

01-2021

public

Kundenmagazin der .msg
für den Public Sector



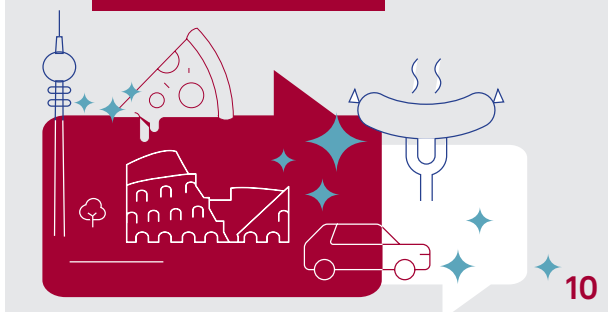
SCHWERPUNKT
DATEN UND KI

KI, entscheide Du mal!
KI-Transparenz

Registermodernisierung
Zieht Deutschland nach?

Sprache ist Magie
Spracherkennung durch KI

INHALT



10



22



42

- 3** **Editorial**
von Dr. Andreas Zamperoni
- 4** **Aus Java wird Jakarta – und die Folgen davon**
Ein Interview mit Werner Keil
- DATEN UND KI**
- 9** **Können Algorithmen diskriminieren?**
Zwischenruf von Werner Achtert
- 10** **Sprache ist Magie – Spracherkennung durch KI**
von John Loutzenhiser
- 16** **KI, entscheide Du mal!**
von Hans-Ulrich Striffler und Dr. Andreas Zamperoni
- 22** **Is It A Match? Dublettenerkennung mit KI**
von Nedislav Nedyalkov und Dominik Müller
- 28** **Chatbots – Bürgerservice neu gedacht**
von Christian Meyer

- 37** **Registermodernisierung – zieht Deutschland nach?**
von Dr. Christian Kiehle

MODERNE VERWALTUNG

- 42** **YouConnect – ein IT-System für behördenübergreifenden Datenaustausch**
von Martin Meyer, Bundesagentur für Arbeit

PROJEKTMANAGEMENT

- 48** **Project Office as a Service – ein langer Weg seit Gantt**
von Bernd Hahn und Ulrich Berghaus

INFORMATIONSTECHNOLOGIE

- 52** **Hochwertige Entwicklungsdokumentation automatisch mit AsciiDoc**
von Martin Sernow

Herausgeber

Jürgen Fritsche, Geschäftsleitung
Public Sector, msg systems ag

Robert-Bürkle-Str. 1
85737 Ismaning
Tel.: +49 89 96101-0, Fax: -1113
E-Mail: info@msg.group
www.msg.group

Verantwortlich

Dr. Stephan Frohnhoff
(Vorsitzender),
Rolf Kranz,
Dr. Aristid Neuburger,
Karsten Redenius,
Dr. Frank Schlottmann,
Dr. Jürgen Zehetmaier

Redaktion

Dr. Andreas Zamperoni (Chefredakteur),
Karin Dohmann,
Dr. Katrin Ehlers

Konzept und Layout

Eva Zimmermann

Bildnachweis

msg systems ag: S. 9, 6, 8
Adobe Stock: Umschlag, S. 9, 10, 16, 22,
28, 34, 35, 37, 42, 47, 48, 51, 52

Produktion

Meisterdruck GmbH,
Kaisheim

Der Inhalt gibt nicht in jedem Fall die
Meinung des Herausgebers wieder.
Nachdrucke nur mit Quellenangabe
und Belegexemplar.



Anthropomorphismus ist der Fachbegriff dafür, Tieren, Göttern, aber auch unbelebten Dingen wie Maschinen oder Software menschliche Eigenschaften zuzuschreiben. Computer „gehören“ uns nicht, KI-Systeme „lernen“, Chatbots „verstehen“ uns (oder auch nicht). Durch diese Vermenschlichung tragen wir auch zur Mystifizierung von KI bei. Dabei ist genau das Gegenteil dringend vonnöten, nämlich Aufklärung! Denn „Aufklärung ist der Ausgang des Menschen aus seiner selbstverschuldeten (!) Unmündigkeit“ (Immanuel Kant).

Aufklärung zum Beispiel bezüglich der sogenannten Maschinenethik: Sie soll eine Antwort auf die Frage geben, ob das selbstfahrende Auto im Zweifelsfall den Familienvater oder die rentenbeziehende Großmutter überfährt. Die Antwort darauf hängt entscheidend von

der in einer Gesellschaft vorherrschenden Ethik ab: utilitaristisch – sehr populär bis ins 20. Jahrhundert und immer noch in autoritären Regimes? Dann die Großmutter. Moralisch im Sinne von Kant, den Humanisten oder christlicher Ethik? Dann warten wir noch ein paar Jahre, bis wir selbstfahrende Autos bauen können, die diese Frage „beantworten“ (auch ein Anthropomorphismus), ohne dass jemand zu Schaden kommt. Doch wir können die unserer Gesellschaft angemessene Ethik nur dann in unsere Maschinen programmieren, wenn wir die ethischen Systeme studieren, durchdringen und verstehen.

Ein spannendes und herausforderndes Thema, das uns noch lange beschäftigen wird. Für uns Grund genug, den Schwerpunkt dieser .public auf das Thema „KI & Daten“ (in der öffentlichen Verwaltung!) zu legen und damit ein wenig zur „Aufklärung“ beizutragen.

Schon heute ist es nicht ungewöhnlich, dass Bürgerinnen und Bürger ihre Fragen an die Verwaltung mit Chatbots „besprechen“. Im Artikel „Chatbots – Bürgerservice neu gedacht“ wirft unser Autor Christian Meyer einen Blick auf die Gegenwart und die Zukunft maschinellen „Verstehens“ in der öffentlichen Verwaltung. In „KI – Entscheide Du mal“ klären Autoren Hans-Ulrich Striffler und ich auf, wie die Akzeptanz automatisierter, KI-basierter Entscheidungen erhöht werden kann. In „Is it A Match? Dublettenerkennung mit KI“ beschreiben unsere Autoren Nedislav Nedyalkov und Dominik Müller ein KI-basiertes Experiment zur Adressdubletten-Erkennung. Und ganz ohne Magie geht es dann doch nicht (oder doch?), wenn KI-Kompetenz-Center-Leiter John Loutzenhiser in seinem Beitrag in „Sprache ist Magie“ uns anschaulich die Grundlagen der Spracherkennung in der KI näherbringt.

