

Spracherkennung im professionellen Einsatz.

New Work: Wie KI und
Cloud-Anwendungen die
Arbeitswelt revolutionieren



Spracherkennung im professionellen Einsatz.

New Work: Wie KI und Cloud-Anwendungen die Arbeitswelt revolutionieren.

Angesichts rasanter Fortschritte bei der Spracherkennung und -analyse erobert sich die Spracheingabe immer mehr Einsatzbereiche. Sie ist nicht nur wesentlich schneller als die Texteingabe per Tastatur, sondern führt auch zu mehr Flexibilität und Mitarbeiterzufriedenheit. Bei der Wahl einer Spracherkennungslösung sollten Unternehmen neben Leistungsfähigkeit und Genauigkeit aber auch Sicherheit und Datenschutz berücksichtigen.

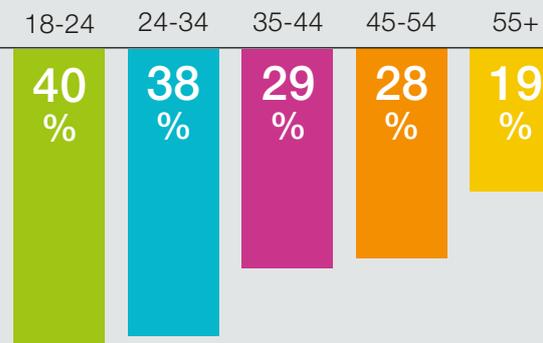
Die Digitalisierung verändert die Art der Arbeit und die Wahl der Arbeitsmittel massiv. Mobile Endgeräte, leistungsfähige Netze und Cloud-basierte Infrastrukturen ermöglichen es Mitarbeitern, von überall auf Unternehmensressourcen zuzugreifen – zumindest theoretisch, denn die Realität sieht häufig noch anders aus. Nur 16 Prozent der deutschen Arbeitnehmer können flexibel von zu Hause oder von unterwegs aus arbeiten, ergab eine aktuelle Studie der Initiative D21¹. Die Weigerung oder das Unvermögen, seinen Mitarbeitern flexibles Arbeiten zu ermöglichen, kann sich für Firmen als handfester Wettbewerbsnachteil erweisen.

Unternehmen, die den Ansprüchen heutiger Arbeitnehmer gerecht werden (Stichwort: „New Work“), haben ein höheres Wachstum und sind laut einer von VMware in Auftrag gegebenen Studie² häufiger Marktführer als Wettbewerber mit einer eher traditionellen Workplace-Ausstattung. Vor allem junge Mitarbeiter drängen auf Veränderungen, wie eine Umfrage³ des Marktforschungsunternehmens YouGov im Auftrag der Intranet-Agentur Hirschtec zeigt. Während 40 Prozent der 18- bis 24-Jährigen im Bereich der digitalen Arbeitsplätze Modernisierungsbedarf sehen, sind dies bei den über 55-Jährigen nur 19 Prozent.

Aber auch was die Büroumgebung angeht, müssen sich Chefs von lieb gewordenen Ansichten über Arbeitsorganisation und Arbeitsmittel verabschieden. Das Großraumbüro, das als Hort für Kreativität und Produktivität galt, gerät zunehmend in die Kritik. Laut den empirischen Untersuchungen von Ethan S. Bernstein und Stephen Turban hemmt es die Interaktion und Kommunikation, statt sie zu fördern⁴.

Modernisierungsbedarf des digitalen Arbeitsplatzes.

Anteil der Personen pro Altersgruppe, die den Modernisierungsbedarf für ihren digitalen Arbeitsplatz als hoch oder sehr hoch einschätzen.



Vor allem junge Mitarbeiter sehen beim digitalen Arbeitsplatz Modernisierungsbedarf.

