



Künstliche Intelligenz mit DMS/ECM-Lösungen – da geht etwas und es kommt noch viel mehr!

Als vor über 40 Jahre die elektronischen Archivierungslösungen die Mikrofilmtechnologie ablösten, war der Begriff Künstliche Intelligenz in Verbindung mit Dokumenten Management noch in weiter Ferne.

Heute wird über Künstliche Intelligenz (kurz „KI“) im Zuge von autonomen Fahren, Verkehrslenkung in Städten, automatisierten Kundenservice mit virtuellen Assistenten, sowie Robotik und digitale Bilderkennung vielfältig berichtet und genutzt.

Im Bereich von elektronischen Dokumenten und dem Enterprise Content Management verbindet man Dokumente und KI oft mit den Themen „intelligente Dokumentenerkennung“, „Wissensextraktion“ oder neuerdings auch mit der Abkürzung RPA für Robotic Process Automation.

Es stellt sich die Frage, wo man nun in einem ECM künstliche Intelligenz bereits findet und wie die Entwicklung weitergehen kann. Der Beitrag nennt

einige Beispiele und gibt einen Ausblick.

Dabei sei zunächst der Kürzel „ECM“ für Enterprise Content Management“ in Erinnerung gerufen. Es geht hierbei um die Planung, Steuerung und Kontrolle des unternehmensweiten **Content** – zumeist sind damit Dokumente in jedweder Form gemeint, obwohl „Content“ auch multimediale Inhalte sein können (Bild 1).

Was vor Jahrzehnten mit Spooldatenaufgabe, Scannen, Volltextindexierung und Indexdatenbanken die Ablösung der Mikrofilme einleitete, ist heute ein einheitliches System oder ein Dienst, der in vielen Fachanwendungen oder in der Cloud angeboten wird.

Es gibt für die Content-Verwaltung viele verschiedene Lösungen – im Idealfall jedoch ein einziges ECM-System, das alle Funktionen anbietet oder integrieren kann. Und heute geht es mehr als nur um Scannen, Ablegen und Suchen – oder doch nicht und was hat Künstliche Intelligenz damit zu tun?

Künstliche Intelligenz – wovon sprechen wir?

Künstliche Intelligenz (KI, auch artificial intelligence (engl.), abgekürzt AI) ist ein Teilgebiet der Informatik, welches sich mit der Automatisierung intelligenten Verhaltens und dem maschinellen Lernen befasst. Der Begriff ist insofern nicht eindeutig abgrenzbar, als es bereits an einer genauen Definition von „Intelligenz“ mangelt. Dennoch wird er in Forschung und Entwicklung verwendet (vgl. https://de.wikipedia.org/wiki/K%C3%BCnstliche_Intelligenz).

Etwas knapper formuliert handelt es sich um selbständig agierende Maschinen oder Programme: Diese denkenden Computer werden beispielsweise beim autonomen Fahren genannt, bei medizinischer Diagnostik, in der modernen Landwirtschaft, aber auch bei computergestützter Beratung und Fallbearbeitung im Versicherungs- und Finanzdienstleistungsbereich.

Klassisches ECM Vereinfachte Darstellung

▪ Komponenten:

- Erfassung (Capture)
- Verwaltung (Manage)
- Speicherung (Store)
- Ausgabe (Deliver)
- Bewahrung (Preserve)

▪ Erfassung →

▪ Anwendungsfelder

- Document Management (DM)
- Collaboration (Groupware)
- Web Content Management (WCM)
- Records Management (RM, Archive)
- Workflow/Business Process Management (WF/BPM)

