

„Sprich mit mir, Maschine!“

Wie künstliche Intelligenz Sprache verarbeitet

Intelligente Maschinen könnten unseren sprachlichen Alltag auf vielerlei Weise erleichtern, und sie tun es zum Teil bereits: Sie könnten von einer Sprache schnell in eine andere übersetzen, uns eine Zusammenfassung eines umfangreichen Buches liefern und vor allem unsere Sprache verstehen und mit uns kommunizieren. Natural Language Processing (NLP) nennt sich in der Fachsprache jene Technik, die intelligente Maschinen wie den Computer in die Lage versetzt, menschliche Sprache zu verstehen und anzuwenden. Die aus Meran stammende Gründerin einer KI-basierten Innovationsplattform Veronika Gamper und Klaus Diepold von der Technischen Universität München erklärten in einem Vortrag am 17.10.2022 in Bozen, was NLP leisten kann und wo die Grenzen, Herausforderungen und auch möglichen Gefahren durch Missbrauch dieser Technik liegen. Ein Gespräch mit Veronika Gamper als Nachlese:

Natural Language Processing (NLP) wird im Deutschen oft mit „linguistische Datenverarbeitung“ übersetzt. Was ist damit gemeint?

Veronika Gamper: Natural Language Processing (NLP) ist ein interdisziplinäres Fachgebiet an der Schnittstelle zwischen Informatik und Linguistik und mittlerweile ein Teilbereich der künstlichen Intelligenz. Ziel ist es, einen Computer zu befähigen, den Inhalt von gesprochener oder geschriebener Sprache zu „verstehen“. Der Computer ist dann in der Lage, die in Texten oder gesprochenen Sätzen enthaltenen Informationen herauszuholen und Erkenntnisse abzuleiten. So kann man beispielsweise Texte gewissen Themen zuordnen, ähnliche Texte finden, Texte als positiv oder negativ behaftet einstufen. Neueste Sprachmodelle können in Texten Antworten auf Fragen finden, Texte automatisiert zusammenfassen, sie vervollständigen und selbst vollkommen neue Texte generieren.

Im persönlichen Alltag begegnen wir NLP z.B. über digitale Assistenten wie Alexa oder Siri. Unternehmen nutzen NLP bisher v.a. für Chatbots, welche die Kommunikation mit Kund*innen teilweise automatisieren, sowie für die Analyse von Kundenfeedback auf Webseiten oder Social Media. Vor allem die neuen KI-basierten Sprachmodelle erlauben unzählige neue Einsatzmöglichkeiten: von Übersetzungen und Zusammenfassungen bis hin zu Antworten auf Fragen oder computergenerierten Textpassagen. Letztlich können sogar in Sätzen formulierte gewünschte Bildinhalte von einer KI verstanden werden, welche daraus das gewünschte Bild in weniger als einer Minute erstellt.

Welche Anwendungsbereiche von NLP betrachten Sie derzeit als die vielversprechendsten?

Aktuell wird an Anwendungen von NLP quer über alle Industrien hinweg gearbeitet. Neueste Sprachmodelle erlauben es, Programmcode verständlich zu erklären und umgekehrt aus Anweisungen in natürlicher Sprache Programmcode zu erstellen. Somit ermöglichen sie es auch Anfänger*innen, komplizierten Code zu entwickeln, oder Laien, Programmcode zu verstehen. Texte werden beinahe einwandfrei in andere Sprachen übersetzt, automatisierte Zusammenfassungen geben die wichtigsten Inhalte aus langen Texten wieder, und aus wenigen Stichworten entstehen heute ganze Paragraphen im gewünschten Sprachstil.