Zurück zur Realität: In „Zieht Deutschland nach?“ kommentiert Dr. Christian Kiehle die Realität der Registermodernisierung in Deutschland, und Martin Meyer von der Bundesagentur für Arbeit stellt „YouConnect“, ein spannendes und erfolgreiches Projekt behördenübergreifenden Datenaustauschs, vor.

Am Beginn unserer aktuellen Ausgabe steht diesmal ein besonderes Interview: Werner Keil, Urgestein der Java-Schöpfer und -Weiterentwickler, erklärt in „Von Java zu Jakarta“ unserem Java-Experten Alexander Salvanos, welche Auswirkungen der Übergang des JEE-Standards von Oracle auf die Eclipse Foundation auch für Ihre Behörde haben könnte. Ich wünsche Ihnen viel Spaß beim Lesen!

Dr. Andreas Zamperoni
Chefredakteur .public



VON JAVA ZU JAKARTA

EIN INTERVIEW MIT WERNER KEIL

Geführt von Alexander Salvanos, Senior IT Consultant bei msg systems ag

Seit 1998 ist die Java Enterprise Edition (Java EE) der Industriestandard, der es sich unter anderem zur Aufgabe gemacht hat, die besonderen Sicherheitsprobleme der geschäftskritischen Unternehmensanwendung von Industrie und Behörden zu lösen. Und auch wenn dieser Industriestandard im Jahr 2018 an die Eclipse Foundation ging und dort auch noch mit Jakarta EE einen neuen Namen bekam, so ist es immer noch der gleiche Industriestandard. Das Ziel von Jakarta EE (ehemals Java EE) ist es nach wie vor, eine Plattform anzubieten, mit der sich hochskalierbare, zuverlässige und sichere Unternehmensanwendungen erstellen lassen.

Und genau wie Java EE ist auch Jakarta EE Teil des Java Community Process (JCP). Sie erinnern sich: Beim JCP werden Java-Neuerungen als Spezifikationsanfrage (JSR) eingereicht. Nimmt das Exekutivkomitee die Anfrage an, so durchläuft sie einen formalen Prozess, bei dem sie eine Expert Group genauer

spezifiziert. Auch alles andere, wie die technologischen Kompatibilitätstests (TCK), wird auf gleiche Weise erstellt. Dennoch gibt es einen kleinen Unterschied, und der liegt im Führungsmodell. Denn Jakarta EE ist nun Community basiert. Hiermit ist nicht nur gemeint, dass sich jeder Java-Programmierer an der Programmierung als gleichgestellter Entwickler beteiligen kann, sondern auch, dass die neuen Spezifikationen und Entwicklungsprozesse für Jakarta EE offen und herstellerneutral sind und gleiche Wettbewerbsbedingungen für alle Teilnehmer bieten. Das klingt alles sehr positiv. Dennoch gab es Hürden: Beispielsweise stellte der Transfer der Java-EE-Quelltexte von Oracle an die Eclipse Foundation eine große Herausforderung dar. Gerade hierbei ist es daher sehr wichtig, dass erfahrene Java-EE-Persönlichkeiten auf dem Schiff mit der nun neuen Flagge mitreisen. Eine der herausragenden Personen hierbei ist Werner Keil, der bereits zu Zeiten von Sun Microsystems Mitglied des Java Community Process Executive Committee

und der Java EE Expert-Group tätig war und dort mit zahlreichen Awards prämiert wurde. Seine Auszeichnungen füllen Regale: JCP Award Winner, JCP Spec Lead/JSR of the Year, Speaker of All Time, JCP Award Winner, Member of the Year JCP Award Winner, Most Significant JSR Spec Lead at JSR 385 JCP Award Winner, JSP Executive Committee Member, Java EE Expert Group Member und nun auch Jakarta EE Specification Committee Member und Committer Representative der Jakarta EE Working Group. Wir hatten die Gelegenheit, mit Werner Keil zu sprechen.

msg: Vielen Dank, Werner, dass Du Dir heute die Zeit genommen hast für uns. Wann und wie bist Du erstmalig mit Java EE in Berührung gekommen und wie bist Du zu diesen hohen Positionen in der Java- und Java-EE-Community gekommen?

Werner Keil: Erstmals bei einer JavaOne-Konferenz war ich 1998, nachdem mir mein Kunde aus dem öffentlichen Sektor, nämlich die österreichische Pensionskasse, freundlicherweise ein Ticket mit Flug zur JavaOne finanzierte. Auf der Konferenz wurde auch das spätere Java EE erstmalig vorgestellt. In das JCP Executive Committee und in die Java EE Expert Group wurde ich aber erst im Jahre 2007 gewählt, nachdem ich für Sun Microsystems und BEA/Oracle gearbeitet habe. Diese Tätigkeit hat mich dann auch inspiriert und in die Lage versetzt, die Standards später aktiv mitzugestalten.

msg: 2018 ging Java EE 8 von Oracle an die Eclipse Foundation über, wo sie kurze Zeit später den Namen Jakarta EE 8 erhielt. Welche Auswirkungen, denkst Du, hat diese Entscheidung für den Industriestandard und somit für unternehmenskritische Geschäftsanwendungen gehabt?

Werner Keil: Ein großer Vorteil des Übergangs von Oracle an die Eclipse Foundation ist, dass seitdem Entscheidungen demokratischer gefällt werden. Vorher hatte Oracle stets das letzte Wort, denn es waren ja meistens Oracle-Mitarbeiter an Java EE beteiligt. Und jede Spezifikation hatte auch wieder einen Spezifikationsleiter von Oracle usw. Dadurch hatte Oracle ein Vetorecht. Das gibt es in dieser Form jetzt nicht mehr. Denn durch den Wechsel haben nun alle das gleiche Mitbestimmungsrecht. Dadurch kann auch niemand mehr diktieren, was er im Standard haben will.

