

Herausforderung im Gesundheitswesen

Bruchfreie Kommunikation zum
Wohle des Patienten



Die Medizin ist seit antiken Zeiten eine Profession der Kommunikation. Der grundlegende Kommunikationsbegriff umfasst hierbei die Interaktion zwischen dem Patienten und dem Arzt. Der Austausch über die Belange des Patienten findet meist im persönlichen Gespräch statt und muss unweigerlich auch in schriftlicher und digitaler Form festgehalten werden. Nicht nur aufgrund rechtlicher Vorschriften, die eine sorgfältige Dokumentation fordern, sondern auch als Erinnerungs- und Wissensgrundlage der behandelnden Ärzte spielen diese Informationen eine große Rolle in der Behandlung des Patienten¹.

Die aktuelle Herausforderung im Gesundheitswesen besteht darin, moderne Kommunikationsformen nutzbar zu machen, ohne das Vertrauensverhältnis zwischen Arzt, Pflegepersonal und Patient durch den Einfluss Dritter zu beeinträchtigen oder gar zu zerstören.

Interaktion Arzt-Patient.

- Persönlicher Austausch und sorgfältige Dokumentation sind essentiell
 - Moderne Kommunikationsformen nutzbar machen
-

Schnittstellen überwinden

Schnittstellen treten immer im Kontakt zwischen zwei Systemen auf. Eine Schnittstelle ist der Kontakt zum Patienten wie der Kontakt zu den Krankenkassen, zu den Krankenhäusern, zu den niedergelassenen Ärzten, zu den nicht medizinischen Berufen, kurz zwischen allen beteiligten Akteuren des Gesundheitswesens untereinander. Laut einer Untersuchung des Instituts für Sozialmedizin der Universität Lübeck entstehen in Deutschland gegenwärtig besondere Herausforderungen durch ein Transparenzdefizit, ein Kommunikationsdefizit und ein Kooperationsdefizit².

Dabei erscheint es illusorisch, diese Defizite einfach zu lösen. Denn die Komplexität des Gesundheitssystems fordert das Aufteilen verschiedener Behandlungsschritte. Dabei erscheint jedoch jeder Prozess wünschenswert, der die an den Schnittstellen entstehenden genannten Defizite abbaut oder reduziert. Neben persönlicher Kompetenz und Kommunikationsvermögen liegt hier die Betrachtung organisatorischer und damit verbundener technische Aspekte sowie die Analyse von Prozessen auf der Hand.

Das universelle Medium

Ein zentraler Aspekt dieser technologischen Möglichkeiten besteht in der Erfassung und Verarbeitung der durch Anamnese, Diagnostik und Therapie des Patienten entstehenden Gesundheitsdaten.

In der Erfassung und Verarbeitung von Daten existieren die verschiedensten Systeme, die idealerweise über Schnittstellen, wie zum Beispiel HL7 miteinander verbunden sind. Es gibt eine Vielzahl von Benutzeroberflächen, Bedienungskonzepten, Formularen, Masken und spezifischen Anpassungen. Dabei sollten die Veränderungen hin zu einer sektorübergreifenden, gemeinsamen Versorgung auch ein Verständnis von einer sektorübergreifenden, gemeinsamen technologischen Organisation mit sich bringen.

Das Erwarten von Sprache

Eine „hörende Technologie“ basiert auf maschinellen Prozessen, die unter dem Begriff Spracherkennung bekannt sind. Bereits dieser Begriff beinhaltet jedoch eine sprachliche Unschärfe: Denn das Erkennen von Sprache ist nur möglich durch das Erwarten von Sprache. Eine Sprache zu erwarten und die Sprachäußerung richtig in Schrift oder Handlung umzusetzen setzt eine Kenntnis der Sprache voraus. Sollte beispielsweise ein Radiologe das Wort „Flair“ verwenden, so meint er mit großer Wahrscheinlichkeit eine FLAIR-Sequenz. Damit hat er die Erwartung, dass das Wort Flair in Großbuchstaben, also FLAIR, geschrieben wird. Der die gleiche Technologie einsetzende Psychotherapeut wird mit dem Wort „Flair“ hingegen die Atmosphäre einer Therapiesitzung beschreiben wollen und hier keine gesonderte Schreibweise erwarten.

Eine gelungene Spracherkennung stellt also die gelungene Erfüllung einer Erwartung dar. Dabei gilt diese Erwartung in zwei Richtungen. Sowohl der Nutzer eines solchen Systems erwartet, dass er „erkannt“ wird, als auch die Maschine „erwartet“ einen spezifischen Sprecher, um die richtige Schreibweise, oder um es plakativ zu formulieren, sogar die richtige Sprache auszuwählen.