In den nächsten Jahren werden wir NLP in unserem Arbeitsalltag als ständigen digitalen Begleiter wahrnehmen, der uns beim Erfassen und Verfassen von Text unterstützt und uns das Erstellen bestimmter standardisierter Textformen abnimmt. Neueste Entwicklungen befassen sich mit dem Übersetzen zwischen Textbeschreibungen und Bildern und ermöglichen es, mit einer



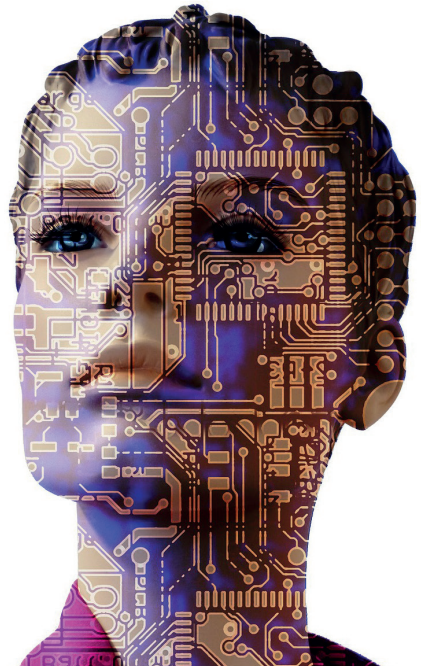
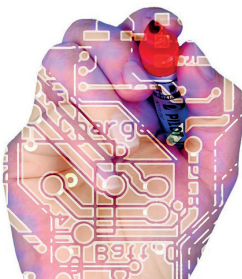
kurzen Anweisung – einem „Prompt“ – ein gewünschtes Bild neu zu generieren. Dies könnte das Berufsbild des Grafikdesigners, der Grafikdesignerin stark wandeln. Da heute viele Inhalte im Netz im Bild- und Videoformat entstehen, sehe ich hier im Zusammenwachsen der Disziplinen NLP und Computer Vision enorme Wachstumsmöglichkeiten.

Und womit befassen Sie sich selbst bei Ihrer Arbeit?

Ich bin Geschäftsführerin und Gründerin von WeDaVinci, einer NLP-basierten Plattform für Innovationsmanagement. Der Schlüssel zu erfolgreicher Innovation liegt in der Zusammenarbeit und dem Feedback von Nutzer*innen und Expert*innen mit unterschiedlichem Wissen. Wir nutzen NLP, um Transparenz zu den Beiträgen auf der Plattform zu generieren. Darüber können wir Projekte und Expert*innen vernetzen, automatisiert Fragen stellen, welche die Projekte voranbringen und Antworten liefern, um Ideen zu ausgereiften Projekten zu entwickeln.

Unsere Sprache ist voll von Mehrdeutigkeiten, wir sprechen nicht nur Standardsprachen, sondern viele Dialekte, und manches, was wir sagen oder schreiben, meinen wir ironisch. Sind das die größten Herausforderungen für Maschinen?

Ein Bild in Worten beschreiben und der Computer erstellt es. Auch das wird zukünftig möglich sein.



Die genannten Themen – Mehrdeutigkeiten, Dialekte bzw. weniger häufig gesprochene Sprachen, und Ironie – sind tatsächlich größere Herausforderungen für Algorithmen. Diesbezüglich wurden letzthin jedoch große Fortschritte erzielt, dank der auf Deep Learning basierenden Sprachmodelle: Das heißt der Computer lernt eigenständig, mithilfe von Unmengen an Daten. Diese Sprachmodelle ermöglichen es, Mehrdeutigkeiten korrekt zu verstehen, da sie den Kontext mit einbeziehen: So können sie z.B. zwischen einer „Bank“ als Geldinstitut und einer „Bank“ als Sitzbank unterscheiden.

Bei Dialekten liegt die Schwierigkeit an der geringeren Anzahl an Trainingsdaten, und das Forschungs- und wirtschaftliche Interesse ist geringer als bei weitverbreiteten Sprachen wie z.B. Englisch, Deutsch, Italienisch. Auch hier dürften die neuen Sprachmodelle jedoch bereits anhand weniger Beispiele imstande sein, einen gewünschten Dialekt zu imitieren. Was bleibt, ist jedoch eine gewisse Verzerrung der Ergebnisse in Bezug auf die ursprünglichen Trainingsdaten. So gehen die Sprachmodelle im Englischen z.B. davon aus, dass ein „doctor“ (Arzt) männlich ist und eine „nurse“ (Krankenschwester) weiblich, obwohl beide Berufsbezeichnungen im Englischen geschlechtsneutral sind. Auch kulturell sind die NLP-Modelle natürlich geprägt von den verwendeten Daten. Es wird intensiv daran geforscht, diese Verzerrungen auszuräumen.

