

Aktennotizerstellung in der Sozialen Arbeit durch Künstliche Intelligenz – Erkenntnisse aus einem Mixed-Method-Forschungsprojekt¹

Christina Plafky, Mitra Purandare, Benjamin Plattner, Svitlana Hrytsai

Abstract: Diese Studie untersucht die Entwicklung und Implementierung einer datenschutzkonformen KI-Anwendung zur automatisierten Erstellung von Aktennotizen in der Sozialen Arbeit. Im Rahmen eines Mixed-Method-Forschungsprojekts wurde ein Prototyp entwickelt, der mittels Speech-to-Text und Large Language Models unstrukturierte Sprachaufnahmen in strukturierte Aktennotizen umwandelt. Die Anwendung wurde als Stand-Alone-Lösung konzipiert, um höchste Datenschutzstandards zu gewährleisten. Durch Usability-Tests mit Fachkräften wurde die Praxistauglichkeit evaluiert. Die Ergebnisse zeigen, dass die KI-Anwendung qualitativ hochwertige Aktennotizen generiert und von den Fachkräften als nützliches Instrument wahrgenommen wird. Allerdings wurden auch Herausforderungen identifiziert, beispielsweise hinsichtlich KI-generierter Halluzinationen. Ob diese Anwendung tatsächlich zur Effizienzsteigerung beiträgt, wird in Frage gestellt, da die generierten Ergebnisse auf Richtigkeit überprüft werden müssen. Die Studie verdeutlicht das Potenzial von KI-Anwendungen in der Sozialen Arbeit und unterstreicht die Bedeutung der interdisziplinären Zusammenarbeit bei deren Entwicklung.

Keywords: Künstliche Intelligenz (KI), Automatisierung, Aktennotiz, Soziale Arbeit, Datenschutz, Dokumentationsunterstützung

1 © Die Autor:innen 2025, veröffentlicht im Verlag Beltz Juventa, Open Access: CC BY 4.0
Gesa Linnemann/Julian Löhe/Beate Rottkemper (Hg.), Künstliche Intelligenz in der Sozialen Arbeit
10.3262/978-3-7799-8562-4_014

1 Einführung²

In der praktischen Sozialen Arbeit ist das Schreiben von Berichten und Aktennotizen ein wesentlicher Bestandteil des Berufsalltags. Es dient der Dokumentation von Fallverläufen der Klient:innen, der effektiven Kommunikation zwischen Fachkräften sowie der Rechenschaftspflicht gegenüber Adressat:innen und anderen Institutionen. Aktennotizen dokumentieren Interaktionen, Einschätzungen und Fortschritte der Adressat:innen. Diese Notizen sind unerlässlich für die Erstellung präziser Fallakten, die Ausarbeitung von Unterstützungsplänen und die Sicherstellung einer kontinuierlichen, zielgerichteten Betreuung. Aus rechtlichen, ethischen und praktischen Gründen sind Aktennotizen von großer Bedeutung für die Fachpraxis. Daher sind qualitative hochwertige Aktennotizen äußerst wichtig. Mittlerweile sind die Einrichtungen zur digitalen Falldokumentation übergegangen. Das genaue Ausmaß der Arbeitszeit, die für das Schreiben von Aktennotizen aufgewendet wird, und die Bewertung dieses Teils des Berufsalltags durch Fachkräfte wurden bisher hauptsächlich im englischsprachigen Raum untersucht (vgl. z. B. Lillis/Leedham/Twiner 2020; MacAlister 2022; Pascoe/Waterhouse-Bradley/McGinn 2022). In einer Studie aus der Schweiz wurde festgestellt, dass die wahrgenommene und tatsächliche Belastung durch administrative Aufgaben je nach Handlungsfeld und Kontext variiert (Plafky et al. 2025).

Die Schwierigkeit, vollständige und qualitativ hochwertige Aktennotizen in der Praxis zu erstellen, wird ebenso in anderen Professionen wie beispielsweise der Medizin thematisiert (vgl. z. B. Earnshaw et al. 2020). Auch wenn sich die Funktion und der Zweck der Aktennotizerstellung in den Professionen Soziale Arbeit, Medizin oder im juristischen Kontext unterscheiden mögen, gibt es doch Überschneidungen in Bezug auf Arbeitsaufwand und Qualitätssicherung. Der Einsatz von KI-Technologien zur medizinischen Dokumentation (inklusive Aktennotizen) wird bereits im Gesundheitswesen, z. B. in der Psychiatrie, Neurochirurgie und Augenheilkunde (vgl. Nguyen/Pepping 2023; Ali et al. 2024; Heilmeyer et al. 2024), sowie im juristischen Kontext (vgl. Regalia 2024) getestet. Die Ergebnisse in Bezug auf Qualität, Vollständigkeit und potenzielle Arbeitserleichterung scheinen vielversprechend. Die Studien zeigen, dass KI-Technologien die Effizienz und Genauigkeit der Dokumentation verbessern können, wobei jedoch Datenschutzaspekte berücksichtigt werden müssen.

In der EU und der Schweiz gelten seit einigen Jahren strenge Datenschutzbestimmungen. Diese sollen den verfassungsmäßigen, grundrechtlichen Anspruch auf informationelle Selbstbestimmung und den Schutz der Persönlichkeit gewährleisten. Dies spiegelt sich in verschiedenen Verordnungen und Gesetzen

2 Zur sprachlichen Verbesserung des Textes wurden ChatGPT und Microsoft Copilot verwendet.

wider (Europäische Datenschutzkonvention des Europarates, Datenschutz-Grundverordnung der EU (DSGVO), Eidgenössisches Datenschutzgesetz (DSG), kantonale Datenschutzgesetze etc.). Die Autonomie der Betroffenen sowie der Schutz vor Missbrauch persönlicher Daten und Diskriminierung durch Datenverarbeitung stehen hierbei im Mittelpunkt (vgl. Mösch Payort/Pärli 2022). (Weitere Details zum Datenschutz in der Sozialen Arbeit finden sich auch im Beitrag „KI und IT-Sicherheit und Datenschutz“ von Jan Pelzl in diesem Band).