Quelle: YouGov/Hirschtec
Stichprobe: 1.000 Berufstätige in Deutschland

1 https://initiated21.de/app/uploads/2019/01/d21_index2018_2019.pdf

2 <https://www.vmware.com/radius/digital-employee-experience-infographic/>

3 <https://hirschtec.eu/studie-digitaler-arbeitsplatz-ganzheitlich-2019/>

4 The impact of the 'open' workspace on human collaboration, <https://doi.org/10.1098/rstb.2017.0239>

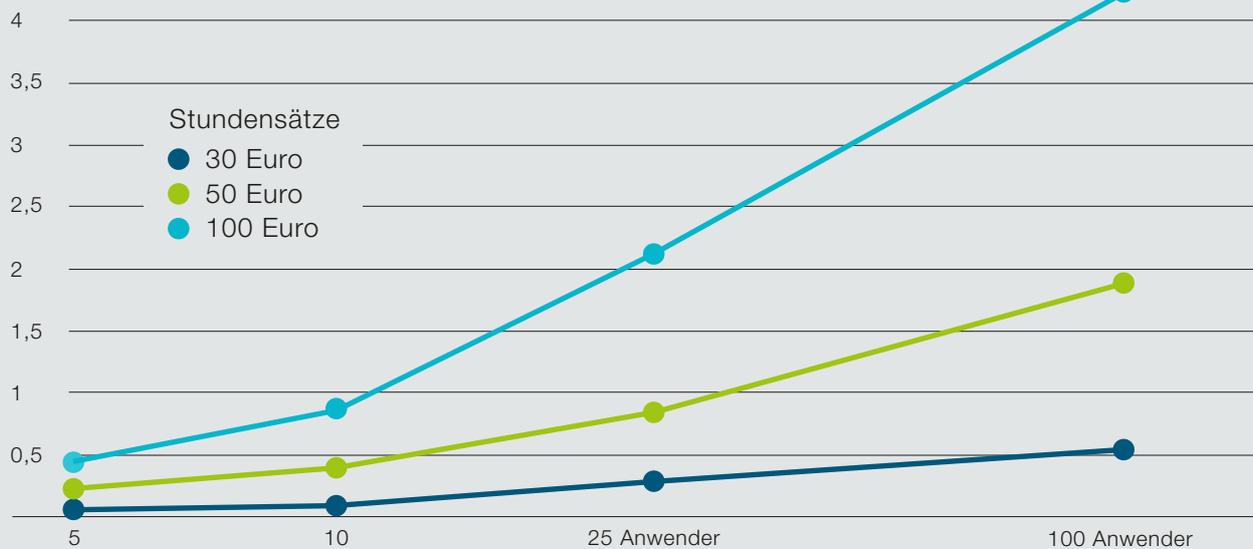
Selbst traditionelle Arbeitsmittel wie die Tastatur könnten vom Aussterben bedroht sein. Fortschritte in der Künstlichen Intelligenz (KI) machen ihren Einsatz zunehmend überflüssig. Aktuelle Spracherkennungssysteme können gesprochenen Text nahezu fehlerfrei erkennen und umsetzen. Wie eine Studie an der Stanford University ergab, ist die Texteingabe per Sprache dreimal schneller als das Tippen⁵.

Durch Spracheingabe lassen sich also textbasierte Aufgaben wie die Bürokommunikation, die Dokumentation oder die Erstellung von Schriftsätzen massiv beschleunigen. Die Investition in eine Spracherkennungslösung rechnet sich daher schnell. Wenn in einem Unternehmen mit 50 Mitarbeitern durch effizientere Texterfassung nur zwei Stunden Arbeitszeit pro Woche und Mitarbeiter gespart werden könnten, beläuft sich das Einsparpotenzial bei einem angenommenen Stundensatz von 50 Euro auf 1,2 Millionen Euro pro Jahr.

Die großen Fortschritte in den Bereichen Maschinelles Lernen (ML) und Künstliche Intelligenz (KI) sind wesentliche Ursachen für die hohe Leistungsfähigkeit aktueller Spracherkennungslösungen. Algorithmen können heute selbstständig aus Erfahrung lernen oder gezielt trainiert werden. So passen sie sich durch die Interaktion mit einem Sprecher immer besser an dessen Spracheigenheiten, etwa einen Dialekt oder Akzent, an und erhöhen kontinuierlich die Treffergenauigkeit. Fachspezifische Vokabulare lassen sich im Vorfeld importieren, explizit trainieren oder im Diktierprozess laufend verbessern. KI-basierte Komponenten wie Natural Language Processing (NLP) und Natural Language Understanding (NLU) erlauben es modernen Spracherkennungsprogrammen, den Inhalt und den Kontext der Spracheingabe zu identifizieren. Zum Beispiel können sie auf Anweisungen wie „nächste Seite“ oder „neuer Abschnitt“ reagieren, ohne dass der Sprecher für Steuerungsbefehle starre Wortkombinationen benutzen muss.

Einsparpotenzial durch Nutzung von Spracherkennung

Einsparung in Mio. Euro pro Jahr bei 2 Stunden pro Woche



Wer durch eine Spracherkennungslösung zwei Stunden pro Woche an Zeit gewinnt, kann einiges an Kosten sparen. (Quelle: Nuance)

⁵ https://hci.stanford.edu/research/speech/paper/speech_paper.pdf

Ein weiterer Faktor, der die Verbreitung und die Einsatzmöglichkeiten der sprachbasierten Texterstellung fördert, ist das Cloud Computing. Eine zentral verwaltete und flexibel nutzbare Infrastruktur erleichtert den Zugang erheblich und lässt sich mit wenig Aufwand beliebig skalieren. Statt einer aufwendigen Installation auf den Clients lässt sich Spracherkennung aus der Cloud schnell und einfach mit wenigen Klicks nutzen. Für die Unternehmen, für die Cloud Computing auf externen Servern nicht infrage kommt, bietet Nuance als einer von wenigen Herstellern eine Private-Cloud-Lösung an, die Spracherkennung in eigenen Rechenzentren möglich macht. Der Provider bzw. die interne IT-Administration stellt dabei nicht nur die notwendigen Server- und Speicherkomponenten zur Verfügung, sondern spielt auch alle notwendigen Aktualisierungen und Sicherheits-Updates ein, ohne dass sich der Anwender darum kümmern muss.