Sprachliche Ironie ist ein komplexes Thema und meines Wissens noch nicht gelöst. Doch zu ähnlich komplexen Themen gab es gerade kürzlich große Entwicklungen: Das von Google trainierte Modell „PaLM“ kann Witze erklären und komplexe Denksportaufgaben lösen. Ich bin überzeugt davon, dass man einer KI das Erkennen von Ironie über Beispiele und gewisse vorgegebene Regeln in einem gewissen Umfang beibringen kann.

Gibt es weitere Herausforderungen?

Zu den größeren Herausforderungen würde ich das Beseitigen von Verzerrungen zählen, wie eben am Beispiel von „doctor/nurse“ beschrieben. Damit einher geht die Notwendigkeit, ethische Grundsätze zu wahren sowie Ergebnisse für die Nutzer*innen nachvollziehbar zu machen, ein großes Forschungsgebiet unter dem Namen „Explainable AI“. Dieser Terminus bezeichnet die Anforderung, dass ein KI-System seine Entscheidungen so trifft und erläutert (explain), dass ein Mensch in der Lage ist, diese Entscheidung nachzuvollziehen. Dies dient dazu, dass Menschen Vertrauen in die Entscheidungsfähigkeit von KI-Systemen entwickeln und den Einsatz dieser Technologie akzeptieren.

Bislang gelingt es Computerprogrammen nur in begrenztem Umfang, selbst sinnvolle Texte zu erstellen. Könnte sich das ändern? Werden wir beispielsweise zukünftig von Suchmaschinen nicht nur die Internetseiten zu einem

Begriff angezeigt bekommen, sondern gleich ein Dossier erhalten, das uns das Wichtigste aus diesen Seiten zusammenfasst?

Dieses Szenario ist durchaus möglich. Heutige KI-basierte Sprachmodelle sind so gut vortrainiert, dass sie anhand weniger Beispiele oder teilweise sogar nur auf eine Anweisung hin die gewünschten Texte generieren. Dazu gehört auch die Zusammenfassung von Texten. Wir sind heute noch etwas begrenzt im Umfang – aktuell sind Eingaben und Ergebnisse auf etwa 1500 Wörter limitiert, doch auch hier hat sich die Länge der handhabbaren Texte in den letzten 1-2 Jahren vervierfacht.

Google, Bing und andere Suchmaschinen heben auf den gefundenen Webseiten bereits die zu unserer Suche relevanten Passagen hervor. Diese Passagen können heute bereits sinnvoll zusammengefasst werden. Vermutlich ist es urheberrechtlich noch zu klären, wie die Quellen gekennzeichnet werden, um ein Dossier, wie von Ihnen gewünscht, zu liefern.

Wenn wir über Suchmaschinen etwas recherchieren oder wir etwas ins Netz stellen – z. B. in den sozialen Medien – geben wir viel über unsere Interessen, Meinungen usw. preis. Wird auch hier NLP eingesetzt, um unsere Interessen zu entschlüsseln und dies für Marketingzwecke zu nutzen?