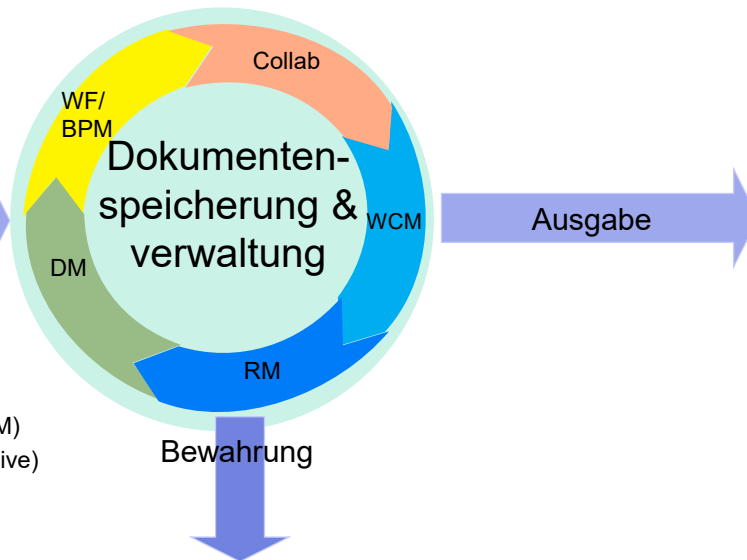


Bild 1: Vereinfachte Darstellung von ECM

Führt man mit einer künstlichen Intelligenz einen Dialog oder Chat und bemerkt dabei keinen Unterschied zu einer natürlichen, menschlichen Intelligenz, dann ist der sogenannte Turing-Test bestanden und es liegt eine „künstlichen Intelligenz“ vor.

Diese KI wird zunehmend auch mit Dokumenten Management bzw. Enterprise Content Management eingesetzt, dabei wird in schwache und starke KI-Ansätze unterschieden.

ECM-Anwender profitieren von KI-Ansätzen

Die Arbeit mit elektronischen Dokumenten ist heute nicht mehr nur mit manuellem Verschlagworten verbunden. Viele Dokumente lassen sich maschinell vorklassifizieren oder weiterleiten. Drei Beispiele seien hierzu vorgestellt.

Erfassung von Content: Intelligente Datenaufbereitung am Beispiel der Rechnungserkennung

Im ECM-Bereich wird das „maschinelle Erkennungs- oder Datenextraktionsverfahren“ gerne als Beispiel für die aktuell praktizierte intelligente Software-Unterstützung angeführt. Dabei werden Rechnungen einer Erkennungssoftware zugeführt und diese gibt Absender, betroffene Gesellschaft, Brutto-, Nettowerte

etc. zurück und prüft gewisse Formalregeln ab. Werden Rechnungen nicht korrekt erkannt, wird über eine Lernroutine oder ein automatisches Lernen die Verbesserung gemerkt bzw. gespeichert.

Dazu merkt sich die Software für das Rechnungslayout eines Lieferanten beispielsweise die Koordinaten des ersten Begriffes oder Wertes, um beim nächsten Mal die Werte von dieser Stelle abzugreifen.

So kann sukzessive die Erkennungsrate erhöht werden und die verschiedenen Dokumenteninhalte den richtigen Attributfeldern zugeordnet werden. Systematisch betrachtet passiert dabei folgendes:

1. Klassifikation – Finden von benötigten Informationen:

Für die Prüfung bestimmter Nummern, z.B. Umsatzsteuer-ID wie DE123456789 kann ein regulierter Ausdruck den Inhalt prüfen, der in etwa wie folgt aussieht: (DE)?[0-9]{9}. Solche Regeln kann es auch für andere Werte geben (Datum, Steuernummern, Postleitzahlen, Vertrags-, Kunden-, Bestellnummern, etc.).

Klassifizierung geht aber darüber hinaus, denn es gibt auch unstrukturierte Dokumente (Informationsquellen) aus denen mit Hilfe o.g. Informationen letztendlich zu bestimmen ist, ob eine Rech-

nung, Bestellung, Beschwerde oder Vertragsdokument vorliegt.

2. Kontextualisierung - Zusammenhang zu anderen Informationen herstellen:

Die erkannten Informationen sind im nächsten Schritt in einen Zusammenhang zu setzen. Bezüglich o. g. Rechnungsbeispiel können beispielsweise drei Datumsangaben auftreten: Rechnungsdatum, Lieferdatum, Bestelldatum und vielleicht das Zahldatum als vierter Wert. „Das Rechnungsdatum ist das Datum, das am wenigsten in der Vergangenheit liegt, das Zahldatum wird in der Zukunft liegen“. Diese Regel kann nur durch Beurteilung der anderen Werte ausgeführt werden.

Ebenso verhält es sich mit Vertragsnummern, Kundendaten oder Personalnummern: Sind diese erkannt und mit einem Vergleichsdatenbestand gegenübergestellt worden, können Sie in Kontext zu anderen strukturierten Daten gesetzt werden.