Wenn beispielsweise CORBA oder EJB aus dem Standard herausgenommen werden sollen, wird diese Entscheidung demokratisch gefällt. Und auch ich als Committer Representative, also als Vertreter der Jakarta EE Committer Members, habe

eine Stimme, genauso wie die Strategic Enterprise Members, die einen Jahresbeitrag bezahlen. Das sind beispielsweise die Unternehmen IBM, Red Hat usw. Zwar hat Oracle als Industrieunternehmen hierdurch auch eine Stimme, aber es hat kein Vetorecht mehr. Und auch die verschiedenen Arbeitsgruppen, wie das Specification Committee, oder führende Funktionäre der Eclipse Foundation, können nicht über diesen demokratischen Prozess hinweg entscheiden.

msg: Das hört sich ja gut an. Hatte diese Änderung auch Nachteile zur Folge?

Werner Keil: Ja, schon ein bisschen. Denn jetzt müssen alle Mitglieder zur Weiterentwicklung des Jakarta-EE-Standards beitragen. Man kann also nicht mehr sagen, dass Oracle und vielleicht auch noch IBM die Arbeit machen und die anderen höchstens mal alle paar Monate ihre Stimme abgeben. Sondern es ist sehr viel mehr aktive Gestaltung gefragt. Früher waren es ja vor allem zahlreiche Mitarbeiter von Oracle, die an der Programmierung beteiligt waren. Jetzt sind es gefühlt nur noch zwei bis vier Oracle-Mitarbeiter, die mitarbeiten.

msg: Sind auch Behörden in irgendeiner Weise an der Weiterentwicklung von Jakarta EE beteiligt?

Werner Keil: Ja, zumindest indirekt. So sind beispielsweise das Deutsche Zentrum für Luft- und Raumfahrt (DLR) und das Bundesinstitut für Risikobewertung (BfR) Eclipse-Mitglieder, ebenso die Deutsche Bahn oder die Schweizer Bundesbahnen (SBB), auch wenn diese derzeit wahrscheinlich nicht aktiv an Jakarta EE mitwirken. Und ich habe gerade heute mit den Eclipse-Vertretern über das von der EU gehostete und gesponserte DSS-Projekt¹ (Digital Signature Service) gesprochen. Aber in der Regel sind Behörden vorwiegend Nutzer.

JAKARTA EE – EIN SOLIDER INDUSTRIE-STANDARD FÜR BEHÖRDEN

msg: Für Behörden waren die standardisierten Java-Technologien stets von großer Bedeutung, denn ihre geschäftskritischen Unternehmensanwendungen müssen besonderen Sicherheitsanforderungen gerecht werden. Denkst Du, dass Jakarta EE immer noch den soliden Industriestandard darstellt, wie er von den Behörden für ihre unternehmenskritischen Geschäftsanwendungen benötigt wird?

Werner Keil: Ja, nun sogar umso mehr. Denn erst jetzt ist vorhersehbarer geworden, wohin sich der Standard bewegt.

Zum Beispiel wird jetzt niemand mehr aus wirtschaftlichen Gründen die Technologien in eine eigene Cloud befördern, nur um seinen eigenen Gewinn zu steigern.

msg: In der Vergangenheit wurden Java-EE-Server von den Großen der IT-Industrie zur Verfügung gestellt. Dadurch konnten sich Behörden darauf verlassen, dass sie ein solides Fundament vorfinden, auf dem sie ihre Anwendungen laufen lassen können. Ist das noch immer so, oder hat sich hieran etwas geändert?

Werner Keil: Hieran hat sich nichts geändert. Ganz im Gegenteil, denn es sind noch mehr Unternehmen wie Fujitsu und weitere Anbieter als solides Fundament hinzugekommen. Gerade aus dem asiatischen Raum kamen zuletzt besonders viele Jakarta-EE-Mitglieder dazu – neben Fujitsu TmaxSoft aus Südkorea sowie sechs Mitglieder aus China, darunter die beiden einzigen Enterprise-Mitglieder der Jakarta EE Working Group.



Fast alle alteingesessenen Hersteller haben ihre Produkte auch schon kurze Zeit nach der Ankündigung von Jakarta EE 8 zertifizieren lassen, so auch IBM WebSphere – jetzt Liberty – und JBoss beziehungsweise Wildfly. Mit Ausnahme von Tomitribe, wo man zwar oft etwas „laut“ ist, was aber verglichen mit Oracle, IBM oder Fujitsu bestenfalls ein KMU (kleine und mittlere Unternehmen) darstellt, haben alle Strategic Members der Jakarta EE Working Group ihre Produkte bereits für Jakarta EE 8 zertifizieren können.

Auch für Jakarta EE 9 ist mit IBM Open Liberty bereits der erste Server, abgesehen von GlassFish, als kompatibel zertifiziert. Andere hingegen sind etwas in den Hintergrund geraten. Beispielsweise ist GlassFish zwar die primäre kompatible Implementation. Aber dadurch, dass bei Jakarta EE der Begriff der Referenz-Implementation weggefallen ist, spielt GlassFish keine so eine große Rolle mehr. Einige Nutzer wie beispielsweise BMW sind dadurch zu Payara gewechselt. Oracle WebLogic ist inzwischen kompatibel mit Jakarta EE 8, aber bei Oracle ist zumindest ein gewisses Fragezeichen vorhanden, weil sie die

Technologien lieber in ihrer eigenen Cloud anbieten. Wobei Oracle hier allerdings weniger erfolgreich unterwegs ist als große Mitbewerber wie Microsoft, Google oder Amazon, die oftmals auch ihre eigenen OpenJDK-Distributionen bereitstellen, wie Coretto bei Amazon.

ÄLTERE, AUF JAVA EE BASIERENDE INSTALLATIONEN WERDEN IMMER WENIGER MIT JAKARTA EE KOMPATIBEL SEIN

msg: Seit vielen Jahren setzt sich Apache Tomcat in Kombination mit Spring und Hibernate in Behörden immer weiter durch. Genauso wird als Frontend immer mehr auf JavaScript-basierte Frameworks wie Angular gesetzt. Ist das nicht ein Trend, der die ganze Idee eines Industriestandards wie Java EE oder jetzt Jakarta EE infrage stellt?

Werner Keil: Es war ja ohnehin schon in der Vergangenheit so, dass Java-EE-Versionen stets neue Java-Server-Versionen mitbrachten und dies auch stets zu neuen Tomcat-Hauptversionen führte. Und Tomcat ist und war ja auch immer eine Implementierung zahlreicher Java-EE-Technologien. Somit sollte man Tomcat nicht als Konkurrent zu Java EE betrachten. Und dieser Anschluss zum Enterprise-Standard setzt sich auch mit Jakarta EE fort. So basieren auch die aktuelleren Tomcat-Versionen mehr und mehr auf Jakarta-EE-Versionen.