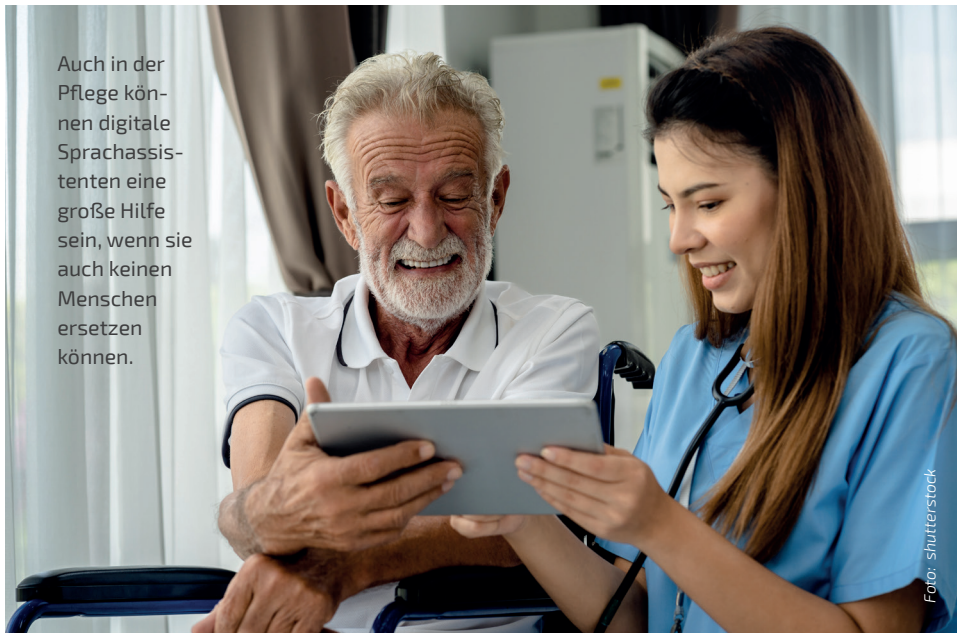
Heutige Empfehlungssysteme basieren größtenteils auf NLP-Algorithmen. Neben Nutzerinteraktionen und Nutzungsmustern wird auch auf die inhaltliche Ähnlichkeit von Texten geachtet, um den Nutzer*innen Inhalte anzubieten, die möglichst ihren Interessen entsprechen. Innerhalb uns bekannter Plattformen wie Amazon oder sozialer Netzwerke wie Facebook und YouTube ist uns dies auch überwiegend bewusst. Darüber hinaus hinterlassen wir Spuren über unseren Besuch von Webseiten, welche Tracker enthalten, die unsere Interessen über diverse Webseiten und Plattformen hinweg sammeln. Hierzu gibt es diverse Initiativen, z.B. auf EU-Ebene, um den Schutz unserer Daten zu gewährleisten und transparenter zu gestalten. Es ist gut vorstellbar, dass der Einsatz derartiger Technologien und die damit verbundenen Praktiken Gegenstand zukünftiger Gesetzgebung und Regulierung werden, z. B. durch eine Erweiterung des Datenschutzes.

Siri, Alexa usw. sind Programme, die unsere Sprache verstehen und uns unterstützen sollten. Nutzen Sie selbst solche Assistenten? Oder birgt es zu viele Gefahren für den Datenschutz?

Ich bin persönlich einfach sehr vorsichtig, was die Preisgabe meiner persönlichen Daten betrifft. Leider bin ich da aber auch nicht immer konsequent, ich nutze beispielsweise Facebook. Beim Surfen im Internet lehne ich die

Einwilligung zur Erfassung nicht notwendiger Daten stets ab, verklücke mich jedoch hin und wieder. Siri, Cortana und Alexa nutze ich nicht. In gewissen Situationen sind sie jedoch sehr hilfreich – z.B. beim Autofahren oder für ältere Personen, die darüber einfach in Kontakt mit Bezugspersonen treten könnten, sofern sie die neue Technologie annehmen und frühzeitig herangeführt werden. Gefahren beim Einsatz von Siri, Alexa und Co. sind auch leicht vorstellbar, wenn wir an Situationen zu Hause denken, wo sich Kinder aufhalten, die mit diesen Systemen interagieren und Kommandos wissentlich oder auch unwissentlich geben.

Der Personalmangel ist in vielen europäischen Ländern bereits spürbar. Werden sprechende und verstehende Maschinen hier helfen – auch in sensiblen Bereichen wie der Pflege?



Auch in der Pflege können digitale Sprachassistenten eine große Hilfe sein, wenn sie auch keinen Menschen ersetzen können.