3. Bedeutung erkennen:

Bei semi-strukturierten Quellen ist die Bedeutung oft ableitbar (Rechnung führt zur Buchung und Bezahlung), aus unstrukturierten Inhalten gelingt dies schwieriger: Liegt eine Adressänderung,

eine Kündigung oder eine neue Lieferadresse vor?

Die Festlegung der Bedeutung ist anspruchsvoller. Zur Lösung sind weitere Regeln, mehr Rechenleistung, eine größere Lernmenge oder KI-Verfahren mit neuronalen Netzwerken notwendig.

Fassen wir die intelligente Erkennung zusammen:

Man erkennt, dass es sich bei Punkt 1) und 2) teilweise um ein Regelwerk handelt, das leicht bis sehr anspruchsvoll formuliert werden kann. Es ist geeignet für festgelegte, semi-strukturierte – und nicht beliebig freie Texte. Es gibt zwar faszinierende Erkennungsergebnisse, aber ebenso auch bizarre Fehlinterpretationen. Technisch betrachtet funktionieren diese Erkennungsverfahren nicht mit Hilfe neuronaler Netze und auch das selbständige Entscheiden ist begrenzt. Der Verfasser meint, dass diese Verfahren als „schwache künstliche Intelligenz“ eingeordnet werden können.

Mit der Fähigkeit zur Interpretation der Bedeutung (Punkt 5) gelangt das Verfahren aber durchaus in Richtung „echte bzw. starke künstliche Intelligenz“.

Ein weiteres Beispiel soll die Beziehung zwischen KI und ECM beschreiben, und zeigen, dass nicht nur die Erfassung, sondern auch die Content-Bearbeitung von KI profitieren kann:

Workflow: Automatisierung durch Robotic Process Automation (RPA)

Liegen extrahierte Daten aus Dokumenten vor, resultieren daraus im Allgemeinen verschiedene Aktivitäten, wie Suchen, Zusammenführen, Ergänzen und Bearbeiten von Daten. Oft erfolgt dies in strukturierten Fällen durch definierte Workflows, häufig auch über verschiedene Anwendungen hinweg. In diesen Fällen werden Daten aus Anwendungen ermittelt, importiert oder manuell erfasst. Für die Schritt-Reihenfolge wird wie bei Workflow-Lösungen ein Prozessmodell erstellt, wonach die Bearbeiter angesteuert werden, um verschiedene Aktivitäten durchzuführen.

Die Übertragung der Daten via Dialogmasken folgt dabei oft einem gleichen Muster, so dass man ähnlich einem Makro, diese Masken automatisch befüllen kann (Screen Scraping).

Fortgeschrittene Verfahren kombinieren nun die Schrittfolgen des Workflows mit den extrahierten bzw. ermittelten Daten und übertragen diese mittels Screen Scraping automatisch in eine andere Anwendung. In diesem Fall spricht man mittlerweile von RPA.

Denn die RPA-Technologie verhält sich wie ein menschlicher Anwender, der sich in Anwendungen einloggt, Daten eingibt, Aufgaben erledigt und sich abmeldet.

Hat man die Zeit für die Analyse und Planung nicht, sollen KI-Ansätze das Verhalten „aufzeichnen“ bzw. „lernen“ und dann selbst durchführen. Diese trendbasierte Analyse stellt fest, dass beim Fehlen bestimmter Angaben, der Geschäftspartner angeschrieben wird und ihm bestimmte Angaben und einem Bearbeitungswunsch mitgeteilt werden.

Ähnliche Ansätze werden bei Service-Centern verfolgt, bei denen die ersten Antworten auf eingehende E-Mails durch die Maschine auf Basis der Inhaltsanalyse erstellt werden. In manchen Versicherungen liegen diesbezüglich bereits umfangreiche Erfahrungen vor: Die eingehende **Briefpost, E-Mails oder auch Anrufe** werden automatisch klassifiziert und weiterbearbeitet. Beispielsweise kann mittels RPA ein Konto oder eine Versicherungspolice automatisch in einem System erfasst werden, soweit alle relevanten Unterlagen und Informationen vorliegen.