Ab Apache Tomcat Version 10 wird die Angleichung zum Jakarta Servlet und somit Jakarta-EE-Standard komplett sein. Ältere, auf Java EE basierende Installationen werden immer weniger mit Jakarta EE kompatibel sein. Dies betrifft im Übrigen auch Spring-Technologien wie Spring Boot, die sich ebenso heute schon auf Jakarta EE beziehen. Was Angular anbelangt, ist es schon so, dass Angular eine bedeutende Rolle spielt. Allerdings handelt es sich bei Angular ja nicht um eine Application-Server-basierte, sondern um eine clientseitige Technologie, die nicht ohne ein serverseitiges Pendant benutzt wird. Somit ist es zwar richtig, dass der Trend zu Angular Bestand hat, aber nicht als Konkurrenz gesehen werden kann, weil für das Backend stets ein Application Server eingesetzt werden muss.



**DIE HÖHERE VERBREITUNG VON SKRIPT-
SPRACHEN UND DIE INTERPRETATION
IHRES QUELLCODES ZUR LAUFZEIT MACHEN
DAS EINSCHMUGGELN BÖSARTIGER
CODESEQUENZEN LEICHTER.**

msg: Für mich ist eine Frage noch besonders interessant. Wenn Du jetzt eine Behörde beraten solltest: Welche Technologien würdest Du zur Nutzung empfehlen, wenn es sich um eine transaktionale, geschäftskritische, sichere Webanwendung handeln würde.

Werner Keil: Es gibt heutzutage immer mehr Programmiersprachen und Technologien, die für sich in Anspruch nehmen, dass mit ihnen geschäftskritische Webanwendungen und APIs besonders gut und effizient entwickelt werden können. Auch durch die meist steilen Lernkurven sind dabei Skriptsprachen wie JavaScript beziehungsweise TypeScript oder Python besonders populär. Die steile Lernkurve und höhere Verbreitung sowie die Tatsache, dass die Skriptsprachen per Definition eben nicht in ein binäres Artefakt wie eine ausführbare Datei oder ein Archiv münden, sondern der Quellcode zur Laufzeit interpretiert wird, machen das Einschmuggeln böstiger Codesequenzen dort leichter.

Die Verbreitung von Angular oder NodeJS gerade auch im unternehmenskritischen Umfeld führt immer öfter zu Angriffen oder der Ausnutzung von Sicherheitslücken in diesem Bereich. Der Node Package Manager (npm), De-facto-Standard für das JavaScript-Ökosystem, war in den vergangenen ein bis zwei Jahren mehrfach Ziel von Angriffen, die ich beispielsweise im Gespräch mit Speaker-Kollegen auf der Developer Week bereits ein paar Jahre früher vorhergesehen habe, als sie mir damals erzählten, dass die Pakete bei npm ohne jede Authentifizierung anonym hochgeladen und beliebig oft ersetzt werden könnten.

Das mag sich seither gebessert haben, aber dank sogenannter „Dependency

Confusion“ konnte sich ein Sicherheitsexperte, wie eben erst Anfang Februar berichtet², in einige der größten IT- und Cloud-Unternehmen weltweit, wie Apple, Microsoft, PayPal, Netflix oder Uber, hacken. Währenddessen setzen die binären Repositories des De-facto-Standards für die Java Virtual Machine, Sonatype Nexus beziehungsweise MavenCentral, seit Jahren auf Zugangsdaten und digitale Signaturen zum Schutz ihrer Artefakte. Und selbst das Deployment eines Pakets in einem neuen oder unbekanntem Namensraum ist nicht ohne vorherige Genehmigung durch die Urheber der jeweiligen Projekte, beispielsweise den Domain-Eigentümer, möglich.

Auch wenn Java und Java EE beziehungsweise Jakarta EE schon mehrfach totgesagt wurden, halte ich auch aus diesem Grund beide für eine solidere Basis, um geschäftskritische, transaktionale und vor allem sichere Webanwendungen zu entwickeln. Es spielt dabei keine so große Rolle, ob man auf Java, Scala, Kotlin oder Clojure schwört und welches Framework letztlich zum Einsatz kommt – ob „pure“ Jakarta EE oder darauf basierende Lösungen wie Spring Boot, Eclipse MicroProfile, Micronaut oder sonstige – solange sie auf der JVM laufen, was dank GraalVM übrigens auch für die meisten Skriptsprachen heute möglich ist.

msg: Welche Rolle wird eine Modularisierung von Jakarta EE in Zukunft spielen?

Werner Keil: Bereits mit Jakarta EE 9 habe ich im Specification Committee geholfen, gut zwei Drittel der Jakarta-EE-Spezifikation-

APIs mit JPMS-Moduldefinitionen (seit Java 9 „Jigsaw“ Modulsystem) zu versehen. Daneben enthalten auch etliche eine OSGi-Bundle-Deklaration. Da die beiden Systeme konkurrieren und manche Committer insbesondere aus dem Apache-Lagerhistorisch bedingt ihre API JARs wiederpacken (seit sie mit Apache Harmony lange Zeit auf Konfrontationskurs mit dem JCP und Sun beziehungsweise Oracle waren), statt auf die offiziellen Java EE oder Jakarta EE JARs zu setzen, ist das nicht immer so einfach.



ZUR PERSON

WERNER KEIL ist Cloud Architekt, API-Designer, Java- und Microservice-Berater für die unterschiedlichsten Branchen. Er entwickelt Enterprise-Systeme mithilfe von Java, Java EE, Jakarta EE, Oracle, IBM oder Microsoft, betreibt Webentwicklung mit JavaScript und anderen dynamischen oder funktionalen Sprachen. Er ist seit über dreißig Jahren als Projektleiter, Software-Architekt, Analyst und Technologie-Berater für führende Unternehmen tätig.

Werner Keil ist unter anderem Committer in der Apache und der Eclipse Foundation, Committer Vertreter im Jakarta EE Specification Committee sowie aktives Mitglied der Java Community Process.

Neben seiner Tätigkeit für große Unternehmen leitet oder unterstützt Werner Keil zahlreiche Open-Source-Projekte, schreibt Songtexte, Romane, technische Bücher und Artikel. Er ist ein gefragter Speaker auf Konferenzen weltweit.