Wie schon in der Antike kommt dabei der Sprache eine besondere Bedeutung zu: Die Sprache ist das universelle Kommunikationsmedium. Die „hörende Technologie“ stellt dabei also auf eine verblüffend einfache Art und Weise eine einheitliche, sektorübergreifende Schnittstelle dar. Somit stellt die zunehmende Nutzung von Sprache in der Technologie einen wichtigen Beitrag zur Überwindung von zahlreichen Schnittstellen, sowohl organisatorischer als auch inhaltlicher Art dar.

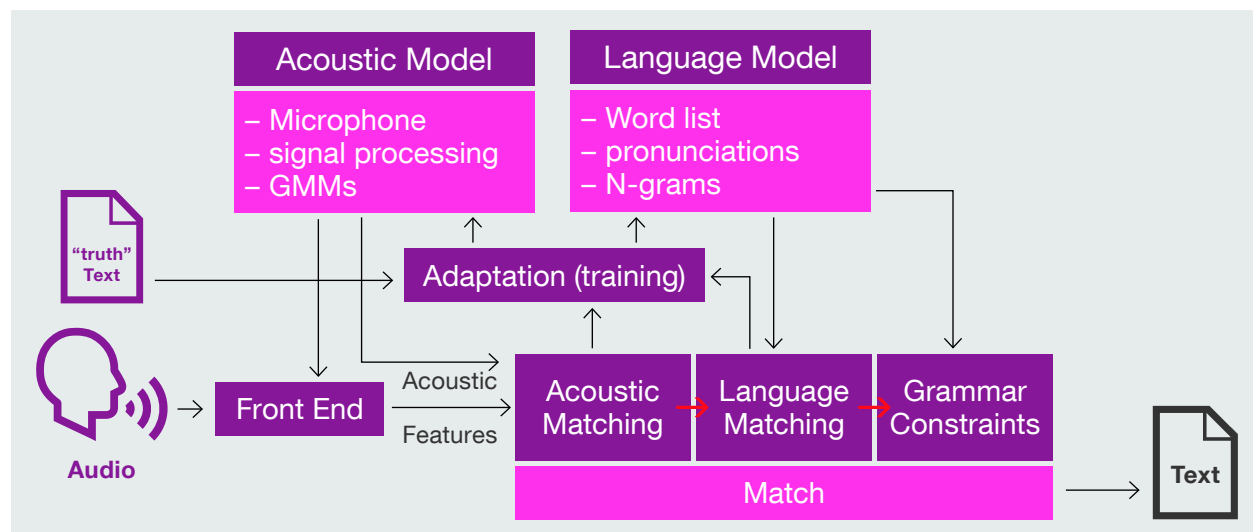


Abbildung 1: Das Erkennen von Sprache basiert auf der richtigen Erwartung, die durch Berücksichtigung von Akustik, Sprache und Grammatik gesetzt wird³

Dieses Beispiel soll deutlich machen, dass eine gelungene Spracherkennung nur möglich ist, wenn die Maschine eine Sprecherin oder einen Sprecher mit seinen spezifischen Eigenschaften richtig erwartet. Neben der grundsätzlichen Zuordnung zu einem beispielsweise medizinischen Umfeld (etwa im Gegensatz zu einem mathematisch-naturwissenschaftlichen Umfeld), geschieht dies durch Zuordnung zu einem sogenannten Sprachprofil.

In den letzten Jahren wurden erhebliche Anstrengungen unternommen – insbesondere im medizinischen Bereich – diese Sprachprofile angepasst an Sprecher aus dem medizinischen Bereich sowohl zu erstellen, als auch vor dem Hintergrund verschiedenster Technologien im Gesundheitssystem an jedem Ort verfügbar zu machen. Dabei sind Lösungen sowohl mit, als auch ohne Nutzung von Netzwerken verfügbar.

Auch eine den Erwartungen entsprechende Spracherkennung hat zunächst keinerlei Verständnis vom Inhalt des Gesagten. Erst eine umfassend funktionierende Spracherkennung öffnet den Anwendungshorizont für eine tiefgehende inhaltliche Analyse des Gesagten, die eine reine Command and Control-Ebene verlässt.

Die technische Meisterleistung

Auch Spracherkennung folgt grundsätzlichen Gesetzmäßigkeiten der Informationstechnologie: Der Eingabe folgt die Verarbeitung, folgt die Ausgabe. Wie in Abbildung 1 dargestellt wird das eingehende Audiosignal verarbeitet. Im Rahmen der Front-end Verarbeitung wird das Signal zunächst prozessiert, wobei die wichtigsten Informationen für die Spracherkennung als Repräsentation der Originalsprache extrahiert werden. Diese Repräsentation wird als Feature Extraction bezeichnet. Die Audioinformation wird dabei zunächst in Blöcke zerlegt. Nach einer Fourier Transformation und Verarbeitung des Spektrums wird die Anzahl der enthaltenen Frequenzbänder reduziert. Der gesamte Prozess führt zu einer kompakten Darstellung des Frequenzspektrums in Form von Mel-Frequenz-Cepstrum-Koeffizienten. Diese Koeffizienten werden auch für die Erkennung von Musik eingesetzt⁴.