Neben diesen „Inhaltsbasierten“-KI-Ansätzen in ECM-Lösungen geht die Entwicklung dann weiter zu automatisch arbeitenden Computerprogrammen (sogenannte „Bots“ von engl. „robot“). Diese Bots empfehlen Policen oder regulieren einfache Schadensmeldungen. Damit ist ein Vertragsabschluss unter einer Minute und eine Schadensregulierung in wenigen Sekunden möglich – so zumindest bei lemonade.com. Werden von den Bots Entscheidungen bei Unsicherheit getroffen – sei es auch nur die Weiterleitung an einen Bearbeiter kann man von Verfahren auf Basis starker KI sprechen.

Die Bots sind dabei ECM-unabhängig und eher fachspezifisch entwickelt. Die ECM-Relation ergibt sich aus der Analyse der eingehenden Mails oder Dokumente und den Workflow-Möglichkeiten in einem ECM.

Neben Erfassung und Bearbeiten suchen viele Anwender in ECM-Lösungen und anderen Quellen unternehmensweit nach Informationen. Auch hier gibt es intelligente Ansätze:

Intelligente bzw. kognitive Suche

Der Begriff intelligente Suche wurde in letzter Zeit durch das frischer klingende „kognitive Suche“ verdrängt.

Der Marktforscher Forrester definiert kognitive Suche als „Kombination von Indexierung, Verarbeitung natürlicher Sprache und Maschine-Learning-Technologien, um eine zunehmend relevante

Wissensbasis aus allen Quellen unstrukturierter und strukturierter Daten aufzubauen. Natürlichsprachliche und verdeckte Abfrageschnittstellen (Bereitstellung von relevanten Informationen ohne explizite Abfrage) werden dabei verwendet, um dem Nutzer Wissen via Text, Sprache, Visualisierungen und / oder sensorisches Feedback zu vermitteln.“ Die Kommunikation in natürlicher Sprache weist auf die Verwendung starker künstlicher Intelligenz hin und erfordert erhebliches technisches Leistungsvermögen.

Dabei ist die Suche durch den Anwender eigentlich der zweite Schritt innerhalb der kognitiven Suche. Denn Voraussetzung für das Auffinden ist in einem ersten Schritt die Text- bzw. Datensammlung und –aufbereitung. Die KI-Technologie kann dabei helfen, Verbindungen zwischen Inhalten aufzubauen und Relevanzen für die Ergebnisverwendung festzulegen.

Unstrukturierte Dokumente, aber auch sogenannte Rich-Media-Inhalte wie Sprachaufzeichnungen oder Videos, sollen hierbei für das Auffinden erschlossen werden. Ebenso soll eine Abfrage mittels natürlicher Sprache verstanden werden. Dass dies sehr schwierig ist, belegen Berichte, wonach in anderen Anwendungsfällen mit Spracherkennung tausende von Sprachmitschnitten durch Amazon-Mitarbeiter transkribiert werden, um die Spracherkennung zu verbessern. Mit solchen Erfahrungen und KI-Technologie wird das Verstehen und Zerlegen und schließlich auch das Auffinden von Texten in Sprach-Aufzeichnungen in Zukunft besser möglich sein. Dass die Spracherkennung und Umwandlung in Text bereits von jedem genutzt werden kann, sieht man an Mailbox-Diensten, die die hinterlassene Sprachnachricht als Text übersetzen.

