermöglicht dieses Instrument auch in zukünftigen Forschungsprojekten eine umfangreiche Codierung von Onlineberatungstexten. So liegt nun ein Standard vor, der vergleichbare Ergebnisse bei der Erforschung verschiedener Quellenbestände zur psychosozialen Onlineberatung ermöglicht. Durch den damit erstellten Datensatz und das trainierte Modell stellt das GeCCo-Projekt der Fachwelt ein Toolset zur Verfügung, mit dem auch andere Forschungsgruppen Onlineberatungstexte mit Hilfe von Machine Learning analysieren können. So ist die Grundlage für umfangreiche Forschungsprojekte in diesem Bereich gelegt.

Auch wenn durch die erstellten Modelle aktuell bereits mehr als zufriedenstellende Ergebnisse erreicht werden, ist hier noch Verbesserungspotenzial vorhanden. So kann durch die Erstellung oder Annotation weiterer Beispieldaten in eher unterrepräsentierten Kategorien die Trefferquote wahrscheinlich deutlich gesteigert werden. Geplant ist darüber hinaus die Untersuchung generativer Transformer-Modelle (wie z.B. die GPT-Modelle) für die Generierung weiterer Trainingsdaten bzw. direkt für die Klassifikation ohne eigenes Training (Few-Shot-Learning, vgl. (BROWN u. a., 2020)).

5. Nachhaltigkeit / Verwertung / wissenschaftliche Arbeiten

Im Bereich der deutschsprachigen Sozialen Arbeit und Onlineberatung kommen KI-unterstützte Ansätze nach wie vor kaum zum Einsatz. Sie werden in großen Teilen der Community als neue Herausforderung wahrgenommen (SCHRÖDTER & BASTIAN, 2019) und eher selten als Chance für die fachliche Weiterentwicklung begriffen. Auch wenn aktuelle Publikationen hier allmählich eine Änderung der Bewertung erkennen lassen (LEHMANN & STIELER, 2022), sind Umsetzungen von KI-basierten Verfahren selten. Ein großes Problem besteht in der Verfügbarkeit von Trainingsdaten. Mit dem nun vorliegenden Datensatz wird die Untersuchung innovativer Anwendungsfelder durch einen breiten und vor allem interdisziplinären Forscher*innenkreis ermöglicht. Die Etablierung dieses neuen Datensatzes stellt demnach eine bedeutsame Innovation für Forscher*innen im deutschsprachigen Raum und ein Novum dar.

Die Publizierung ist als erste deutschsprachige Referenz im Themenbereich als Grundlageninvestition zu werten, die einen wissenschaftlichen Ruf begründet. Der entstehende Datenkorpus kann vielschichtig beforscht und Grundlage für neue Forschungsfragen werden. Beispielhafte Fragestellungen sind die Untersuchung von Wirkfaktoren in der Beratung, sprachlichen Mustern, Persönlichkeitsmerkmalen, emotionalen Ausnahmezuständen oder gefährdender Faktoren (Suizid, Kindeswohl etc.).

Jede Publikation, die diesen Datensatz verwendet, ergibt eine weitere Referenz für die Forschungsaktivitäten an unserer Hochschule. Im Fortgang kann der Datensatz vernetzt werden, indem weitere Forschende die Daten ergänzen und adaptieren. Datensatz und Modell sind dabei nicht nur für die sozialwissenschaftliche Forschung interessant, sondern erweitern auch das Spektrum an Szenarien für die Untersuchung maschineller Lernverfahren auf natürlichsprachlichen Texten allgemein.

Das Projektergebnis kann als Referenz für dieses Forschungsgebiet im deutschsprachigen Raum dienen. Es schafft ebenfalls die Voraussetzungen für die Beantragung von Anschlussforschung im Rahmen der neuen DFG-Förderlinie für HAW sowie für weitere Förderprogramme zur Veränderung der (sozialen) Arbeitswelt durch Digitalisierung (BMBF, BMFSFJ, Stiftungen) oder zur Digitalisierung und Inklusion vulnerabler Zielgruppen (Bildungsferne, Menschen mit Einschränkungen) (BMAS, BMG). Praxispartner*innen aus unterschiedlichen Beratungsfeldern können über das Institut für E-Beratung sowohl zur Datenbereitstellung als auch als praxisnahe Anwender*innen für wesentliche Fragestellungen eingebunden werden.