Es stellt sich heraus, dass Fachkräfte derzeit verschiedene kommerzielle Anwendungen wie ChatGPT, Diktier-Apps auf dem Handy oder andere Tools zur Unterstützung bei der Aktennotizerstellung nutzen. Diese sogenannte Shadow IT (vgl. Haag/Eckhard 2017) bezeichnet die Nutzung von IT-Ressourcen (Software, Hardware oder Cloud-Dienste) durch Mitarbeiter:innen ohne das Wissen oder die Genehmigung der IT-Abteilung der Einrichtung. Shadow IT birgt erhebliche Risiken, da die IT-Abteilung keine Kontrolle über diese Ressourcen hat. Dies kann zu Sicherheitslücken, Datenverlust und der Missachtung von Datenschutzbestimmungen führen. Die Nutzung von kommerziellen Applikationen wie ChatGPT ist aus Datenschutzgründen ohne vorherige Überprüfung und Freigabe der Einrichtung zu vermeiden.

Es liegt nahe, über den Einsatz einer KI-Technologie nachzudenken, die Fachkräfte der Sozialen Arbeit bei der Erstellung qualitativ hochwertiger Aktennotizen unterstützen kann. Allerdings ist es wichtig, dass diese so konzipiert ist, dass sie den aktuellen Datenschutzbestimmungen für die Nutzung in der Praxis entspricht. Aus diesem Grund steht bei dem Projekt eine datenschutzkonforme KI-Lösung im Fokus. Der vorliegende Beitrag berichtet über die Erfahrungen aus technischer und anwendungsspezifischer Perspektive.

2 Methoden

Im Rahmen des Projekts „Automatisierte Erstellung von Aktennotizen durch eine KI-Anwendung“, finanziert durch die ITBO-Bildungsinitiative des Kantons St. Gallen (Schweiz), wurde von September 2023 bis August 2024 in der Zusammenarbeit zwischen dem Departement Soziale Arbeit, dem Interdisciplinary Center for Artificial Intelligence und dem Institut für Software der Ostschweizer Fachhochschule ein Prototyp im Sinne eines Proof-of-Concept für eine KI-gestützte Applikation zur strukturierten Erstellung von Fallnotizen entwickelt.

Die KI-Anwendung wurde im Prozess mittels Usability-Tests (vgl. Dumas/Reddish 1992; Barnum 2019) auf Praxistauglichkeit, technische Tauglichkeit, Nutzerfreundlichkeit und Qualität überprüft. Usability-Tests umfassten im Entwicklungsprozess Tests mit drei Fachkräften der Sozialen Arbeit in den Räumlichkeiten der Hochschule. Nach der Fertigstellung der KI, in die die Rückmeldungen der drei Personen aus unterschiedlichen Handlungsfeldern

eingeflossen sind, wurde die KI-Anwendung in einer Einrichtung der Sozialen Arbeit mehrere Tage von acht Personen vor Ort getestet. Im Anschluss wurden durch eine Fokusgruppe per Microsoft Teams die Erfahrungen mit vier Personen reflektiert und mit der Template-Analyse-Methode (vgl. Brooks et al. 2015) ausgewertet. Bei den Usability-Tests wurde auf Einhaltung der forschungsethischen Standards der Sozialen Arbeit (vgl. Franz/Unterkofler 2021) geachtet.

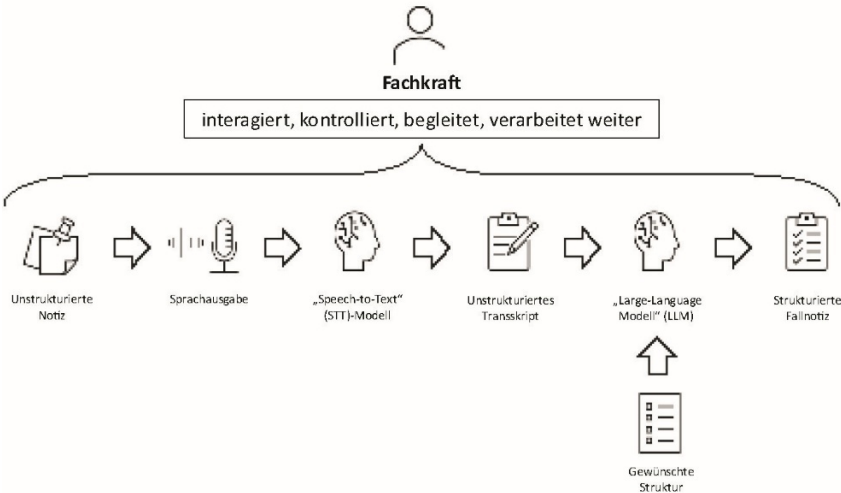
Um den strengen Datenschutzvorgaben gerecht zu werden, wurde diese KI-Anwendung als Stand-Alone-Lösung entwickelt, d. h. ohne Anbindung an Server oder das Internet. Die Schnittstellenproblematik zwischen der KI-Anwendung und einer Fallführungssoftware ist in diesem Proof-of-Concept-Projekt nicht bearbeitet worden und wird durch eine kommerzielle Anwendung gelöst werden müssen, da sehr viele unterschiedliche Fallführungssoftware in den Einrichtungen in Gebrauch sind und für jede Software eine eigene Schnittstellenlösung erarbeitet werden muss. In der aktuellen Version ist es notwendig, die Notizen per Copy-and-Paste zu übertragen.