Spätestens mit dem Java-Modulsystem wurde aber vom JDK eine gewisse Grundlage bereitgestellt, von dem nicht nur Oracle selbst in seinem Helidon Microservice Framework intensiven Gebrauch macht. Auch die erste zertifizierte Implementierung einer einzelnen Jakarta-EE-Spezifikation (JSON Processing) „Joy“ setzt ganz auf Java 14 und höher mit JPMS.

msg: Ich habe kürzlich davon gehört, dass Du zusammen mit anderen Jakarta-EE-Experten an einem Buch schreibst. Worum geht es dabei?

Werner Keil: Das Buch „The Definitive Guide to Jakarta EE Security“ im Apress Verlag, an dem ich zusammen mit dem Leiter der Utrecht Java User Group Thodoris Bais und Java Champion Arjan Tijms arbeite, ist für Frühjahr 2021 geplant. Wie der Titel

erahnen lässt, deckt es alle Facetten und wesentlichen Technologien zum Thema Sicherheit auf Jakarta-EE-Basis ab – angefangen bei Java-Security-Grundlagen, digitaler Signatur oder der Java Cryptography Extension über JSON Web Token (JWT), OAuth beziehungsweise OpenID Connect bis zu den Jakarta-Sicherheitsstandards Authentication, Authorization und Jakarta Security API. Wen das Thema interessiert, der kann sehr gerne dem entsprechenden Twitter-Account³ folgen, dort finden sich die neuesten Ankündigungen und, sobald Apress die Buch-Website freigeschaltet hat, auch Links darauf.

msg: Vielen Dank, Werner, für das sehr informative Interview.

Werner Keil: Vielen Dank für die Einladung zu diesem Interview. ●



msg: Bist Du immer noch als Speaker so aktiv wie in der Vergangenheit? Welche IT-Konferenzen würdest Du unseren Lesern empfehlen?

Werner Keil: Ich wurde in den letzten zehn bis fünfzehn Jahren auf so gut wie alle größeren Entwickler- und IT-Konferenzen eingeladen. Mit Ausnahme von Australien/Ozeanien auf jedem Kontinent und auf alle größeren Veranstaltungen, die 2000 oder mehr Besucher anziehen, in der D-A-CH-Region. Deshalb wäre es fast etwas unfair oder auch zu umfangreich, hier alle zu erwähnen oder einzelne hervorzuheben. Ich möchte dennoch, mit Schwerpunkt auf Mitteleuropa, einige erwähnen, abhängig von der jeweiligen Plattform und Thematik. Wer großteils auf der Java Virtual Machine (JVM) zu Hause ist, der kommt an Devovx (ehemals JavaPolis) in Antwerpen oder Javaland in Brühl bei Köln kaum vorbei. Devovx hat seit einigen Jahren auch mehrere Ableger in England, Frankreich oder Osteuropa sowie kleinere – Vovxed Days genannte – meist eintägige Gastkonferenzen in anderen Städten und Ländern. Das JavaLand ist zwar noch nicht ganz so groß wie die größten Devovx-Events, die meist um die 4.000 Besucher anziehen, wird aber – mit Ausnahme von 2020 – jedes Jahr etwas größer. Zuletzt kamen 2019 wohl schon über 2.000 Besucher ins Phantasia-land. Im Coronajahr 2020 wurden bis auf die Devovx Ukraine online fast alle anderen Konferenzen abgesagt. In diesem Jahr findet das JavaLand corona-bedingt online statt. Welchen Einfluss das auf die Teilnehmerzahlen hat, bleibt abzuwarten. Aber da Anreise und Hotel wegfallen, könnte es Manchen die Teilnahme sogar erleichtern. In deren Schatten ist die JCON

in Düsseldorf, die sich in einem großen Kino vom Ambiente her an der belgischen Devovx orientiert, in den letzten Jahren auch beständig gewachsen. Und konnte trotz oder gerade wegen Corona 2020 als Onlineevent mit mehr Besuchern als je zuvor punkten. Die 2.000er-Marke hat sie ebenfalls überschritten.



Abseits der JVM ist die Codemotion-Serie ebenfalls zu empfehlen. Sie ist mehr polyglott und deckt alle wesentlichen Plattformen von JavaScript/NodeJS über Python oder .NET bis zu JVM-Sprachen ab. Sie hat in Rom den größten Standort mit mehr als 4.000 Besuchern. Nach weitgehenden Absagen finden alle Codemotion-Konferenzen nun wohl online statt. Ansonsten gab es unter anderem in Berlin meist eine etwas kleinere Codemotion. Ähnlich polyglott, mit mehr als 2.500 Besuchern in einem normalen Jahr, sind die Developer Week (DWX) in Nürnberg oder die OOP in München. Die IT-Tage in Frankfurt waren mit 1.000 Besuchern vor Ort etwas kleiner. Sie fanden 2020 auch virtuell statt und sind für Ende 2021 geplant. Ich kenne hier keine Teilnehmerstatistik, aber ähnlich wie bei der JCON dürfte sich das Ausweichen in die Cloud kaum negativ auf die Teilnehmerzahl ausgewirkt haben.

1 <https://ec.europa.eu/cefdigital/DSS/webapp-demo> (abgerufen am 29.03.2021).

2 <https://medium.com/@alex.birsan/dependency-confusion-4a5d60fec610> (abgerufen am 05.03.2021).

3 <https://twitter.com/jakartasecbook/> (abgerufen am 05.03.2021).

KÖNNEN ALGORITHMEN DISKRIMINIEREN?

| von WERNER ACHTERT, Geschäftsleitung Public Sector, msg

Künstliche Intelligenz (KI) hat das Potenzial, Einfluss auf das alltägliche Leben vieler Menschen zu nehmen. Deshalb findet aktuell abseits der technischen Diskussionen ein breiter gesellschaftlicher Diskurs über die Notwendigkeit einer stärkeren Transparenz beim Einsatz von KI-Systemen statt (siehe auch Seite 16 ff.).