Es ist daher nicht verwunderlich, dass sich immer mehr Unternehmen für die sprachbasierte Textfassung interessieren und damit dem Konzept „New Work“ als neue Arbeitsweise der heutigen Gesellschaft folgen wollen. Dies zeigt eine Studie, für die vom Marktforschungsunternehmen Censurwide im Auftrag von Nuance 400 IT-Entscheider im Finanz- und Rechtsbereich befragt wurden. Fast 30 Prozent der Umfrageteilnehmer halten die Tastatur für nicht geeignet, um die Transformation in das digitale Geschäftszeitalter zu unterstützen. Fast alle Beteiligten waren der Ansicht, dass sie von mehr Freiheit am Arbeitsplatz gesundheitlich profitieren würden und dass Spracherkennung ein zukunftsfähiges Werkzeug für ihre tägliche Arbeit darstellt. Nahezu 90 Prozent waren außerdem davon überzeugt, dass Spracherkennungs-Tools ihre Abhängigkeit von Transkriptionsbüros und anderen Text-Dienstleistern senken würden. Mehr als 80 Prozent äußerten die Ansicht, dass sie mithilfe von Spracherkennung produktiver arbeiten könnten.

Vorsicht Falle – Siri, Alexa & Co. im professionellen Einsatz.

Aktuelle Betriebssysteme für mobile Endgeräte und PCs wie Android, iOS, Windows 10 oder Mac OS X bringen ihre eigenen Sprachassistenten schon mit. Auch Amazons Spracherkennungssystem „Alexa“ erfreut sich bei Konsumenten im privaten Umfeld großer Beliebtheit. Was liegt also näher, als diese vorhandenen Tools auch professionell einzusetzen?

Ein genauer Blick in die Nutzungsbedingungen zeigt jedoch, dass hier unabsehbare Gefahren lauern. Die erfassten Sprachdaten werden in aller Regel auf Servern der Anbieter verarbeitet, die in den USA stehen. Werden personenbezogene Daten erfasst und übermittelt, ohne dass ein Einverständnis der Betroffenen und ein Vertrag mit dem Anbieter über eine Auftragsverarbeitung der Daten vorliegt, kann es sich um einen Verstoß gegen die Europäische Datenschutzgrundverordnung (DSGVO) handeln, der in schweren Fällen mit Bußgeldern von bis zu vier Prozent des weltweiten Brutto-Jahresumsatzes eines Unternehmens geahndet werden kann⁶. So können die Einsparungen durch die Nutzung der kostenlosen Tools schnell zunichte gemacht werden.

Besonderen Risiken sind sogenannte Berufsgeheimnisträger ausgesetzt, zu denen Ärzte, Rechtsanwälte und Versicherungsunternehmen gehören. Wer als Mitglied einer solchen Gruppe fremde Informationen offenbart, kann laut § 203, Absatz 1 des Strafgesetzbuches (StGB) mit einer Freiheitsstrafe bis zu einem Jahr bestraft werden. Rechtsanwälte sind darüber hinaus durch § 43a der Bundesrechtsanwaltsverordnung (BRAO) zur Verschwiegenheit verpflichtet.

Aber selbst wenn keine rechtlichen Bedenken gegen die Nutzung von Spracherkennungssoftware aus dem Endkundenbereich sprechen, handeln sich Unternehmen beim Gebrauch solcher Assistenten zahlreiche Nachteile ein. Es gibt weder eine garantierte Verfügbarkeit noch einen professionellen Support für Business-Anwender. Auch eine Integration in bestehende Software-Landschaften oder Workflows ist nicht möglich. Eine besondere Schwachstelle der Consumer-Apps ist ihre mangelnde Anpassbarkeit an fachspezifische Anforderungen. In den meisten Branchen gibt es ein spezifisches Vokabular an Fachausdrücken und Abkürzungen, mit denen Siri & Co. nichts anfangen können. Professionelle Spracherkennungssoftware lässt sich dagegen um bestehende fachspezifische Sprachpakete erweitern oder kann diese sehr schnell lernen.

⁶ <https://dsgvo-gesetz.de/themen/bussgelder-strafen/>

Wie Finanzdienstleister von Spracherkennung profitieren.



Mit der Finanzmarktrichtlinie „MiFID II“ (Markets in Financial Instruments Directive)⁷ hat die Europäische Union 2018 die Dokumentationspflichten für Banken und Versicherungen deutlich ausgeweitet. Sämtliche Kommunikation mit Kunden, die im Zusammenhang mit Aufträgen, etwa dem Kauf von Wertpapieren steht, muss dokumentiert und fünf Jahre aufbewahrt werden. Kunden können verlangen, dass ihnen die Aufzeichnungen oder Kopien davon zur Verfügung gestellt werden. Bei Beschwerden oder behördlichen Untersuchungen stellen diese Dokumente darüber hinaus wichtige Beweismittel dar. Für die betroffenen Unternehmen bedeutet dies einen erheblichen persönlichen Mehraufwand. Aufgezeichnete Telefonate müssen transkribiert, handschriftliche Gesprächsprotokolle abgetippt werden. Hier kann Spracherkennung einen erheblichen Beitrag zu mehr Effizienz und Produktivität, aber auch zu mehr Rechtssicherheit leisten.