Während des Projekts kam die Frage auf, ob es möglich wäre, das gesamte Gespräch mit den Adressat:innen (oder auch einer Teamsitzung) aufzuzeichnen und anschließend durch eine KI-Anwendung bearbeiten zu lassen. Technisch wäre dies machbar, jedoch stellt es professionsethisch eine Herausforderung dar. Es erfordert die informierte Zustimmung der Adressat:innen, und die Aufnahme eines Gesprächs ist juristisch ein anderer Sachverhalt, da es sich dann um ein Wortprotokoll (Vollprotokoll) und nicht um ein Gedächtnisprotokoll handelt. Zudem ist zu beachten, dass ein Aufnahmegerät und das anschließende Transkript die Qualität des Gesprächs und den Inhalt des Gesagten beeinflussen können. Ob dies langfristig zu einer guten Beziehungsqualität zwischen Fachkräften und Adressat:innen führt, ist fraglich. Aus diesen Gründen haben wir uns gegen diese Variante entschieden. Der vorliegende Prototyp arbeitet mit Gedächtnisprotokollen, die die Fachkräfte einsprechen müssen.

Diese Applikation ermöglicht es Fachkräften, unstrukturierte Notizen durch Spracheingabe zu erfassen. Die aufgenommene Sprache wird mittels eines sogenannten Speech-to-Text (STT)-Modells in ein unstrukturiertes Transkript umgewandelt. Dieses Transkript, zusammen mit der gewünschten Struktur für die resultierende Fallnotiz, dient als Eingabe für ein zweites Modell, das allgemein als „Large Language Model“ (LLM) (siehe auch den Beitrag von Rottkemper in diesem Band) bekannt ist. Die Spracheingabe und somit das Transkript können unstrukturiert und in Alltagssprache verfasst sein, da das LLM in der Lage ist, die relevanten Fakten aus dem Transkript zu identifizieren und zusammenzufassen. Im letzten Schritt generiert das LLM daraus die strukturierte Fallnotiz. In Abbildung 1 ist der Ablauf schematisch dargestellt.

Fallnotizen enthalten sensible personenbezogene Daten, die als privat und vertraulich gelten. Entscheidend bei deren Verwendung und Aufbewahrung ist, diese Daten angemessen zu schützen und die Kontrolle über den Daten-

Abbildung 1: Schematische Darstellung der KI-unterstützten Erfassung von Fallnotizen



Quelle: Eigene Darstellung

fluss innerhalb der eigenen Organisation – wo möglich – zu behalten. Wenn personenbezogene Daten aus Fallnotizen an externe LLM-Dienste gesendet werden (z. B. OpenAIs ChatGPT oder Anthrops Claude), besteht das Risiko, dass unbefugte Dritte Zugang zu diesen Informationen erhalten, sei es durch Datenlecks, Sicherheitslücken oder andere Schwachstellen. Des Weiteren können LLM-Dienstanbieter Benutzereingaben auswerten, was ebenfalls ein Datenschutzrisiko darstellt. Solche Vorfälle würden die Datenschutzanforderungen verletzen.

Darüber hinaus ist oft unklar, ob externe LLM-Dienstanbieter die Daten aufbewahren und für das Training ihrer eigenen Modelle verwenden. Dies könnte theoretisch dazu führen, dass spezifische Fälle von einem LLM gelernt und unter bestimmten Umständen vom LLM preisgegeben werden. Um diese Risiken zu minimieren und den Datenschutz, Datenbesitz und die Datenkontrolle zu gewährleisten, wird sowohl das STT-Modell als auch das LLM auf der gleichen lokalen Hardware gehostet, auf der die KI-gestützte Anwendung für Fallnotizen betrieben wird.

Die für diesen Prototypen verwendete lokale Hardware kann vollständig offline gehalten werden, da die Anwendung keine Verbindung zum Internet benötigt, um die Fallnotizen zu erstellen. Dies bietet eine zusätzliche Sicherheit, indem verhindert wird, dass Daten die Hardware verlassen. Mit diesem Ansatz wird sichergestellt, dass alle sensiblen Daten innerhalb der Organisation bleiben und nicht an externe LLM-Dienste übermittelt werden müssen.

Ein limitierender Faktor beim Einsatz von Sprachmodellen auf eigener Hardware sind die verfügbaren Ressourcen. Bekannte Sprachmodelle stellen sehr hohe Anforderungen an Hardwarekomponenten, insbesondere an Grafikkarten. Für Geräte mit eingeschränkten Ressourcen (z. B. der im Rahmen dieses Projekts verwendete leistungsstarke Laptop) existieren jedoch Möglichkeiten, Sprachmodelle in kleinerer und komprimierter (quantisierter) Form zu betreiben (vgl. Han/Mao/Dally 2016; Xu et al. 2024). Obwohl die Performanz und die Qualität der Ausgaben dieser Modelle dadurch limitiert sind, generieren sie dennoch ansprechende Resultate. Aktuelle Forschungsarbeiten konzentrieren sich darauf, weitere sogenannte Small Language Models (SLM) zu entwickeln und diese den Endnutzer:innen auf deren Geräten zur Verfügung zu stellen.

Die vorgeschlagene Lösung basiert auf einem weit verbreiteten Open Source-Sprachmodell in der Kategorie der SLM mit der Bezeichnung Mistral 7B in einer 5-Bit quantisierten Form, entwickelt von der Firma Mistral (vgl. Jiang et al. 2023). Zusätzlich wird ein Open Source-STT-Modell namens Whisper von OpenAI verwendet (vgl. Radford et al. 2022). Da beide Modelle Open Source sind und aufgrund ihrer kleineren Größe lokal auf der eigenen Hardware bereitgestellt werden können, ermöglicht die vorgeschlagene KI-basierte Lösung eine umfassende Kontrolle über die Modelle und den Datenfluss, von der Eingabe bis zur Ausgabe und der Speicherung, womit die Integrität der Daten und die Privatsphäre der Adressat:innen gewährleistet wird.