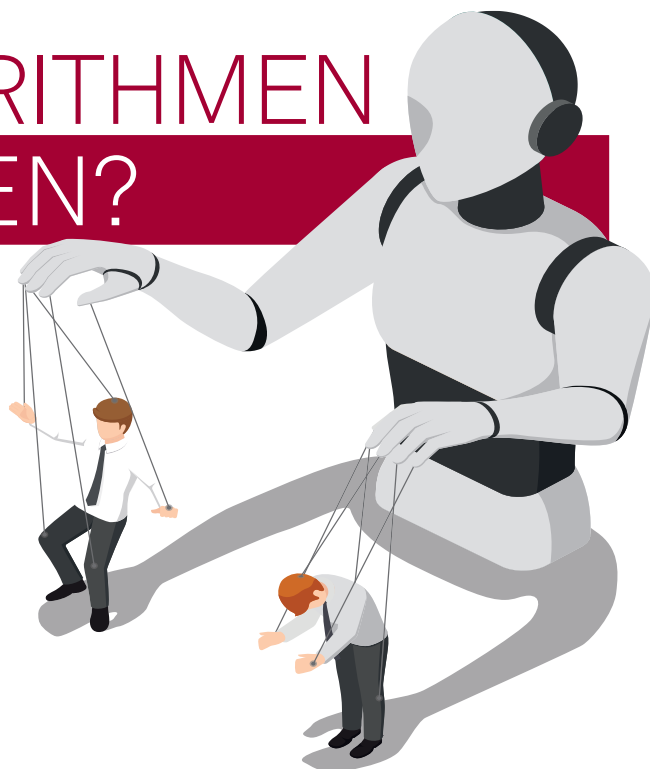
Ein zentrales Thema dabei ist der Vorwurf, KI-Systeme würden einzelne Personengruppen diskriminieren, indem zum Beispiel dunkelhäutige Menschen schlechter erkannt werden, Bewohner bestimmter Stadtteile bei Versandhäusern nur gegen Vorkasse bestellen können und Frauen bei Einstellungen schlechter bezahlt werden.

Dabei machen KI-Systeme eigentlich nur eines: Sie werten Daten der Vergangenheit aus und ziehen daraus – sehr vereinfacht ausgedrückt – durch Korrelationsanalyse Rückschlüsse zur Bearbeitung von weiteren Daten.

Was KI-Systeme nicht können, ist, Kausalzusammenhänge zu erkennen. Ein KI-System kann nicht „erklären“, warum zwei Mitarbeiterinnen oder Mitarbeiter bei vergleichbaren Tätigkeiten unterschiedliche Gehälter bekommen und schon gar nicht, ob das gerechtfertigt ist.

Wir dürfen bei der Diskussion über die Anwendung von KI nicht Ursache und Wirkung verwechseln. KI-Systeme schaffen keine gesellschaftlichen Probleme, sie können allerdings bestehende Ungerechtigkeiten, Benachteiligungen etc. offensichtlich machen. Der Einsatz von KI-Systemen birgt das Risiko, solche Verzerrungen zu verfestigen, aber auch die Chance, sie zu erkennen und zu beseitigen.

Gesellschaftliche Probleme können durch KI-Systeme nicht aus der Welt geschafft werden – sie können nur gesellschaftlich (also politisch) gelöst werden. In unserem Rechtssystem sind bereits Regeln zur Vermeidung von Diskriminierung vorhanden, wie beispielsweise das Allgemeine Gleichbehandlungsgesetz



(AGG). Diese können auf manuelle Entscheidungsprozesse genauso wie auf KI-Systeme angewendet werden.

Nicht das KI-System diskriminiert also, sondern die Auswahl der verwendeten Trainingsdaten, der Klassifikationen und der einer Bewertung zugrunde liegender Attribute (siehe auch Seite 22 ff.).

Was wir brauchen, sind mehr Sachverstand und gesellschaftliches Bewusstsein bei der Auswahl der Daten, mit denen wir KI-Systeme trainieren. Ein KI-System kann beispielsweise aus der Menge der eingegangenen Bewerbungen eine Auswahl für persönliche Einladungen treffen. Um Chancengleichheit zu gewährleisten, müssen in diesem Fall die Trainingsdaten ausgewogen in Bezug auf beispielsweise Alter und Geschlecht sein.

Wir müssen durch kontinuierliches Qualitätsmanagement, also zum Beispiel dem Hinterfragen der durch die KI getroffenen Entscheidungen, sicherstellen, dass die Trainingsdaten für den Einsatzzweck signifikant sind und keine ungewollten (oder gar gewollten) Verzerrungen enthalten. Und dies muss transparent gemacht werden, um bei den Betroffenen in der Öffentlichkeit Vertrauen zu schaffen.

Und dafür bedarf es – nach wie vor – menschlicher Intelligenz. ●



Grundlagen der Spracherkennung durch KI

| von **JOHN LOUTZENHISER**

Indem Sie diesen Artikel lesen, erlauben Sie, liebe Leser, mir, dem Autor, Eindrücke, Ideen und Bilder in Ihrem Kopf entstehen zu lassen. Sie gewähren mir praktisch unbegrenzten Einfluss auf Ihr Inneres, Ihr geistiges Leben – zumindest so lange, wie Sie für diese Lektüre brauchen. Wann immer wir lesen, sprechen oder zuhören, gestehen wir Anderen diese Macht zu – oder üben sie aus. Und die allermeisten von uns halten zu keinem Zeitpunkt inne, um über dieses alltägliche, seltsame und wunderbare Instrument nachzudenken, das uns so viel Macht gibt – die menschliche Sprache. Sprache ist für uns selbstverständlich. Vielleicht, weil Sprache zutiefst menschlich und nur der menschlichen Spezies zu eigen ist. Und weil sie ein so wesentlicher Bestandteil unseres Daseins ist, ist es schwierig Abstand, zu nehmen und sie als etwas von uns Getrenntes zu untersuchen.

Sprache erweist sich bei näherer Betrachtung als schwer durchschaubar und geheimnisvoll. Ihre Eigenschaften sind einzigartig und vielfältig. Als exakter Forschungsgegenstand ist sie daher kaum geeignet.

Doch woher kommt Sprache? Hat der Mensch sie erfunden, wie das Rad oder das Internet? Ist sie ein kulturelles oder ein soziologisches Phänomen? Oder eher ein natürliches, biologisches? Ein Teil von uns, wie das Herz, die Augen oder die Daumen? In welchem Zusammenhang stehen Sprache und Denken? Klären wir unsere Ideen, Absichten und Meinungen für uns selbst bereits in sprachlicher Form, oder denken wir ohne Sprache und nutzen sie erst im zweiten Schritt als Werkzeug, um uns mitzuteilen?