Wie Finanzdienstleister von professioneller Spracherkennung profitieren, zeigt das Beispiel der NATIONAL Bank in Essen. Viele der rund 600 Mitarbeiter der bereits 1921 gegründeten Traditionsbank verbringen einen großen Teil ihres Arbeitstages mit dem Schreiben von Kredit-Statements, Unternehmensanalysen, Besprechungsprotokollen und anderen umfangreichen Textdokumenten. Durch den Einsatz der Spracherkennungssoftware Dragon

Professional Anywhere (DPA) konnte das Unternehmen die Effizienz der Mitarbeiter deutlich steigern. Mithilfe der Software lassen sich 160 Wörter pro Minute transkribieren – eine Steigerung gegenüber der Tastatureingabe um den Faktor 3. Die Fehlerrate ist sehr gering, rund 99 Prozent des gesprochenen Textes werden korrekt erkannt. Aber auch die Gesundheit der Mitarbeiter profitierte vom Einsatz der Spracherkennungslösung. Statt acht Stunden und länger Texte einzugeben, während sie verkrampft über die Tastatur gebeugt sind, können sie nun ihre Arbeit flexibel gestalten und beispielsweise auch im Stehen diktieren.

Die Vermögens- und Unternehmensberatung Courté J. & Co. KG profitiert ebenfalls vom Einsatz professioneller Spracherkennung. Bisher wurden Geschäftsbriefe und Schriftsätze diktiert und dann von externen Schreibbüros oder dem internen Sekretariat transkribiert. Dieses Vorgehen war aufwendig und langsam. Oft vergingen Tage zwischen der Erstellung des Dokuments und seiner Fertigstellung. Mithilfe der Dragon-Spracherkennung kann Courté nun Schriftsätze direkt aus dem Diktat erstellen, was die Verarbeitung enorm beschleunigt hat.

⁷ https://www.bafin.de/DE/PublikationenDaten/Jahresbericht/Jahresbericht2017/Kapitel5/Kapitel5_1/Kapitel5_1_1/kapitel5_1_1_node.html

Spracherkennung in **Rechts-** **anwaltskanzleien** – damit **Klienten** **schneller zu ihrem Recht kommen.**



Die zunehmende Digitalisierung macht auch vor den Rechtsanwaltskanzleien nicht Halt. Wie eine repräsentative Studie des Anwalt-Netzwerks AdvoAssist in Zusammenarbeit mit dem Prozessfinanzierer Foris ergab⁸, sind 69 Prozent der Rechtsanwälte davon überzeugt, dass der Einsatz digitaler Technologien in den kommenden fünf Jahren ihre Kanzlei „sehr stark“ oder „stark“ beeinflussen wird. Klassische Anwaltstätigkeiten wie Mandantenberatung, Vertragserstellung und Dokumentation geraten zudem durch digitale Dienstleister, die sogenannten Legal-Techs, unter Druck. Dabei handelt es sich meistens um Internet-Portale.

Während hinter diesen Portalen oft noch menschliche Rechtsexperten stehen, die das Internet im Wesentlichen als neuen Vertriebsweg nutzen, droht zunehmend die maschinelle Konkurrenz. Systeme, die auf Künstlicher Intelligenz (KI) beruhen, zeigen auch im Umfeld der Rechtspflege erstaunliche Leistungen. In manchen Bereichen ist die KI menschlichen Anwälten bereits deutlich überlegen, wie ein Beispiel aus Amerika zeigt. Um die Leistungsfähigkeit seiner KI-basierten Analysen zu testen, ließ das Portal LawGeex⁹ seinen Algorithmus gegen 20 Anwälte antreten. Sie sollten fünf Geheimhaltungsvereinbarungen (Non-Disclosure Agreements, NDA) auf rechtliche Risiken untersuchen. Die Anwälte benötigten im Durchschnitt 92 Minuten für diese Aufgabe und fanden 85 Prozent der Probleme. Die KI erledigte den Job dagegen in nur 26 Sekunden, bei einer Genauigkeit von 94 Prozent.