Foto: shutterstock

In gewissem Umfang haben wir das heute schon. Viele Banken, Versicherungen, Taxizentralen und andere Unternehmen betreiben Call Center so, dass Sie die Unternehmen gar nicht mehr direkt anrufen können. Stattdessen filtern automatisierte Sprachsysteme die eingehenden Anrufe und beantworten häufig auftretende Fragen, ohne dass ein Mensch involviert

ist. Dieser Trend wird sich sicher fortsetzen. Maschinen werden uns helfen können, wenn sie auch den menschlichen Kontakt nicht ersetzen. So können Sprachassistenten wie Alexa Personen dabei unterstützen, über natürliche Sprache Informationen einzuholen, den telefonischen oder Video-Kontakt zu Angehörigen aufzunehmen oder in Notfällen, beispielsweise bei Stürzen, über Sprachbefehl den Notdienst zu kontaktieren. Andere Anwendungsszenarien gibt es in der Krankenpflege, wo Patient*innen ihre Bedürfnisse über ein sprachgesteuertes System an das Pflegepersonal kommunizieren und dadurch das traditionelle Klingeln ersetzt wird. Diese Funktion kann auch zur verbesserten Planung des Pflegepersonals und zu mehr persönlicher Ansprache der Patient*innen führen.

Wie weit könnte die Technik aus Ihrer Sicht in zehn Jahren sein?

Aktuell schreitet die Entwicklung im Bereich NLP in Siebenmeilenstiefeln voran. Der ungeheure Zuwachs an Rechenressourcen in den vergangenen 10-15 Jahren hat zur Entwicklung von Neuronalen Netzen geführt, die in Verbindung mit Deep Learning Dinge ermöglichen, die vor wenigen Jahren selbst bei enormem Programmieraufwand und hohen Computer-Ressourcen undenkbar gewesen wären. Forschung und Industrie sind dabei, die aktuellen neuen Funktionalitäten – wie das Vervollständigen von Texten, das Generieren von Texten anhand weniger Beispiele, das automatisierte Zusammenfassen – in Anwendungen umzusetzen. Ich gehe davon aus, dass uns in zehn Jahren die künstliche Sprach-Intelligenz als ständiger Begleiter durch den Arbeitsalltag führen, uns unterstützen und viele standardisierbare Aufgaben abnehmen wird. Doch wir werden weiterhin die menschliche Intelligenz benötigen, die dann vor allem dort eingesetzt wird, wo Maschinen nicht mithalten können: Das ist die konzeptionelle Steuerung von kreativem Inhalt sowie das Bewerten der Ergebnisse vor dem Hintergrund komplexer, auch „softer“ Entscheidungsfaktoren: Wie passt das Ergebnis zur strategischen Ausrichtung des Unternehmens, wie wirkt es sich auf Mitarbeiter*innen, Kunden*innen und die öffentliche Wahrnehmung aus?

Auf jeden Fall blicke ich den weiteren Entwicklungen mit Spannung entgegen. Letztendlich ist der Mensch ein soziales Wesen. In der Disziplin des Sozialen und allem, was davon abhängt, wird der Mensch sich trotz der sich stetig weiterentwickelnder Technik noch lange Zeit behaupten können, sofern er oder sie nicht vorher den Planeten für den Menschen unbewohnbar gemacht hat.

Dr.-Ing. Veronika Gamper

gebürtig aus Meran, ist Gründerin und Geschäftsführerin der WeDaVinci GmbH, einer NLP-basierten Kollaborationsplattform für Innovationsprojekte. Sie studierte Informatik an der TU Wien und promovierte bei Prof. Dr.-Ing. Klaus Diepold an der TU München, mit Studien- und Forschungsaufenthalten am IIT Chicago und an der UC Berkeley. Nach ihrem Studium arbeitete sie bei der Boston Consulting Group, am Center for Digital Technology and Management und für einen führenden europäischen Medienkonzern.



Prof. Dr.-Ing. Klaus Diepold

arbeitet am Lehrstuhl für Datenverarbeitung der Fakultät für Elektrotechnik und Informationstechnik an der Technischen Universität München.

Das Interview erschien in der Sprach_info
der Sprachstelle im Südtiroler Kulturinstitut, Ausgabe Herbst 2022.