Die Antwort lautet: Sprache ist all dies zusammen: menschliche Erfindung ebenso wie ein natürliches Phänomen. Sie existiert in unserer innersten individuellen Gedankenwelt. Und wir nutzen sie, um mit Anderen, um mit der Außenwelt in Beziehung zu treten.

„DIGITALE UNENDLICHKEIT“

Sprache ist all dies und auch noch etwas gänzlich anderes. Sprache gehört zum Reich der Information. Sie weist eine komplexe Struktur aus Signalen, Symbolen, Regeln und Hierarchien auf – auf den ersten Blick ist die „reine“ Sprache der Logik und Mathematik nicht unähnlich.

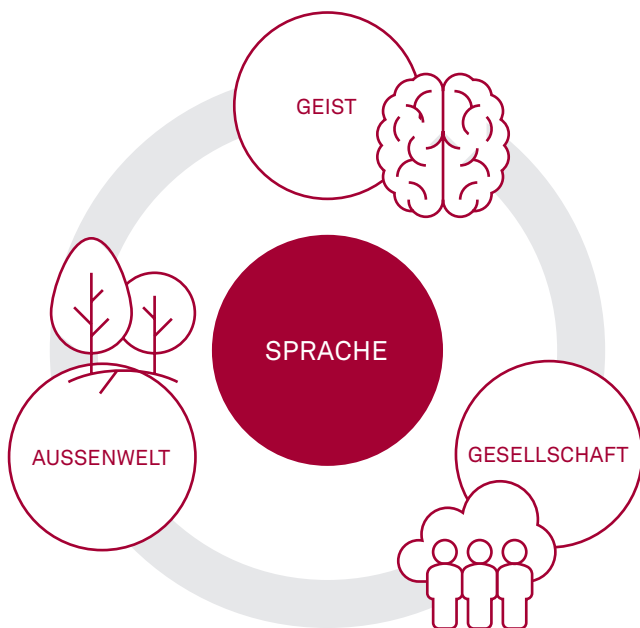


Abbildung 1: Sprache existiert in unserem Kopf, bezieht sich auf die Außenwelt und ist Werkzeug für Kommunikation. Sie ist ein System, das Außenwelt, Gesellschaft und Geist verbindet und von allen drei geformt wird.



„Sprache ist im Grunde ein System, das sowohl digital als auch unendlich ist. Soweit ich weiß, gibt es kein anderes biologisches System mit diesen Eigenschaften.“

Noam Chomsky¹

Sprache kann eine unendliche Bandbreite unterschiedlichster Erfahrungen darstellen und „kodieren“: unsere Wahrnehmung der äußeren Welt, unsere inneren mentalen, emotionalen und kreativen Erfahrungen, soziale Interaktion und Kommunikation. Dafür verwendet sie einen erstaunlich effizienten und kompakten Code – die Laute und Striche, aus denen gesprochene und geschriebene Sprache besteht. Uns stehen nur wenige Dutzend Schriftzeichen zur Verfügung, um sie zu Worten unseres großen, aber dennoch begrenzten Wortschatzes zusammenzufügen. Wir verbinden diese Worte und Formen aus ihnen potenziell unendlich viele verständliche Ausdrücke. So gesehen macht Sprache „unendlichen Gebrauch endlicher Mittel“.

SPRACHE UND KÜNSTLICHE INTELLIGENZ

Diese unbegrenzte expressive Kraft der Sprache und ihre kompakte Darstellungsweise haben sie zu einem fesselnden Untersuchungsgegenstand in der Erforschung künstlicher Intelligenz gemacht – und zwar seit den Anfängen dieser Disziplin vor gut siebenzig Jahren. Damals wie heute war und ist Sprache der naheliegende, ebenso intuitive wie verführerische Untersuchungsgegenstand: Gelingt es, den Computer „ganz einfach“ zu lehren, die Zeichen und Regeln der Sprache in einer ähnlichen Weise zu handhaben (zu verstehen, einzusetzen), wie wir es tun? Dann wäre er „intelligent“ – oder würde uns zumindest Intelligenz vorgaukeln. Dies ist bekanntlich die Annahme hinter dem Turing-Test², einem der frühesten und wichtigsten Impulse in der Geschichte der künstlichen Intelligenz. Einem Computer beizubringen, Sprache zu verwenden, hat sich inzwischen als eine weit größere Herausforderung erwiesen als in den Anfangstagen der KI angenommen. Trotz vieler entscheidender Durchbrüche gibt es auch heute noch keine Software, die Sprache in überzeugender Weise „intelligent“ nutzt, zumindest nicht in dem Sinne, als dass sie den Turing-Test bestehen würde.

Nichtsdestotrotz: Die seit rund zehn Jahren herrschende neue Begeisterung für künstliche Intelligenz geht ganz wesentlich auf jüngere bahnbrechende Fortschritte im Feld der künstlichen Intelligenz beziehungsweise in der „Verarbeitung natürlicher Sprache“, wie es inzwischen heißt (engl.: Natural Language Processing, kurz: NLP), zurück. Diese signifikanten Fortschritte öffnen viele praktische Anwendungsmöglichkeiten für sprachverarbeitende Software.

EINE PRAGMATISCHE WENDE

Die Verarbeitung natürlicher Sprache ist eine Disziplin der Ingenieurwissenschaften und beschäftigt sich mit der Erstellung von „nützlichen“ oder in irgendeiner Weise „wertschöpfenden“ Artefakten. Dennoch: Ingenieurskunst und Naturwissenschaft gehören eng zusammen. Der Bau von Brücken oder die Raumfahrt beispielsweise erfordern nicht nur einen wissenschaftlichen Hintergrund, sondern bestätigen ihrerseits wissenschaftliche Erkenntnisse und bringen sie weiter voran. Das gilt auch für NLP: Das Bestreben, Software zu entwickeln, die mit einer „intelligenten“ Sprachverarbeitung Nutzen und Mehrwerte schafft, hat auch zu Erkenntnissen über das Wesen und die Mechanismen menschlichen Denkens und die menschliche Sprachfähigkeit geführt.