Die finale Datenbank wird Studierenden beider Fakultäten, Sozialwissenschaften und Informatik, vorgelegt. Sie werden aktiv ermutigt, daran zu forschen und die Entwicklung weiterer Datensätze und Modelle voranzutreiben. Konkret ist bereits die Einbindung in Lehrveranstaltungen zur Wissenschaft der Sozialen Arbeit geplant, in der Studierende der Sozialen Arbeit zusammen mit Mitarbeiter*innen des Instituts für E-Beratung Fragestellungen der Sozialen Arbeit anhand dieses Datensatzes und weiterer Modelle aus dem Feld der Künstlichen Intelligenz erproben. So werden niedrigschwellig auch außerhalb der Informatik Kompetenzen und Erfahrungen im Umgang mit Künstlicher Intelligenz bei zukünftigen Fachkräften der Sozialen Arbeit vermittelt. Weiterhin können durch studentische Arbeiten neue Datensätze entstehen oder der bestehende Datensatz könnte angereichert werden. Denkbar wären hier z. B. ein E-Beratungsdatsatz zu den Themen Bildung oder Gesundheit. Es ist geplant, die Ergebnisse jeweils in Arbeiten mit technischem und sozialwissenschaftlichem Fokus zu publizieren. Es wird auf kostenfreien elektronischen Zugriff (Open Access) geachtet.

Literatur

- ALBRECHT, JENS; LEHMANN, ROBERT; ZAUTER, SIGRID; INSTITUT FÜR E-BERATUNG: *Digitale Assistenz in der psychosozialen Beratung: Projekt-Abschlussbericht ; 9.2019-2.2021* : Technische Hochschule Nürnberg Georg Simon Ohm, 2021
- ALTHOFF, TIM; CLARK, KEVIN; LESKOVEC, JURE: Large-scale Analysis of Counseling Conversations: An Application of Natural Language Processing to Mental Health. In: *Transactions of the Association for Computational Linguistics* Bd. 4 (2016), S. 463–476
- BERTAGNOLLI, NICOLAS: *Counsel chat: Bootstrapping high-quality therapy data*. URL <https://towardsdatascience.com/counsel-chat-bootstrapping-high-quality-therapy-data-971b419f33da>. — Towards Data Science
- BROWN, TOM B.; MANN, BENJAMIN; RYDER, NICK; SUBBIAH, MELANIE; KAPLAN, JARED; DHARIWAL, PRAFULLA; NEELAKANTAN, ARVIND; SHYAM, PRANAV; U. A.: Language Models are Few-Shot Learners. In: *Proceedings of 34th Conference on Neural Information Processing Systems (NeurIPS)*. Vancouver, Canada, 2020
- CAO, JIE; TANANA, MICHAEL; IMEL, ZAC E.; POITRAS, ERIC; ATKINS, DAVID C.; SRIKUMAR, VIVEK: Observing Dialogue in Therapy: Categorizing and Forecasting Behavioral Codes, arXiv (2019). — arXiv:1907.00326 [cs]
- DEMASI, ORIANNA; LI, YU; YU, ZHOU: A Multi-Persona Chatbot for Hotline Counselor Training. In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online : Association for Computational Linguistics, 2020, S. 3623–3636
- DEVLIN, JACOB; CHANG, MING-WEI; LEE, KENTON; TOUTANOVA, KRISTINA: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the ACL*. Minneapolis, Minnesota : Association for Computational Linguistics, 2019, S. 4171–4186
- ENGELHARDT, EMILY M.: *Lehrbuch Onlineberatung. 2., erweiterte Auflage* 2021. Göttingen : Vandenhoeck & Ruprecht, 2021 — ISBN 978-3-525-71155-2
- GHANEM, CHRISTIAN; ECKL, MARKUS; LEHMANN, ROBERT: Big Data und Forschungsethik in der Sozialen Arbeit. In: *EthikJournal* Bd. 8 (2022), Nr. 1
- GRANDEIT, PHILIPP; HABERKERN, CAROLYN; LANG, MAXIMILIANE; ALBRECHT, JENS; LEHMANN, ROBERT: Using BERT for Qualitative Content Analysis in Psychosocial Online Counseling. In: *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*. Online : Association for Computational Linguistics, 2020, S. 11–23
- GRAWE, KLAUS 1943-2005: *Psychologische Therapie* : Hogrefe, Verlag für Psychologie, 1998 — ISBN 978-3-8017-0978-5
- HILDEBRAND, ANJA; WEISS, MAREN; BRAUN-SCHARM, HELLMUTH; STEMMLER, MARK: *Evaluierung [U25] - Evaluierung der Online-Suizidpräventions-Beratung* : FAU Erlangen-Nürnberg, 2020
- HORCH, KERSTIN: Suche von Gesundheitsinformationen im Internet – Ergebnisse der KomPaS-Studie (2021)
- LAMNEK, SIEGFRIED; KRELL, CLAUDIA: *Qualitative Sozialforschung: Mit Online-Materialien*. 6., vollständig überarbeitete Edition. Weinheim Basel : Beltz, 2016 — ISBN 978-3-621-28269-7
- LEHMANN, R; STIELER, M: Chancen und Risiken KI gestützter Evaluationsmethoden textbasierter Kommunikation. In: *Suchttherapie* Bd. 23 (2022), Nr. S 01, S. S39_4. — S39_4
- LEHMANN, ROBERT; ALBRECHT, JENS; ZAUTER, SIGRID: Die Computerunterstützte Analyse Sozialwissenschaftlicher Texte – Ergebnisse des Forschungsprojekts „Casotex“. In: *Gegenwart und Zukunft sozialer Dienstleistungsarbeit: Chancen und Risiken der Digitalisierung in der Sozialwirtschaft, Perspektiven Sozialwirtschaft und Sozialmanagement*. Wiesbaden : Springer Fachmedien Wiesbaden, 2021 — ISBN 978-3-658-32555-8
- LEHMANN, ROBERT; ALBRECHT, JENS; ZAUTER, SIGRID; INSTITUT FÜR E-BERATUNG: *CaSoTex, Computerunterstützte Analyse Sozialwissenschaftlicher Texte* : Technische Hochschule Nürnberg Georg Simon Ohm, 2020
- LIU, SIYANG; ZHENG, CHUJIE; DEMASI, ORIANNA; SABOUR, SAHAND; LI, YU; YU, ZHOU; JIANG, YONG; HUANG, MINLIE: Towards Emotional Support Dialog Systems. In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Online : Association for Computational Linguistics, 2021, S. 3469–3483
- MALHOTRA, GANESHAN; WAHEED, ABDUL; SRIVASTAVA, ASEEM; AKHTAR, MD SHAD; CHAKRABORTY, TANMOY: Speaker and Time-aware Joint Contextual Learning for Dialogue-act Classification in Counselling Conversations. In: *Proceedings of the Fifteenth ACM*