⁸ <https://www.advo-assist.de/rechtsmarktforschung/>

⁹ <https://www.lawgeex.com/resources/aivslawyer/>

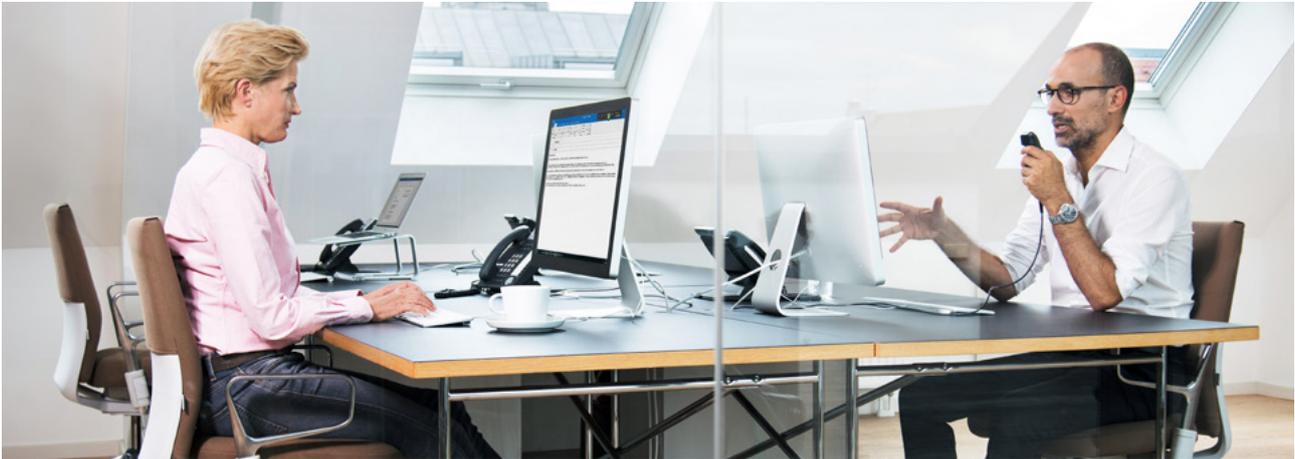


Neben dem zunehmenden Wettbewerb durch die LegalTechs plagen die Rechtsanwälte auch Nachwuchssorgen beim Personal. Der Beruf der Rechtsanwaltsfachangestellten drohe auszusterben, warnte Rechtsanwalt Christoph Möllers, Mitglied im Berufsbildungsausschuss der Bundesrechtsanwaltskammer (BRAK) auf der Veranstaltung „Zukunft der Anwaltschaft“¹⁰. Zu unattraktiv seien die Berufsaussichten, zu groß der Wettbewerb durch Justiz und die Rechtsabteilungen der Unternehmen. In der Folge müssen Rechtsanwälte Schriftsätze immer häufiger durch externe Schreibbüros erstellen lassen. Neben Problemen der Rechtssicherheit und Vertraulichkeit sorgt dieser umständliche Prozess auch für lange Wartezeiten und hohe Kosten – angesichts der schnellen und kosteneffizienten digitalen Konkurrenz sind das erhebliche Wettbewerbsnachteile.

Spracherkennungssoftware ist daher ein ideales Mittel für Rechtsanwälte, um schneller und effizienter zu werden und so im Wettbewerb mit der digitalen Konkurrenz bestehen zu können. So setzt beispielsweise die Rechtsanwaltskanzlei Heussen, eine der großen Wirtschaftskanzleien in Deutschland, auf die Spracherkennung Dragon von Nuance. Zuvor nutzten die Anwälte analoge oder digitale Diktiergeräte, um ihre Schriftsätze zu erfassen, die dann von einer Schreibkraft transkribiert wurden. Dieser Vorgang dauerte mehrere Stunden. In der Zwischenzeit musste der Anwalt den Sachverhalt im Kopf behalten, um den fertigen Schriftsatz auf Korrektheit überprüfen zu können – bei drei oder mehr parallelen Vorgängen eine enorme kognitive Belastung. Heute kann der Anwalt seinen Text direkt diktieren und sofort überprüfen. Danach kann er den Vorgang abschließen und sich sofort voll auf die nächste Aufgabe konzentrieren. Durch die enorme Einsparung an Zeit und Kosten können sich die Anwälte zudem viel intensiver mit einem Rechtsfall befassen und schneller reagieren als bisher.

¹⁰ <https://www.brak.de/thesen-zur-zukunft-der-anwaltschaft/>

Die Digitalisierung in Behörden – ein Weg mit Hindernissen.



In der öffentlichen Verwaltung tut man sich mit der digitalen Transformation derzeit noch besonders schwer. Wie der Tagesspiegel unter Berufung auf eine Studie des Beratungsunternehmens Ernst & Young berichtete¹¹, ist nur bei jeder fünften deutschen Behörde eine Online-Terminvereinbarung möglich. Gleichzeitig können in nur sechs Prozent der Fälle Gebühren online bezahlt werden und nur drei Prozent nutzen die Funktionen des E-Personalausweises. Ein wesentliches Hindernis zu mehr Digitalisierung ist die Personalstruktur. Es fehle an einer konsequenten Ausrichtung der Personalstrategie an den Erfordernissen der Digitalisierung, so die befragten Verantwortlichen.

Aber nicht nur die Personalstrategie, auch die Personalausstattung insgesamt verhindert eine schnellere Transformation. Vor allem in Großstädten fehlt es an Verwaltungspersonal, und die vorhandenen Fachkräfte sind den ganzen Arbeitstag derart in operative Aufgaben wie die Erfassung und Dokumentation von Vorgängen eingebunden, dass für eine strategische Weiterentwicklung weder Kraft noch Zeit bleibt.

Der Einsatz intelligenter Spracherkennungslösungen kann hier erheblich dazu beitragen, Personal zu entlasten und die nötigen Freiräume für eine Weiterentwicklung zu schaffen. Das beginnt schon bei der Anrufannahme im Contact Center. Mithilfe von Spracherkennung und einer automatisierten Kommunikationslenkung lassen sich nach Untersuchungen¹² von Nuance 54 Prozent der Anrufe direkt erledigen,

ohne dass ein menschlicher Eingriff nötig wäre. Die durchschnittliche Bearbeitungsdauer sinkt um 42 Prozent. Darüber hinaus können Routineaufgaben wie die Erstellung von Schriftsätzen und Bescheiden oder die Kommunikation per E-Mail wesentlich beschleunigt werden.