Insgesamt wird eine Lösung vorgeschlagen, die darauf abzielt, den Fachkräften eine Unterstützung bei der Erstellung von qualitativ hochwertigen Fallakten zu bieten. Basierend auf realen und erfundenen Aktennotizen wurde die KI nach einer konkreten Struktur instruiert:

3 Ergebnisse der Usability-Tests

Die auf diesen Ansätzen basierende Applikation stieß in den Usability-Tests auf positive Resonanz. Die Fachkräfte schätzen die Möglichkeit, Informationen zum Fall auf unstrukturierte Weise einzugeben, da dies den Aufwand für die Erstellung von Fallnotizen reduziert. Die Qualität der generierten Zusammenfassungen wird als erstaunlich gut, äußerst nützlich und zeitsparend bewertet. Die strukturierten Fallnotizen können sich laut den Testpersonen insbesondere bei internen Übergaben und der Aktenführung als besonders wertvoll erweisen. Dies liegt daran, dass das LLM unscheinbare Zusammenhänge innerhalb einer unstrukturierten Eingabe erkennen und damit Fallnotizen in klar formulierter Form erstellen kann. Gleichzeitig wurde festgestellt, dass die KI-Anwendung Inhalte erfindet, wenn der eingesprochene Text sehr kurz ist. Die KI versucht, die Vorlage der Aktennotiz zu füllen, und erfindet fehlende Inhalte, falls nicht genügend Informationen vorhanden sind (zu einer Einordnung des Themas

Tabelle 1: Strukturvorgabe für die KI-Anwendung

Struktur der Aktennotiz	
Wer?	Alle Beteiligten benennen
Wann?	Uhrzeit, Datum und Dauer
Wo?	Ort
Grund des	Treffens/Telefonanrufs/virtuellen Meetings
Ablauf des	Treffens/Telefonanrufs/virtuellen Meetings
Diskussionsinhalte/Besprechungsinhalte	alle relevanten Inhalte benennen, eventuell Kontroversen notieren
Getroffene Vereinbarungen	benennen und auflisten
Fachliche Interpretationen und Wahrnehmungen	
Wichtige Interpretation	von Sequenzen, Ereignisse, Verhaltensweisen, Auffälligkeiten, Veränderungen etc.
Wichtige getroffene Entscheidung	benennen und fachlich begründen
Nächster vereinbarter Termin	

Quelle: Eigene Darstellung

Halluzinationen in generativen KI-Systeme siehe den Beitrag von Rottkemper in diesem Band). Einmal wurde im Transkript plötzlich eine unbekannte Sprache verwendet oder es wurden einfach Buchstaben aneinandergereiht, wobei die anschließend daraus generierte Zusammenfassung jedoch korrekt war, da z. B. sich wiederholende Buchstaben vom LLM für eine Zusammenfassung als unwichtig betrachtet werden. Manchmal wurden spezifische Ortsnamen oder Personennamen verändert, was offensichtlich auf eine unklare Aussprache zurückzuführen ist.

Das Einsprechen in den Computer wurde als unproblematisch beschrieben, da alle es gewohnt sind, mit Headsets zu arbeiten. Die Fachkräfte fänden eine Auswahl verschiedener Aktennotiztypen mit unterschiedlichen Vorlagen hilfreich, z. B. eine Vorlage für einen Hausbesuch oder eine für ein kurzes Telefonat zur Terminabstimmung. Ebenso hilfreich wäre die Möglichkeit, Audiodateien direkt in die KI einzuspeisen (beispielsweise von einer Handy-App), ohne den Umweg über das Abspielen und Aufnehmen gehen zu müssen.

Diese fehlende Möglichkeit wurde als unbedeutend angesehen, solange die KI-Anwendung auf dem eigenen Rechner genutzt werden kann. Das Kopieren und Einfügen von Text werden nicht als hinderlich für die Effizienz angesehen. Es wurde festgestellt, dass die Anwendung einfach und intuitiv funktioniert und nach einer sehr kurzen Einweisung keine Schwierigkeiten bereitet. Allerdings erfordert die Anwendung eine gewisse Lernkurve, da man erst durch Ausprobieren herausfindet, wozu die KI fähig ist und wozu nicht. Zur Effizienzsteigerung wur-

de in den Usability-Tests erwähnt, dass der Arbeitsaufwand aufgrund des Kontrollbedarfs gegenüber der KI nicht geringer ist, als wenn man die Aktennotiz selbst schreibt. Es wurde jedoch eingeräumt, dass man die Aktennotiz bei genügend Vertrauen nicht mehr lesen würde. Begrüßt würde eine Erweiterung, sodass nicht Inhalte eingesprochen, sondern auch Texte aus anderen Quellen hinein kopierbar wäre. Ebenso wichtig wäre die Möglichkeit, das Einsprechen zu unterbrechen und später fortzusetzen, da es im Arbeitsalltag häufig zu Unterbrechungen kommt. Derzeit muss das Einsprechen komplett von vorne begonnen werden. Die Erkennung des schweizerdeutschen Dialekts funktioniert teilweise. Dies ist vom Dialekt abhängig, wobei ein Dialekt, der dem Hochdeutschen ähnlicher ist, besser erkannt wird. Dies wurde aber in diesem Projekt nicht tiefgründig analysiert (vgl. Paonessa et al. 2023).

Die Testpersonen sehen jedoch Potenzial darin, dass die Anwendung zur Qualitätssicherung und -steigerung beitragen kann, da die KI gute Zusammenfassungen liefert und die Vorlagen sinnvoll erscheinen.