International Conference on Web Search and Data Mining. Virtual Event AZ USA : ACM, 2022 — ISBN 978-1-4503-9132-0, S. 735–745

MAYRING, PHILIPP; FENZL, THOMAS: Qualitative Inhaltsanalyse. In: BAUR, N.; BLASIUS, J. (Hrsg.): *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden : Springer Fachmedien, 2019 — ISBN 978-3-658-21308-4, S. 633–648

MILLER, WILLIAM R.; MOYERS, THERESA B.; ERNST, DENISE; AMRHEIN, PAUL: Manual for the Motivational Interviewing Skill Code (MISC), Version 2.1. unpublished, 2008

MILLER, WILLIAM R.; ROLLNICK, STEPHEN: *Motivierende Gesprächsführung: Motivational Interviewing: 3. Auflage des Standardwerks in Deutsch* : Lambertus-Verlag, 2015. — Google-Books-ID: ffv1DwAAQBAJ — ISBN 978-3-7841-2546-6

MODE-SCHEIBEL, BRIGITTE: Analyse verschiedener Beratungsansätze und Interventionen in der psychosozialen Onlineberatung. In: *e-beratungsjournal.net* Bd. 10. Jahrgang (2014), Nr. Heft 2, Artikel 4, S. 88–102

PAULUS, CHRISTOPH: Empathie und die Big Five (2016)

RAT FÜR SOZIAL- UND WIRTSCHAFTSDATEN (RATSWD): Erhebung und Nutzung unstrukturierter Daten in den Sozial-, Verhaltens- und Wirtschaftswissenschaften. In: *RatSWD Output Paper Series* (2023)