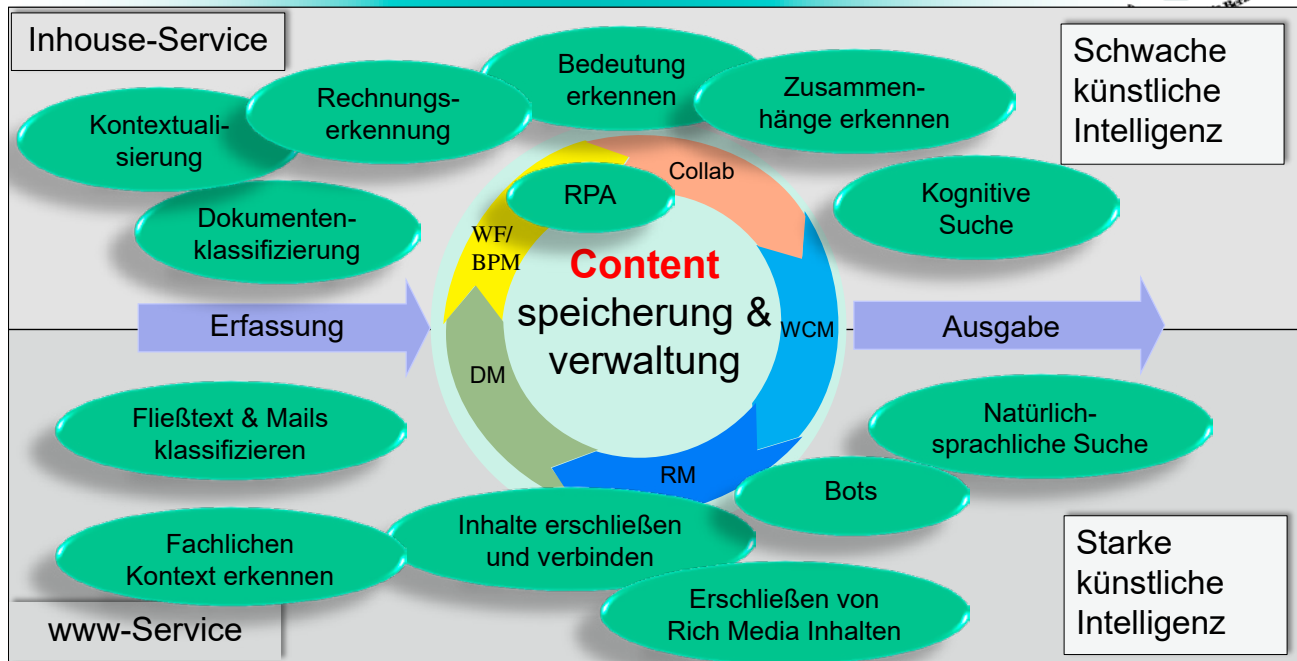
Die heutigen Suchverfahren und -lösungen verstehen sich dabei als eigenständige unternehmensweite Systeme, die ECM-Lösungen nur als einen von mehreren Datenquellen verwenden und auswerten. Daher gibt es solche anspruchsvollen Suchverfahren in ECM-Lösungen eher selten, obwohl sie dort vermutlich den höchsten Nutzen hätten.

Moderne ECM-Lösungen sollten solche Dienste indes anbinden und nutzen können.

Zwischenfazit: Ist das wirklich alles oder kommt noch mehr KI?

Die Beispiele zeigen, dass im Bereich ECM und KI kleine und große Anwendungsmöglichkeiten und -potenziale bestehen. Dabei geht es aktuell mehrheit-

ECM und KI



info@dr-weiss.com

DWB Dr. Dietmar Weiß Beratung
www.dr-weiss.com

Bild 2: Schwache und starke KI-Ansätze für ECM-Lösungen

lich um die Texterschließung als um die Automatisierung. Bei strukturierten und semistrukturierten Inhalten gelingt die Klassifikation und Kontextualisierung auch mittels Regelwerk. Standardisierungsentwicklungen gehen allerdings in Richtung XML-Beschreibung und ersetzen dadurch die „intelligente Erkennung“. Die Feststellung der Bedeutung entfällt bei XML-Beschreibungen ebenfalls, bei Fließtexten ist dazu indes „richtige“ KI-Technologie notwendig.

Daher sieht der Verfasser bei unveränderter ECM-Nutzung eine begrenzte Partnerschaft von KI und ECM-Lösungen:

Bleibt der Fokus von ECM-Lösungen auf der Ablage und Bearbeitung von **schriftlichen Inhalten**, wird der Entwicklungsgrad in Verbindung mit KI-Technologien bald erschöpft sein: Die Standardisierung von Inhalten mittels XML senkt den Bedarf an KI-Technologie. Zudem scheint die Fokussierung der ECM-Lösungen auf schriftlichen Content zwar ein großes, aber begrenztes Einsatzfeld zu sein.

Perspektivisch wächst die Verwendung von Rich-Media-Inhalten und diese sind folglich zu erschließen.