4 Diskussion

Das Problem von Halluzinationen stellt sowohl bei der Transkription mittels STT als auch bei der Erstellung von Fallnotizen durch LLM eine bedeutende Herausforderung dar. Es gibt jedoch verschiedene Ansätze, um diese Halluzinationen zu minimieren. Eine Möglichkeit besteht in der Anpassung spezifischer Parameter, beispielsweise des Temperaturparameters, der die Kreativität der Modelle steuert. Der Temperaturparameter ist ein Skalierungsfaktor, der die Wahrscheinlichkeitsverteilung der Wortvorhersagen beeinflusst und somit ein Gleichgewicht zwischen der Generierung kreativer (aber potenziell fehleranfälliger) und konservativer (aber möglicherweise weniger interessanter) Ausgaben ermöglicht. Ein weiterer Ansatz zur Verbesserung der Resultate ist die Anwendung von Beam Search, einer Technik, die alternative Resultate erkundet und dadurch die Qualität der Ausgaben erhöhen kann (vgl. Peng/Morton 1988). Darüber hinaus kann das Finetuning eines Sprachmodells auf die Struktur der Aktennotiz oder mit fachspezifischen Begriffen die Genauigkeit weiter steigern. Finetuning verändert die Parameter des Modells so, dass gewünschte Resultate wahrscheinlicher werden (siehe den Beitrag von Rottkemper in diesem Band) (vgl. Wei et al. 2021). Verschiedene Faktenüberprüfungsverfahren, darunter Wissensgraphen (vgl. Hogan et al. 2020), Retrieval-Augmented Generation (RAG) (vgl. Lewis et al. 2021) und Faktenchecks, die von anderen LLM durchgeführt werden (LLM-as-a-Judge, vgl. Zheng et al. 2023), haben das Potenzial, das Risiko von Halluzinationen bei der Erstellung von Fallakten zu senken. Wissensgraphen speichern Beziehungen zwischen Fakten oder Entitäten und können, ähnlich wie der RAG-Ansatz, der verschiedene Dokumente in einer separaten Daten-

bank speichert, das LLM bei der Abfrage von Fakten unterstützen und so zu zuverlässigeren Resultaten führen. Beispielsweise könnte ein Wissensgraph, der fachspezifisches Wissen enthält, verwendet werden, um die Richtigkeit von Diagnosen oder Behandlungsvorschlägen in automatisch generierten Aktennotizen zu überprüfen. Beim LLM-as-a-Judge-Ansatz kann ein weiteres Sprachmodell hinzugezogen werden, um die generierten Resultate zu überprüfen und ggf. korrigierend einzugreifen. Da die Modelle auf Wahrscheinlichkeiten basieren und mit Unsicherheiten behaftet sind, kann eine hundertprozentig korrekte Generierung von Text oder anderen Modalitäten mit heutigem Wissensstand nicht garantiert werden. Das Problem der Halluzinationen ist ein aktuelles Forschungsgebiet, das weiterhin intensiv untersucht wird (vgl. Ji et al. 2023). Eine potenzielle Weiterentwicklung zur verbesserten Identifikation von Halluzinationen könnte die Berücksichtigung des Konfidenzgrades der generierten Wörter oder Wortteile (sogenannte Token) umfassen. Ein vordefinierter, anpassbarer Schwellenwert würde festlegen, ab wann die Konfidenz eines generierten Tokens als eher hoch oder niedrig eingestuft wird. Wörter, bei denen das Modell eine geringe Konfidenz aufweist, könnten auf der Benutzeroberfläche farblich hervorgehoben werden. Dies würde Fachkräften ermöglichen, gezielt jene Stellen zu identifizieren, die wahrscheinlich von einer Überprüfung auf Richtigkeit profitieren würden. Es ist jedoch wichtig zu beachten, dass auch Wörter, die vom Modell mit hoher Sicherheit als korrekt eingestuft werden, fehlerhaft sein können. Daher ist es unerlässlich, dass Fachkräfte die generierte, strukturierte Fallakte vor der Weiterverarbeitung oder dem Einfügen in das Fallaktensystem sorgfältig überprüfen. Aus professionsethischen und rechtlichen Gründen muss die Fachkraft zwingend die Verantwortung für die Aktennotiz übernehmen. Das bedeutet, dass sie auch für potenzielle Fehler, die durch die KI generiert werden, verantwortlich ist, sofern diese nicht vor der Ablage in der Akte korrigiert werden. Dies könnte zur Folge haben, dass die Erstellung einer Notiz durch die KI möglicherweise nicht schneller ist, wenn umfangreiche Korrekturen oder ein abschließendes Durchlesen erforderlich sind. Weitere Forschung ist erforderlich, um diese Fragestellungen genauer zu untersuchen.

Das User-Interface, beispielsweise durch eine Ergänzung der Auswahlmöglichkeiten für unterschiedliche Formen und damit auch Struktur (siehe Tabelle 1) von Aktennotizen (Telefonanruf, Hausbesuch, Meeting etc.), und die Bedienung der Applikation sollten generell nach den Rückmeldungen der Fachkräfte sowie weiteren gezielten Tests mit ausgewählten Fachkräften kontinuierlich erweitert und verbessert werden.

Verschlüsselung der Daten

Der Datenschutz ist ein zentraler Aspekt im Umgang mit Daten von Adressat:innen. Daher wurde ein Ansatz gewählt, bei dem die Modelle lokal und vollständig offline auf einer Maschine betrieben werden. Dies gewährleistet mit hoher Wahr-

scheinlichkeit, dass der Datenfluss die Maschine nicht verlässt. Dennoch können Sicherheitslücken im System oder das Übertragen generierter Fallnotizen mittels USB-Stick auf ein unsicheres oder kompromittiertes System zu einem Datenabfluss führen. Ein möglicher Ansatz, um die Daten auch während der Übertragung zu schützen, ist die Verschlüsselung. Insbesondere im Umgang mit Daten von Adressat:innen ist es wichtig, dass die Daten sowohl im Ruhezustand („at rest“) als auch während der Übertragung („in transit“) verschlüsselt werden, um unbefugtes Auslesen zu verhindern. Eine derartige Lösung kann mit dem entsprechenden Fachwissen und angemessenem Aufwand implementiert werden und wird nachdrücklich empfohlen. Insgesamt stellt die Verschlüsselung einen unverzichtbaren, jedoch technisch anspruchsvollen Prozess dar, der sorgfältige Planung und kontinuierliche Überwachung erfordert.

Zentrale Lösung zum Betreiben der Applikation

Als Alternative zum Betrieb der Applikation auf einem mobilen Endgerät wie einem Laptop kann die Nutzung eines zentralen Servers in Betracht gezogen werden. Diese Option bietet den Vorteil, leistungsfähigere Hardware zu verwenden, die es ermöglicht, größere Sprachmodelle zu betreiben und dadurch schnellere sowie qualitativ hochwertigere Ergebnisse zu erzielen. Zudem trägt der Einsatz leistungsstärkerer Modelle dazu bei, Halluzinationen zu reduzieren (Brown et al. 2020). Ein zentraler Aspekt dieser Lösung ist jedoch die Verschlüsselung der Daten sowohl während der Übertragung als auch bei der Speicherung.