So nutzt beispielsweise das Landesamt für Gesundheit und Soziales in Berlin die Spracherkennungssoftware Dragon. Mitarbeiter im Außendienst können ihre Berichte direkt vor Ort oder auf der Fahrt ins Büro in ein mobiles Diktiergerät sprechen. Am Arbeitsplatz angekommen, liegen die fertigen Informationen bereits vor. Der Mitarbeiter kann sich neuen Aufgaben zuwenden, statt erst einmal ein Protokoll tippen zu müssen. Darüber hinaus stellt ein zeitnah diktiert Text den Sachverhalt in der Regel deutlich genauer und richtiger dar als ein Stunden später aus dem Gedächtnis verfasstes Dokument.

Behörden kommt auch eine wesentliche Aufgabe zu, wenn es um die Integration von Menschen mit Behinderungen in den Arbeitsmarkt geht. Wenn körperliche Einschränkungen die Arbeit mit Tastatur und Bildschirm erschweren oder gar unmöglich machen, kann Spracherkennungssoftware eine wesentliche Hilfe darstellen, um solchen Mitarbeitern den Arbeitsalltag zu erleichtern oder ihnen die Teilhabe am Berufsleben überhaupt erst zu ermöglichen.

¹¹ <https://www.tagesspiegel.de/wirtschaft/digitalisierung-verwaltung-weiterhin-wenig-digitalisiert/23770040.html>

¹² <https://www.nuance.com/de-de/industries/government.html>

Fazit.

Die digitale Transformation stellt Unternehmen jeder Größe und Branche, aber auch die öffentliche Verwaltung vor große Herausforderungen. Eine der größten liegt darin, im täglichen Geschäft genügend Zeit und Freiräume für eine strategische Ausrichtung und Weiterentwicklung zu schaffen. Spracherkennung kann hier einen wesentlichen Beitrag leisten, denn sie beschleunigt das Erfassen und Verarbeiten von Dokumenten und Nachrichten, sorgt für ein ergonomischeres und damit ermüdungsfreies Arbeiten und vereinfacht textbasierte Workflows erheblich. Mitarbeitern bleibt damit mehr Zeit und Energie, sich weiterzubilden, neue Geschäftsfelder zu erschließen und die Digitalisierung voranzutreiben.

Eine Spracherkennung ist auch ein Schritt in Richtung Arbeitswelt der Zukunft („New Work“) mit mehr Freiheit und Selbstbestimmung der Arbeitnehmer. Unternehmen und Behörden sollten bei der Wahl einer Lösung aber auch Datenschutz und Sicherheit nicht außer Acht lassen. Nur so lässt sich ein böses Erwachen verhindern und ein langfristiger und zukunftsfähiger Einsatz gewährleisten.

Worauf Sie bei der Wahl einer **Spracherkennungslösung** achten sollten.

Folgende Kriterien sind für die Kaufentscheidung besonders wichtig:

Schnelligkeit:

Ihr gesprochener Text sollte nahezu in Echtzeit transkribiert werden. Eine Spracherkennung, die nach jedem Satz erst ewig „nachdenkt“, bringt Sie aus dem Konzept und verzögert die Texterstellung eher, als dass sie diese beschleunigt.

Genauigkeit:

Eine Spracherkennung hilft wenig, wenn das transkribierte Dokument vor Fehlern strotzt und aufwendig nachbearbeitet werden muss. Achten Sie daher auf eine nachweislich hohe Erkennungsrate. Aktuelle marktführende Programme können bis zu 99 Prozent des diktierten Textes korrekt erkennen.

Anpassungsfähigkeit:

Die Software sollte sich schnell und einfach an die Gegebenheiten und Anforderungen Ihrer Umgebung anpassen lassen. Das reicht vom Training auf individuelle Sprecherprofile über den in Ihrer Branche gängigen Fachwortschatz bis hin zur Integration in bestehende Workflows, etwa im Dokumenten- oder Genehmigungsmanagement.

Flexible Bereitstellung:

Je nach Anforderung und Budget bietet sich eine Bereitstellung über eine Public oder Private Cloud an. In manchen Fällen kann aber auch heute noch die direkte Installation einer lokalen Desktop-Lösung sinnvoll sein. Der Hersteller sollte Ihnen daher alle Optionen bieten und Ihre Entscheidungsfreiheit nicht durch eine zu kleine Auswahl an Bereitstellungs- und Lizenzmodellen einschränken.

Einfache Installation und Wartung:

Eine Spracherkennungssoftware nützt wenig, wenn der Zeit- und Produktivitätsvorteil durch eine komplexe, aufwendige und personalintensive Administration zunichte gemacht wird. Achten Sie daher darauf, dass die Lösung Ihrer Wahl möglichst einfach zu konfigurieren und zu installieren ist. Zusätzlich senken automatische Updates den Verwaltungsaufwand.

Skalierbarkeit:

Nur wenn sich die Spracherkennungslösung schnell und flexibel an einen erhöhten Bedarf anpassen lässt, ohne dass es zu Leistungseinbußen kommt, sind Sie auch für die Zukunft gerüstet.

Sicherheit:

Beim Diktat werden häufig sensible Informationen weitergegeben und verarbeitet. Achten Sie daher darauf, dass die Kommunikation zwischen den Komponenten der Lösung durch eine starke 256-Bit-Verschlüsselung geschützt ist. Auch bei der Speicherung der Daten sollte eine starke Verschlüsselung zum Einsatz kommen.