Rich Media: Die Kommunikation ändert sich – ECM-Lösungen auch? Blickt man genau auf die Kommunikationsentwicklung könnte man einen Trend zum multimedialen Austausch und Mobilgerätekommunikation erkennen:

- Kundenkontakte via Call-Center-Ansprache erfolgen über Sprache. Werden Sprachnachrichten hinterlassen, sind diese abzuhören und in Inhalte bzw. Daten zu wandeln und in Geschäftsvorfälle zu erfassen.
- Dort wo Sprachaufzeichnung vorgesehen oder verpflichtend ist, werden oft nur einzelne Dateien abgelegt und mit Datum und Rufnummer verschlagwortet, der Inhalt bleibt verschlossen.
- Kommunikations- und Kollaborationstools für Unternehmen nutzen ebenfalls Sprachkommunikation: Statt einer schriftlichen Meldung gibt es eine mündliche Nachricht.
- Durch die Nutzung von mobilen Endgeräten werden Inhalte abfotografiert statt gescannt. Es gibt Verzerrungen, Schrägaufnahmen, Schatten auf fotografierten Dokumenten und andere Störungen, die beim Scannen nicht auf-

treten und verarbeitet werden müssen – zur Lösung werden andere Fehlertoleranzen und neue Korrekturverfahren benötigt.

Hieraus resultieren neue Anforderungen:

Zum einen gelangen Informationen nicht regelmäßig via Postanlieferungen über eine „zentrale Poststelle“ mit Scanning-Infrastruktur und Verschlagwortung in das Unternehmen, sondern dezentral, digital, dauerhaft und ungeplant (im Gegensatz zur täglichen Postanlieferung).

Zum anderen liegen - technisch betrachtet - die Informationen immer mehr in Bildern oder Sprachfiles „verpackt“ und nicht recherchierbar vor. D.h. eine Menge an Wissen ist unbekannt bzw. nicht recherchierbar oder muss sehr aufwändig ausgewertet und indexiert werden.

So kann aus einem Podcast, einem Interview oder einer aufgezeichneten Telefonkonferenz momentan kein schriftliches Verlaufsprotokoll (mit wörtlicher Wiedergabe) oder Ergebnisprotokoll (mit inhaltlichem Fazit) in Form eines schriftlichen Dokumentes maschinell

erzeugt werden. Die Suche nach Begriffen ist ebenfalls nicht möglich.

Dabei bietet sich die Ablage von Dateien jedwedem Formats in digitalen Akten sowie die Bearbeitung durch Workflows in einem ECM geradezu an, da sowohl Akten als auch Prozesse hier schon vorhanden und definiert sind – aber das ist alles auf Dokumente und deren Inhaltsextraktion via OCR ausgelegt. Denn multimediale Inhalte können aus mehreren Gründen nicht erschlossen werden:

- a) „Anzeigen“ oder „Abspielen“ ist nur eingeschränkt möglich: Während es gute Programme zum Blättern, Anzeigen und Navigieren in großen Dokumenten gibt, fehlt vergleichbares für andere Formate
- b) Schnelles Arbeiten ist mit diesen Formaten nicht immer möglich
- c) Die Handhabung und Bearbeitungsmöglichkeiten sind im Vergleich zu Dokumenten unterentwickelt (Ausschneiden & Zusammenfügen von Inhalten, Hervorheben von Inhalten oder Stellen etc.)
- d) Suchen und Finden von Inhalten (in einer Sprachnachricht) gelingt nicht.
- e) Spracherkennung wird im Vergleich zu Texterkennung kaum von einem ECM-Anbieter angeboten.

Es soll hier kein Missverständnis auftreten:

Formulare, Rechnungen und relevanter Schriftverkehr werden weiterhin auf analogen oder elektronischen Dokumente ausgetauscht, denn die Schriftform ist für viele Geschäftsvorgänge vorgeschrieben. Wenn aber Ergebnisse aus Mobiltelefonen (Fotos, Sprachnachrichten) oder Sprachaufträgen verwendet werden, gibt es neue Anforderungen an ECM-Lösungen. Was Siri, Alexa, Cortana oder Diktierprogramme verstehen, sollte auch bei Sprachnachrichten extrahiert werden können, so dass die Wörter recherchierbar sind. Ebenso verhält es sich bei Fotos, Podcast oder anderen Aufzeichnungen (z. B. von Web-Konferenzen).