Datenschutz und Nutzung auf dem Mobiltelefon

Die Verschlüsselung der Daten wird insbesondere wichtig, wenn die Daten zentral in das Fallaktensystem integriert werden. Ein möglicher Anwendungsfall ist die Aufnahme einer Sprachnotiz von unterwegs in einer ungestörten Umgebung (z. B. im eigenen Auto) nach einem Hausbesuch. Dazu kann das Mikrofon des Mobiltelefons als Aufnahmegerät dienen. Dabei bietet sich eine zentrale Lösung an, bei der die beiden Sprachmodelle (STT sowie SLM respektive LLM) auf einem Server laufen, der über das mobile Endgerät, z. B. via Virtual Private Network (VPN), erreichbar ist. Daneben ist es von zentraler Wichtigkeit, dass die Daten trotz der sicheren Verbindung eines VPN zusätzlich verschlüsselt werden. Die übertragenen sowie die generierten Daten werden zur Verarbeitung auf dem Server gespeichert, was eine sichere Verschlüsselung unumgänglich macht. Grundsätzlich sind mobile Lösungen eine von Fachkräften gewünschte und nutzenbringende Funktion, wobei unklar ist, wie viele Fachkräfte tatsächlich über dienstliche Mobiltelefone verfügen. Die Nutzung von unverschlüsselten, hochsensitiven Daten auf privaten Geräten birgt erhebliche Sicherheits- sowie Datenschutzrisiken, ausgehend von Malware oder bei Verlust des Gerätes. Für die Nutzung einer KI-Anwendung zur automatisierten Aktennotizerstellung ist es außerdem wichtig, klare Regeln festzulegen, an welchen Orten diese App verwendet werden darf. Beispielsweise

wäre eine Nutzung in einem Café oder auf einer belebten Straße aus datenschutzrechtlichen Gründen problematisch.

Schnittstellen

Funktionen von Applikationen können über sogenannte „Application Programming Interfaces“ (APIs) oder Schnittstellen für andere Anwendungen zugänglich gemacht werden. Ein weit verbreiteter Architekturstil für Schnittstellen ist „Representational State Transfer“ (REST) (Fielding 2000). Ohne näher auf die technischen Details einzugehen, wurde die vorgestellte Applikation unter Verwendung dieses Architekturstils entwickelt, sodass andere Applikationen beispielsweise die Sprachmodelle ansteuern und nutzen könnten. Umgekehrt könnte die Applikation selbst wiederum andere Anwendungen, z. B. ein Fallnotizenverwaltungssystem, ansteuern und die generierten strukturierten Fallnotizen übertragen. Dieser REST-Architekturstil ist ein zentraler Baustein, um neue Applikationen in bestehende Umgebungen integrieren zu können. Der Einsatz einer solchen KI-Applikation kann womöglich dazu führen, dass Arbeitsprozesse im Kontext der Aktennotizerstellung als einfacher und effizienter wahrgenommen werden. Eine Integration in bereits bestehende Fallführungssoftware durch entsprechende Schnittstellen wäre ideal, wurde aber, wie bereits erwähnt, im Rahmen dieses Projekts nicht verfolgt.

Veränderte Arbeitsprozesse

Fragen zu den veränderten Arbeitsprozessen, die sich durch die Nutzung solcher Technologien entwickeln werden, sind noch unbeantwortet. Es lässt sich jedoch festhalten, dass dieses Beispiel einer KI-Anwendung langfristig zu einer effizienteren und qualitativ einheitlicheren Fachpraxis führen könnte. Dennoch werden dadurch neue Herausforderungen entstehen und vor allem andere Fragen an die Fachpraxis gestellt, die derzeit noch schwer abzuschätzen sind, jedoch möglicherweise bereits in ähnlicher Form durch andere Digitalisierungsprozesse aufgetreten sind (vgl. z. B. Kutscher 2024). An dieser Stelle entstehen neue (Forschungs-)Fragen: z. B. Wie werden sich die Rollen und Verantwortlichkeiten der Fachkräfte durch den Einsatz von KI-Anwendungen verändern? Welche neuen Aufgaben und Kompetenzen werden erforderlich sein? Welche Schulungs- und Weiterbildungsmaßnahmen sind notwendig, um Fachkräfte auf den Einsatz von KI-Anwendungen vorzubereiten? Wie können diese Maßnahmen effektiv gestaltet werden? Wie kann die Zusammenarbeit zwischen IT-Spezialist:innen und Fachkräften der Sozialen Arbeit verbessert werden, um die Implementierung von KI-Anwendungen zu unterstützen?

Nicht zu unterschätzen sind hier die Anforderungen an die Fachkräfte, sich mit immer neuen IT-Anwendungen konfrontiert zu sehen und kontinuierlich digitale Kompetenzen aufbauen zu müssen. Dies erfordert nicht nur technisches Wissen, sondern auch die Fähigkeit, sich an ständig wechselnde Technologien

und Arbeitsprozesse anzupassen. Gleichzeitig erfordert es von Einrichtungen, neue Handlungsleitlinien zu entwickeln, die die Nutzung und Anwendung solcher Applikationen festlegen. Intuitiv bedienbare und sorgfältig integrierte Applikationen erleichtern dabei die Akzeptanz.

5 Schlussfolgerung

Zusammenfassend lässt sich durch dieses Forschungsprojekt sagen, dass eine KI-Lösung für die automatisierte Aktennotizerstellung als datenschutzkonforme Stand-Alone-Lösung erfolgreich entwickelt werden kann. Es hat sich gezeigt, dass diese Anwendung in der Fachpraxis auf Interesse stößt und zur Erhöhung der Arbeitseffizienz sowie zur Sicherstellung von Qualitätsstandards bei Aktennotizen innerhalb einer Einrichtung beitragen könnte.