Datenschutz:

Bei der Verarbeitung personenbezogener Daten sind die Anforderungen der Europäischen Datenschutzgrundverordnung (DSGVO) zu beachten. Achten Sie bei der Wahl einer Cloud-Lösung darauf, dass sich die Rechenzentren des Providers im Europäischen Rechtsraum befinden und nach gängigen Standards wie der ISO 27001 zertifiziert sind. Auch der Anforderungskatalog Cloud Computing (C5)¹³ des Bundesamtes für Sicherheit in der Informationstechnik (BSI) liefert eine valide Bewertungs- und Entscheidungsgrundlage.

Integrationsfähigkeit:

Die Lösung Ihrer Wahl sollte sich sowohl lokal installieren als auch nahtlos in bestehende Thin-Client-Infrastrukturen, virtualisierte Umgebungen und Workflows einbinden lassen.

¹³ https://www.bsi.bund.de/DE/Themen/DigitaleGesellschaft/CloudComputing/Anforderungskatalog/Anforderungskatalog_node.html

Checkliste für die Auswahl einer Spracherkennung.

- Welche Bereitstellungsmodelle gibt es?
- Muss ich die Software selbst installieren und managen oder kann ich sie auch als Service beziehen?
- Wie viele Nutzer können parallel mit der Software arbeiten? Welche Skalierungsoptionen gibt es?
- Wie ist die Software aufgebaut? Muss sie auf jedem Client installiert werden oder gibt es eine zentrale Serverkomponente?
- Lässt sich die Lösung ohne Cloud auf einem lokalen Server installieren?
- Wie lange dauert die Installation? Wie aufwendig ist die Wartung?
- Wie genau ist die Spracherkennung? Kann sich die Software an einzelne Sprecher anpassen?
- Wie schnell kann die Software fachspezifische Vokabulare lernen?
- Muss die Software trainiert werden? Falls ja, wie hoch ist der Aufwand?
- Lässt sich die Software in bestehende Workflows integrieren? Wie aufwendig ist die Integration?
- Wie sicher ist die Software? Sind Übertragung und Speicherung verschlüsselt?
- Werden alle Belange des Datenschutzes eingehalten? Welche Zertifizierungen und Belege gibt es dafür?
- Falls die Software als Service (SaaS) aus einer Public Cloud bereitgestellt wird: Welche Zertifizierungen kann der Cloud-Provider nachweisen? Werden die Daten beim Transport und bei der Speicherung verschlüsselt? Befinden sich die Rechenzentren im Europäischen Rechtsraum?

Glossar.

Cloud Computing

Nach der Definition¹⁴ des National Institute of Standards and Technology (NIST) zeichnet sich Cloud Computing durch folgende fünf Eigenschaften aus:

On Demand Self Service: Anwender können jederzeit direkt auf Ressourcen zugreifen.

Broad Network Access: Der Zugriff erfolgt über eine Netzwerk-Infrastruktur. Dabei kommen standardisierte Protokolle zum Einsatz, die einen Zugriff über alle gängigen Client-Plattformen (PC, Smartphone, Thin-Clients) ermöglichen.

Resource Pooling: Ressourcen werden geteilt und stehen Anwendern je nach Anforderungen dynamisch zur Verfügung.

Rapid Elasticity: Die Zuordnung der Ressourcen erfolgt schnell und flexibel. Anwendern steht jederzeit nahezu beliebig viel Leistung zur Verfügung. Eine Skalierung ist jederzeit möglich. Skalierungseffekte können effizient ausgeschöpft werden.

Measured Service: Der Cloud-Betreiber überwacht die Systeme automatisch und kontinuierlich und sichert so die Servicequalität.

Cloud-Bereitstellungsmodelle

Private Cloud/On Premise: Die Cloud-Infrastruktur steht bei diesem Bereitstellungsmodell einem Unternehmen oder einer Abteilung zur alleinigen Nutzung zur Verfügung. Sie kann im unternehmenseigenen Rechenzentrum lokalisiert sein, aber auch von einem Provider betrieben werden.

Public Cloud: Bereitstellungsmodell, bei dem die Cloud-Infrastruktur über das öffentliche Internet erreichbar ist und von jedem genutzt werden kann. Nutzerkonten werden logisch separiert und durch Verschlüsselung auf dem Transportweg und bei der Speicherung zusätzlich gegen unberechtigte Zugriffe geschützt. Der Anwender hat keinen Einfluss darauf, in welchem Rechenzentrum der oft weltweit verteilten Infrastruktur seine Daten verarbeitet werden. Häufig lässt sich die Datenverarbeitung aber auf bestimmte Zonen, etwa EU, einschränken.

Cloud-Service-Modelle

Software-as-a-Service (SaaS): Service-Modell, bei dem der Anbieter aus seiner Cloud-Umgebung Applikationen bereitstellt. Der Anwender bezahlt in der Regel eine monatliche oder jährliche Gebühr und kann die Lösung dann direkt nutzen, ohne sich über Updates, Absicherung oder Skalierung Gedanken machen zu müssen.

Platform-as-a-Service (PaaS): Der Cloud-Betreiber stellt dabei eine Plattform mit Betriebssystem, Middleware und Entwicklungswerkzeugen zur Verfügung, auf der ein Kunde eigene Applikationen entwickeln und/oder betreiben kann.