Ausblick: KI als Technologie für die Erschließung der neuen Medien in ECM-Lösungen!

Die oben beschriebenen Szenarien (Erkennung, Workflow, RPA etc.) stehen vielen Unternehmen als Funktionalität in einem ECM-System oder durch entsprechende Speziallösungen zur Verfügung. Sie haben alle gemein, dass die Dienste Dokumente betreffen und in der

IT-Landschaft des Unternehmens eingerichtet werden können. Teilweise können die Dienste mittlerweile aus Cloud-Lösungen gebucht werden. Es handelt sich in Bild 2 dabei um die Dienste in der oberen Bildhälfte und damit um **schwache KI-Ansätze**, die sich auf „Dokumente“ spezialisiert haben und unternehmensintern eingesetzt werden können.

Blickt man auf die **weitere Entwicklung in der unteren Bildhälfte**, gilt es künftig **Rich Media** zu integrieren, was komplexere Unterstützung in Form **starker KI** erfordert, die aufgrund der Rechenleistung, Lernmenge und Erfahrung oft nur als externer Dienst eingebunden werden kann.

Beispielsweise sei die OCR-Technologie (Optical Character Recognition) für die Zeichenerkennung angeführt. Diese Technologie war zu Zeiten des Mikrofilms revolutionär und konnte die Dokumenteninhalte erschließen. Mit dem Aufkommen von Rich Media besteht eine ähnliche Situation: Der Inhalt kann vom klassischen ECM nicht erschlossen werden – und das bei wachsender Menge!

Eine neue KI-basierte Technologie für **Open Content Recognition** („OCR 4.0“) ist notwendig und kann mit KI-Technologie ein Quantensprung darstellen.

Das wird zwar nicht jedes Unternehmen zu jedem Zeitpunkt benötigen. Aber gerade die flexible Nutzung ist eine Stärke web- oder cloudbasierter Dienste, die hohe Rechenleistung oder neuronale Netz-Technologie anbieten, die sich ein „normales Unternehmen“ nicht leisten könnte. Und diese buchbaren Cloud-Dienste machen es jedem möglich, an solche Spitzentechnologie zu gelangen.

Von den KI-Technologien profitieren bereits jetzt viele Unternehmen, auch

Der Autor



Foto:

Dr. Dietmar Weiß (Mail dwb@dr-weiss.com) ist ECM-Spezialist und unterstützt Unternehmen bei der Erstellung von Fachkonzepten, Prozessoptimierung, Einführung, Auswahl und Integration von ECM-Lösungen. Er hat Eingangsrechnungsbearbeitungssysteme bereits in 15 europäischen Ländern eingeführt und für Installationen entsprechende Verfahrensbeschreibungen erstellt.

Ein weiteres Spezialgebiet ist die Planung und Durchführung der Migration von ECM- und Archivsystemen.

in ihren vorhandenen ECM-Lösungen. Da sich die Kommunikation und deren Medien ändern, sollte die IT-Landschaft und auch die ECM-Infrastruktur darauf vorbereitet sein und sich anpassen. Die KI-Technologie kann diese Lösung sein, insbesondere wenn es um moderne Dokumente und Rich Media geht.

Gelingt dies nicht – dann wird etwas Neues kommen, wie vor 40 Jahren gegen Ende der Mikrofilm-Ära. ◀

Quellen:

https://de.wikipedia.org/wiki/K%C3%BCnstliche_Intelligenz

Googles Sprachassistent: Mitarbeiter hören und bewerten Audioaufnahmen
<https://www.heise.de/newsticker/meldung/Googles-Sprachassistent-Mitarbeiter-hoeren-und-bewerten-Audioaufnahmen-4467985.html?view=print>
Abruf am 12.07.2019

Amazon-Mitarbeiter tippen zum Teil Alexa-Sprachbefehle ab
<https://www.heise.de/newsticker/meldung/Amazon-Mitarbeiter-tippen-zum-Teil-Alexa-Sprachbefehle-ab-4374871.html>
Abruf am 12.07.2019

Amazon Workers Are Listening to What You Tell Alexa
<https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio>
Abruf am 12.07.2019