Grundsätzlich wurde die Qualität der Aktennotizen als gut empfunden, aber nur durch die Auswertung durch weitere Fachkräfte und über einen längeren Zeitraum kann dies und eine potenzielle Effizienzsteigerung umfänglich beurteilt werden. Halluzinationen treten unausweichlich auf. Mit Faktenüberprüfungsverfahren oder Wissensgraphen können die Halluzinationen allenfalls reduziert, aber mit heutigem Wissensstand nicht beseitigt werden.

Strenge Datenschutzanforderungen bringen für die Anwendung von KI-Technologien in der Praxis Sozialer Arbeit einige Nachteile mit sich. Beispielsweise erfordert die Einrichtung einer sicheren App zur Aufnahme von Aktennotizen auf einem Mobiltelefon eine sachkundige Umsetzung durch Fachpersonen. Ein weiteres Problem sind bei Stand-Alone-Lösungen die Schnittstellenlösungen.

Der Einsatz einer solchen KI-Anwendung in der Praxis wurde von den Testpersonen begrüßt und als hilfreich bezeichnet, dies gilt insbesondere für die Zusammenfassungen und die Qualitätssicherung. Konkret ist eine datenschutzkonforme KI-Anwendung eine Möglichkeit für Einrichtungen, zu verhindern, dass Fachkräfte andere, kommerzielle KI-Anwendungen nutzen, die den Datenschutz nicht sicherstellen können.

Gleichzeitig gibt es jedoch noch einige offene Fragen, die für den konkreten Einsatz in der Praxis durch weitere Projekte geklärt werden müssen. Dazu gehören die langfristigen Auswirkungen solcher technischen Lösungen auf den Arbeitsalltag und die Fallarbeit, die konkrete, messbare Effizienzverbesserung, der Schulungsbedarf/ die Lernkurve sowie die Bereitschaft, solche Anwendungen in verschiedenen Handlungsfeldern und Arbeitskontexten dauerhaft zu nutzen. Festzuhalten ist, dass bei der Entwicklung solcher KI-Lösungen zwingend Fachkräfte der Sozialen Arbeit von Beginn an beteiligt sein müssen. Die Hürden, die bei der Entwicklung auftauchen, haben einen direkten Bezug zum Arbeitsalltag der Fachkräfte und können nur gemeinsam gelöst werden (Co-Creation). Dies erhöht die Wahrscheinlichkeit signifikant, dass solche Lösungen in der Praxis

nicht nur akzeptiert, sondern auch intensiv genutzt werden, da die Anwendung optimal auf die Bedürfnisse der Fachkräfte vor Ort angepasst ist. Die im Rahmen dieses Projektes initiierte Zusammenarbeit zwischen der Sozialen Arbeit und der Software-Entwicklung ist ein erfolgreiches Beispiel für eine interdisziplinäre Entwicklung von KI-basierten Lösungen für die Praxis.

Literatur

- Ali, Abdullah/Kumar, Rohit Prem/Polavarapu, Hanish/Lavadi, Raj Swaroop/Mahavadi, Anil/Legarreta, Andrew D./Hudson, Joseph S./Shah, Manan/Paul, David/Mooney, James/Dietz, Nicholas/Fields, Daryl P./Hamilton, D. Kojo/Agarwal, Nitin (2024): Bridging the Gap: Can Large Language Models Match Human Expertise in Writing Neurosurgical Operative Notes? In: *World Neurosurgery* 192, S. e34-e41.
- Brooks, Joanna/McCluskey, Serena/Turley, Emma/King, Nigel (2015): The utility of template analysis in qualitative psychology research. *Qualitative Research in Psychology* 12(2), S. 202–222. <https://doi.org/10.1080/14780887.2014.955224>
- Brown, Tom/Mann, Benjamin/Ryder, Nick/Subbiah, Melanie/Kaplan, Jared D./Dhariwal, Prafulla/Neelakantan, Arvind/Shyam, Pranav/Sastry, Girish/Askell, Amanda/Agarwal, Sandhini/Herbert-Voss, Ariel/Krueger, Gretchen/Henighan, Tom/Child, Rewon/Ramesh, Aditya/Ziegler, Daniel/Wu, Jeffrey/Winter, Clemens/Hesse, Chris/Chen, Mark [...] Amodei, Dario (2020): Language Models are Few-Shot Learners. 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada. arXiv preprint <https://arxiv.org/abs/2001.02038>
- Earnshaw, Charles H./Pedersen, Amanda/Evans, Jo/Cross, Tina/Gaillemin, Oliver/Vilches-Moraga, Arturo (2020): Improving the quality of discharge summaries through a direct feedback system. *Future Healthcare Journal* 7(2), S. 149–154.
- Fielding, Roy Thomas (2000). Architectural styles and the design of network-based software architectures (Doctoral dissertation, University of California, Irvine). <https://ics.uci.edu/~fielding/pubs/dissertation/top.htm> (Abfrage: 15.06.2025).
- Franz, J., & Unterkofer, U. (Eds.). (2021). *Forschungsethik in der Sozialen Arbeit: Prinzipien und Erfahrungen*. Verlag Barbara Budrich.
- Haag, Steffi/Eckhardt, Andreas (2017): Shadow it. *Business & Information Systems Engineering* 59, S. 469–473.
- Han, Song/Mao, Huizi/Dally, William J. (2015): Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. arXiv preprint <https://arxiv.org/abs/1510.01545>
- Heilmeyer, Felix/Böhringer, Daniel/Reinhard, Thomas/Arens, Sebastian/Lysenko, Lisa/Haverkamp, Christian (2024): Viability of Open Large Language Models for Clinical Documentation in German Health Care: Real-World Model Evaluation Study. In: *JMIR Medical Informatics* 12, e59617.
- Hogan, Aidan/Blomqvist, Eva/Cochez, Michael/Melo, Gerard D./Gutierrez, Claudio/Kirrane, Sabrina/Labra Gayo, José Emilio/Navigli, Roberto/Neumaier, Sebastian/Ngonga Ngomo, Axel/Cyrille/Polleres, Axel/Rashid, Sabbir M./Rula, Anisa/Schmelzeisen, Lukas/Sequada, Juan F./Staab, Steffen/Zimmermann, Antoine/d'Amato, Claudia (2021): Knowledge graphs. *ACM Computing Surveys (Csur)* 54(4), S. 1–37. <https://doi.org/10.1145/3447772>
- Ji, Ziwei/Lee, Naon/Frieske, Rita/Yu, Tiezheng/Su, Dan/Xu, Yan/Ishii, Etsuko/Bang, Yejin/Madotto, Andrea/Fung, Pascale (2023): Survey of hallucination in natural language generation. *ACM Computing Surveys* 55(12), S. 1–38. <https://doi.org/10.1145/3571730>
- Jiang, Albert Q./Sablayrolles, Alexandre/Mensch, Arthur/Bamford, Chris/Chaplot, Devendra S./de las Casas, Diego/Bressand, Florian/Lengyel, Gianna/Lample, Guillaume/Saulnier, Lucile/La-