REINDL, RICHARD: Zum Stand der Onlineberatung in Zeiten der Digitalisierung. In: *E-Beratungsjournal* Bd. 14 (2018), S. 16–26

SCHLAPBACH, MARTINA; ETTLIN, RICARDA; SPIESS, MANUELA; RUFLIN, REGULA: *Anwendung von Wirkfaktoren in der Email-Beratung von SafeZone.ch* (Schlussbericht) : Bundesamt für Gesundheit, Sektion Prävention in der Gesundheitsversorgung, 2018

SCHRÖDTER, MARK; BASTIAN, PASCAL: Risikodiagnostik durch Big Data Analytics im Kinderschutz. In: *ARCHIV für Wissenschaft und Praxis der sozialen Arbeit* (2019), Nr. 02, S. 40–49

STIELER, MARA: Zugänge, Einflussfaktoren und Effekte der psychosozialen Onlineberatung von Jugendlichen - eine empirische Erhebung unter User*innen der Plattform Jugendnotmail (2022)

STIELER, MARA; LIPOT, SARAH; LEHMANN, ROBERT: Zum Stand der Onlineberatung in Zeiten der Corona Krise. Entwicklungs- und Veränderungsprozesse der Onlineberatungslandschaft. In: *Fachzeitschrift für Onlineberatung und computervermittelte Kommunikation* Bd. 18 (2022), Nr. 1, S. 50–65

VASWANI, ASHISH; SHAZEER, NOAM; PARMAR, NIKI; USZKOREIT, JAKOB; JONES, LLION; GOMEZ, AIDAN N; KAISER, ŁUKASZ; POLOSUKHIN, ILLIA: Attention is All you Need. In: GUYON, I.; LUXBURG, U. V.; BENGIO, S.; WALLACH, H.; FERGUS, R.; VISHWANATHAN, S.; GARNETT, R. (Hrsg.): *Advances in Neural Information Processing Systems*. Bd. 30 : Curran Associates, Inc., 2017

WATSON, JEANNE C.: Reassessing Rogers' necessary and sufficient conditions of change. In: *Psychotherapy: Theory, Research, Practice, Training* Bd. 44 (2007), Nr. 3, S. 268–273

WU, ZIXIU; BALLOCCU, SIMONE; KUMAR, VIVEK; HELAOUI, RIM; REFORGIATO RECUPERO, DIEGO; RIBONI, DANIELE: Creation, Analysis and Evaluation of AnnoMI, a Dataset of Expert-Annotated Counselling Dialogues. In: *Future Internet* Bd. 15 (2023), Nr. 3, S. 110

WU, ZIXIU; BALLOCCU, SIMONE; KUMAR, VIVEK; HELAOUI, RIM; REITER, EHUD; REFORGIATO RECUPERO, DIEGO; RIBONI, DANIELE: Anno-MI: A Dataset of Expert-Annotated Counselling Dialogues. In: *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Singapore, Singapore : IEEE, 2022 — ISBN 978-1-66540-540-9, S. 6177–6181

XIAO, BO; CAN, DOĞAN; GIBSON, JAMES; IMEL, ZAC E.; ATKINS, DAVID C.; GEORGIU, PANAYIOTIS; NARAYANAN, SHRIKANTH S.: Behavioral Coding of Therapist Language in Addiction Counseling Using Recurrent Neural Networks. In: *Interspeech 2016 : ISCA, 2016*, S. 908–912

YU, JIHYE; LEE, SUKYUNG; KIM, MIRAN; LIM, KIYOUNG; CHANG, KIHONG; LEE, MIJIN: Relationships Between Perspective-Taking, Empathic Concern, and Self-rating of Empathy as a Physician Among Medical Students. In: *Academic Psychiatry* Bd. 44 (2020), Nr. 3, S. 316–319

ZAUTER, SIGRID; LEHMANN, ROBERT: Schwer auffindbare Beratungsangebote in der psychosozialen Onlineberatung. In: FREIER, C.; KÖNIG, J.; MANZESCHKE, A.; STÄDTLER-MACH, B. (Hrsg.): *Gegenwart und Zukunft sozialer Dienstleistungsarbeit: Chancen und Risiken der Digitalisierung in der Sozialwirtschaft, Perspektiven Sozialwirtschaft und Sozialmanagement*. Wiesbaden : Springer Fachmedien, 2021 — ISBN 978-3-658-32556-5, S. 129–140