³ Nach: Makhoul J, Schwartz R. Proc Natl Acad Sci USA. 1995 Oct 24;92(22): 9956–63.

⁴ Schukat-Talamazzini, Ernst Günter. Automatische Spracherkennung: Grundlagen, statistische Modelle und effiziente Algorithmen. Springer-Verlag, 2013.

Damit die entstehenden die Sprache der präsentierenden Vektoren zur Suche des entsprechenden Textes verwendet werden können ist eine zugrundeliegende Datenbank notwendig, die grundlegende Informationen über die Phoneme als akustische Bestandteile, den Dialekt, das Alter, die Qualität des aufgenommenen Signals, einfache Interpretation des Textes (zum Beispiel durch Formation von Einheiten), die phonetische Beschreibung der Wörter selbst sowie Informationen über verschiedene Schreibweisen enthält.

Mit diesen Informationen kann bereits eine hohe grundlegende Erkennungsqualität erzielt werden, sofern der Sprecher über eine flüssige Sprache mit korrekter Grammatik verfügt. Dabei stellt eine falsche Grammatik die Spracherkennung vor weitaus höhere Herausforderungen, als ein Dialekt, der beispielsweise durch Deutsch als Fremdsprache entsteht. Ursache hierfür ist das in Abbildung 1 dargestellte Language Model (Sprachmodell). Denn das Sprachmodell sagt Wortfolgen vorher. Im Handybereich ist dieses Prinzip gut an den Tastaturen zu erkennen, die bereits während der Eingabe das nächste wahrscheinliche Wort vorschlagen. Im Bereich der Spracherkennung hilft das Sprachmodell, akustisch ähnliche Alternativen voneinander zu unterscheiden (z. B. „FLAIR“ oder „Flair“, „Kante“ oder „kannte“, usw.). Dies erklärt, warum die Spracherkennung vor eine größere Herausforderung gestellt ist, wenn die Wörter an einer falschen Stelle im grammatikalischen Kontext erscheinen.

Die Ausgabe des Textes erfolgt also nach akustischer, sprachlicher und grammatikalischer Verarbeitung. Dabei wird für jeden Sprecher durch die Benutzung des Systems ein Sprachprofil aufgebaut, in dem die persönlichen Modifikationen des statistischen Modells erfasst werden. Es wurden erhebliche Fortschritte gemacht, wobei es zurzeit nicht verwunderlich ist, dass der menschliche Erkennungsprozess insbesondere bei Vorliegen von Umgebungsgeräuschen oder kreativen Neuschöpfungen von Wörtern weiterhin überlegen ist.

Talk and Touch statt Point, Klick and Type

Eine hörende Maschine bietet insbesondere im Gesundheitsbereich faszinierende Möglichkeiten. Der technologische Übergang von Maus und Tastatur hin zu Sprechen und Berühren kann eine neue Ära der Datenerfassung und Bedienung von Maschinen einleiten.

Dazu ist neben der Perfektionierung der im letzten Abschnitt genannten Funktionsweise offensichtlich auch der Zugriff zu einer Erkennungsmaschine mit dem darauf hinterlegten Sprachprofil notwendig. Aktuelle Technologien sorgen insbesondere im komplexen Umfeld von Krankenhäusern und Arztpraxen für eine Verfügbarkeit der Spracherkennung und eine Verbindung zu dem Sprachprofil in jeder erdenklichen Software-Umgebung.

Doch damit entstehen nicht nur neue Möglichkeiten für die Textverarbeitung. Eine Maschine, die Sprache perfekt erfasst, kann über einen Hub aus der Sprache Steuerungsbefehle oder in Echtzeit strukturierte Information extrahieren. Für Gerätehersteller entstehen vollkommen neuartige Möglichkeiten, eine Sprachinfrastruktur zu nutzen und damit dem Gerät das Hören beizubringen.

5 Gute Gründe für Spracherkennung in der Medizin finden Sie unter

 www.nuance.de/healthcare

Über Nuance Healthcare

Nuance Healthcare, ein Unternehmensbereich von Nuance Communications, bietet Spracherkennungslösungen für Ärzte in 22 Sprachen. Damit können diese Patientendaten exakt erfassen und in aussagekräftige, verwertbare Information umwandeln. Heute nutzen weltweit mehr als 10.000 Einrichtungen und 450.000 Benutzer Spracherkennungstechnologie von Nuance – mit positiven Auswirkungen auf die Versorgungsqualität, Profitabilität und Rechtssicherheit. Wenn Sie mehr darüber erfahren möchten besuchen Sie uns auf www.nuance.de oder schreiben Sie uns an info.dragonmedical@nuance.com.