- vaud, L  lio Renard/Lachaux, Marie-Anne/Stock, Pierre/le Scao, Teven/Lavril, Thibaut/Wang, Thomas/Lacroix, Timoth  e/Sayed, William E. (2023): Mistral 7B. arXiv preprint <https://arxiv.org/abs/2308.08839>
- Kutscher, Nadia (2024): Digitalit  t und Digitalisierung als Gegenstand der Sozialen Arbeit. In: Kurtz, Thomas/Meister, Dorothee M./Sander, Uwe (Hrsg.): *Digitale Medien und die Produktion von Wissenschaft*. Wiesbaden: Springer VS, S. 121–146. https://doi.org/10.1007/978-3-658-42542-5_8
- Liang, Percy/Bommasani, Rishi/Lee, Tony/Tsipras, Dimitris/Soylu, Dialara/Yasunaga, Michihiro/Zhang, Yian/Narayanan, Deepak/Wu/Kumar, Ananya/Newman, Benjamin/Yuan, Binhang/Yan, Bobby/Zhang, Ce/Cosgrove, Christian/Manning, Christopher D./R  , Christopher/Acosta-Navas, Diana/Hudson, Drew A./Zelikman, Eric/Durmus, Esin [...] Koreeda, Yuta (2022): Holistic Evaluation of language models. arXiv preprint <https://arxiv.org/abs/2205.12244>
- Lillis, Theresa/Leedham, Maria/Twiner, Alison (2020): Time, the written record, and professional practice: The case of contemporary social work. In: *Written Communication* 37(4), S. 431–486.
- Lillis, Theresa/Twiner, Alison/Balkow, Michael/Lucas, Gillian/Smith, Miriam/Leedham, Maria (2023): Reflections on the procedural and practical ethics in researching professional social work writing. In: *Journal of Applied Linguistics and Professional Practice* 18(3), S. 315–342. <https://doi.org/10.3138/jalpp.20014>
- MacAlister, Josh (2022): The independent review of children's social care – Final report, S. 178–187. webarchive.nationalarchives.gov.uk/ukgwa/20230308122535mp_/https://childrensocialcare.independent-review.uk/wp-content/uploads/2022/05/The-independent-review-of-childrens-social-care-Final-report.pdf (Abfrage: 15.06.2025).
- M  sch Payot, Peter/P  rli, Kurt (2022): Datenschutz in der Sozialen Arbeit. Eine Praxishilfe zum Umgang mit sensiblen Personendaten. *AvenirSocial I – Berufsverband Soziale Arbeit Schweiz*. https://avenirsocial.ch/wp-content/uploads/2023/01/Datenschutz-i-d-SA_db_120123.pdf (Abfrage: 15.06.2025).
- Nguyen, Josh/Pepping, Christopher A. (2023): The application of ChatGPT in healthcare progress notes: A commentary from a clinical and research perspective. In: *Clinical and translational medicine* 13(7), e1324. <https://doi.org/10.1002/ctm2.1324>
- Ow, Peng Si/Morton, Thomas E. (1988): Filtered beam search in scheduling. In: *The International Journal of Production Research* 26(1), S. 35–62.
- Paonessa, Claudio/Schraner, Yanick/Deriu Jan/H  rliemann Manuela/Vogel, Manfred/Cieliebak, Mark (2023): Dialect Transfer for Swiss German Speech Translation. arXiv preprint <https://arxiv.org/abs/2308.08839>
- Pascoe, Kathryn Margaret/Waterhouse-Bradley, Bethany/McGinn, Tony (2023): Social workers' experiences of bureaucracy: A systematic synthesis of qualitative studies. *The British Journal of Social Work* 53(1), S. 513–533. <https://doi.org/10.1093/bjsw/bcac106>
- Radford, Alec/Kim, Jim Wook/Xu, Tao/Brockman, Greg/McLeavey, Christine/Sutskever, Ilya (2023): Robust speech recognition via large-scale weak supervision. In: *International conference on machine learning*, S. 28492–28518. PMLR.
- Regalia, Joe (2024): From Briefs to Bytes: How Generative AI is Transforming Legal Writing and Practice. In: *Tulsa Law Review* 59, 193. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4371460 (Abfrage: 15.06.2025.).
- Wei, Jason/Bosma, Maartens/Zhao, Vincent Y./Guan, Kelvin/Yu, Adams Wei/Lester, Brian/Du, Nan/Dai, Andrew M./Le, Quoc V. (2021): Finetuned language models are zero-shot learners. arXiv preprint <https://arxiv.org/abs/2109.00859>
- Xu, Jiajun/Li, Zhiyuan/Chen, Wei/Wang, Qun/Gao, Xin/Cai, Qi/Ling, Ziyuan (2024): On-Device Language Models: A Comprehensive Review. arXiv preprint <https://arxiv.org/abs/2406.12244>