Infrastructure-as-a-Service (IaaS): In diesem Fall greift der Kunde auf (meist virtualisierte) Hardware-Komponenten wie Server, Storage und Arbeitsspeicher zu und stellt sich daraus seine eigene Laufwerkumgebung zusammen. Installation und Verwaltung von Betriebssystem und Applikationen liegen in der Verantwortung des Anwenders.

Künstliche Intelligenz (KI)

Unter KI versteht man informationsverarbeitende Systeme, die den Bedeutungsinhalt von Sprache, Texten oder Bildern erfassen und daraus Schlüsse ziehen können. KI-Systeme kennen das Prinzip von Ursache und Wirkung, erkennen Konzepte sowie Kontext und können aufgrund ihrer Beobachtung Hypothesen für die Zukunft aufstellen. Sie sind zudem in der Lage, diese Hypothesen anhand neuer Daten zu überprüfen und aus den Ergebnissen zu lernen. KI-Systeme kommunizieren mit Menschen auf natürliche Art und Weise über Sprache oder Gesten und können angemessen auf die Handlung und Anweisungen ihres Gegenübers reagieren.

¹⁴ <https://csrc.nist.gov/publications/detail/sp/800-145/final>

Maschinelles Lernen (ML)

Machine-Learning-Algorithmen können große Datenmengen analysieren, Muster erkennen und daraus Regeln zur Klassifizierung ableiten. Im Wesentlichen lassen sich drei Arten des Maschinellen Lernens unterscheiden: Beim „Unsupervised Learning“ wertet der Algorithmus Daten ohne menschliches Zutun aus, während beim „Supervised Learning“ ein menschlicher „Trainer“ die Trainingsdaten vorgibt und die Ergebnisse bewertet. Das „Reinforcement Learning“ stellt einen Mittelweg dar. Der Algorithmus lernt selbständig, seine Ergebnisse werden aber an von Menschen vorgegebenen Zielwerten gemessen. Durch wiederholtes Ausprobieren und Verändern der Parameter versucht das System, diesen Zielwerten immer näher zu kommen und so seine Lernfunktion zu optimieren.

Natural Language Processing (NLP)

Eine Methode, bei der gesprochene Sprache oder Text auf Schlüsselwörter untersucht wird, um die Bedeutung zu erkennen und eine passende Antwort formulieren zu können. So könnte ein Chatbot beispielsweise auf das Wort „Wetter“ die Wettervorhersage für die nächsten Tage wiedergeben oder beim Schlüsselwort „kostet“ den Preis eines Produkts nennen.

Natural Language Understanding (NLU)

Eine semantische Analyse eines gesprochenen oder geschriebenen Textes, die es erlaubt, das Thema eines Dokuments zu erkennen und kontextgerecht zu reagieren. Damit lassen sich Stimmungen sowie Bewertungen erfassen und Absichten unterscheiden, etwa ob der Gesprächspartner eine Frage stellen, etwas bestellen oder sich beschweren möchte.

New Work

Mit diesem Begriff wird die neue Art, Leben und Arbeiten zu verbinden, umschrieben. Durch die allgegenwärtige Verfügbarkeit des Internets und von Cloud-Anwendungen ist eine physische Anwesenheit bei vielen Arbeiten nicht mehr erforderlich. Arbeitnehmer wollen nicht länger in festen Strukturen, sondern „agil“ arbeiten, und profitieren von mehr Freiheit und Selbstbestimmung. Hierarchien in Unternehmen sind abgebaut, strategische Entscheidungen werden gemeinschaftlich getroffen.

Speech-to-Text/Text-to-Speech

KI-Services, die gesprochene Sprache in Text übersetzen und umgekehrt.

Thin-Client-Infrastruktur

Verteilte Infrastruktur, bei der die Hauptfunktionen in der zentralen Serverkomponente lokalisiert sind. Auf dem Client wird nur ein Browser benötigt, über den der Anwender auf die Serverkomponenten zugreifen kann. Der Thin-Client-Ansatz ermöglicht ein weitgehend Hardware-unabhängiges Arbeiten, da alle wesentlichen Schritte auf dem Server ausgeführt werden. Er erleichtert auch das Deployment von Applikationen und reduziert den Wartungsaufwand.

Nuance (NASDAQ: NUAN) ist Pionier und Marktführer im Bereich der dialogorientierten und kognitiven KI-Innovationen, die Intelligenz in die tägliche Arbeit und das tägliche Leben bringen. Der Marktführer liefert Lösungen, die die menschliche Sprache verstehen, analysieren und darauf reagieren können, um die Produktivität zu steigern und die menschliche Intelligenz zu verstärken. Auf Basis jahrzehntelanger Erfahrung in den Bereichen Domains und Künstliche Intelligenz arbeitet Nuance mit Tausenden von Unternehmen zusammen, die in globalen Branchen wie Gesundheitswesen, Telekommunikation, Automobil, Finanzdienstleistungen und Einzelhandel tätig sind, um stärkere Beziehungen und bessere Erfahrungen für ihre Kunden und Mitarbeiter zu schaffen.



Weitere Information finden Sie online unter www.nuance.de/dragon, Xing, LinkedIn oder Facebook

Nuance Communications Germany GmbH

Willy-Brandt-Platz 3
81829 München