Den Fokus weniger auf theoretische Eleganz zu legen als vielmehr auf den pragmatischen Wert und Nutzen, hat ebenso im Feld der Linguistik und Sprachphilosophie an Bedeutung gewonnen. In weiten Teilen des 20. Jahrhunderts hat sich Linguistik mit formalen, logischen und symbolischen Zugängen zu Syntax (Grammatik) und Semantik (Bedeutung) beschäftigt – in der Hoffnung, eine „reine“ und nachvollziehbar adäquate Theorie der Sprache hervorzubringen.

Die Niederungen alltäglichen Sprachgebrauchs, der Einfluss linguistischen und realen Kontextes auf die Bedeutung von Wörtern und Sätzen, der Einfluss von Sprache auf Andere im Verlauf der Kommunikation waren in Randgebiete der Linguistik verbannt – in die Unterdisziplin einer „pragmatischen Linguistik“. Heute ist die Pragmatik kein Randgebiet mehr, sondern von zentraler Bedeutung für NLP.

In diesem Artikel stehen zwei der maßgeblichen Ideen der linguistischen Pragmatik und Philosophie der Sprache im Fokus: Kontext und Verwendung.

KONTEXT IST KÖNIG

Kontext ist König: Dieser Gedanke aus der pragmatischen Schule der Linguistik und Sprachphilosophie hat NLP – bezogen auf die Bedeutung von Wörtern und Sätzen – am stärksten geprägt. Er ist im berühmten Zitat von Ludwig Wittgenstein zusammengefasst: „Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.“⁴³ Der Linguist J. R. Firth formuliert das so: „Die vollständige Bedeutung eines Wortes ist immer kontextabhängig, und keine kontextunabhängige Untersuchungen über Bedeutungen kann ernst genommen werden.“ Oder noch schärfer: „Man erkennt ein Wort an seinen Begleitern.“⁴⁴

Aus Sicht dieser „Kontextualisten“ sind Wörter nur im Kontext verständlich. Aber wieso? Bedeuten sie nicht auch an sich etwas, unabhängig vom Zusammenhang? Stimmt es nicht auch, dass Sprache und Wörter auf Dinge in der echten Welt „verweisen“ oder „für etwas stehen“? Das Wort „Bedeutung“ impliziert genau dieses Sprachverständnis: „auf etwas deuten“. Das Wort „Tür“ ist in einer befriedigenden Weise ganz einfach ein Symbol, das auf ein Ding in der Welt hinweist, nämlich auf eine Tür. Aber ist das wirklich so einfach? Welche Tür genau? Das lässt sich nur beantworten, wenn es aus dem Zusammenhang heraus, sei er visuell oder verbal, Hinweise auf eine spezifische Tür gibt. Und was genau bedeutet „Tür“ im Satz: „Den Bürgermeister getroffen zu haben, hat uns Türen geöffnet“? Ist diese Bedeutung von „Tür“ in irgendeiner Weise weniger wirklich als die im Satz „Die Haustür ist offen“?

Offensichtlich gibt es Bezeichnungen mit eindeutigen, feststehenden Bedeutungen. Eigennamen sind ein Beispiel. Es wird sich schwerlich ein Kontext finden, in dem in „München“ eine „Brooklyn Bridge“ steht. Aber wenn wir Bedeutung ernst nehmen, so wie wir es tun müssen, wenn wir moderne Systeme zur Verarbeitung natürlicher Sprache entwickeln wollen, dann erkennen wir, dass die Bedeutung von Wörtern eine komplexe und unscharfe Angelegenheit ist und sich nicht adäquat über diskrete, also eine begrenzte Anzahl eindeutiger, unterscheidbarer Symbole abbilden lässt. Das Ding, auf das ein Wort vielleicht hinweist, ist nur ein kleiner Aspekt seiner Bedeutung.

Wörter haben komplexe, vieldimensionale Schattierungen und Aspekte. Wir Sprachnutzer spielen damit in kreativer Weise, indem wir eine Teilbedeutung auf neue Situationen anwenden und damit neue, abgeleitete, kontextabhängige Bedeutungen generieren. Deshalb kann „Tür“ ebenso „Gelegenheit“ oder „Möglichkeit“ bedeuten wie das hölzerne Ding mit Angeln (auch so ein Wort) in unserem Haus – immer abhängig vom Kontext.

DIE „EINBETTUNG“ ALS DARSTELLUNGSMÖGLICHKEIT VON BEDEUTUNG

Wenn also die Symbole alleine nicht zur Beschreibung der Bedeutung von Wörtern taugen, wie können wir Bedeutung dann fassen, um bessere NLP-Systeme zu entwickeln? Einer der wichtigsten Durchbrüche sind die „Worteinbettungen“ (word embeddings), das heißt die kontextabhängige und mehrdimensionale Kodierung von Wörtern. Worteinbettungen ermöglichen einen kontextbezogenen Zugang zu Sprache, so wie oben beschrieben („Sie werden ein Wort an seinen Begleitern erkennen“).

```

text_classification_demo > wordvectors > glove.6B.300d.w2v.txt
cnn_text_classification_demo.py | spacy_test.py | glove_fasttext_mixed.py | codecs.py | glove_demo.py | glove.6B.300d.w2v.txt
The file is too large: 1.04G. Showing a read-only preview of the first 2.56M.
277 help -0.080441 0.072276 -0.47046 -0.39104 -0.0016158 0.067585 0.17371 -0.010269 0.34673 -1.8027 0.28699 0.10734 0.18789 -0.23229 0.077246 0.071
278 chief 0.070806 -0.61458 0.054454 -0.55143 0.2895 -0.88188 -0.32588 0.59543 -0.1193 -1.5105 0.040473 -0.41871 -0.29206 0.018349 -0.13419 0.91431
279 saturday 0.17015 0.25273 -0.2255 0.13639 -0.17897 -0.073514 0.0376 0.05338 0.068511 -0.85258 0.069401 -0.373 -0.25872 0.19209 0.042546 0.5245 +
280 system -0.20788 -0.077677 0.23521 -0.21965 0.29176 0.031269 0.55304 0.37834 0.32832 -2.6476 0.15075 0.5719 -0.16453 -0.35184 0.58232 0.2604 -0.
281 john 0.058563 0.093696 -0.16048 -0.16589 -0.1032 0.016987 -0.15772 -0.26285 -0.40447 -1.1235 0.38675 -0.44901 -0.026734 0.059817 0.012804 0.411
282 support -0.28085 -0.23965 -0.13137 -0.62986 0.032218 0.28209 0.21912 -0.25771 -0.12819 -2.1128 0.32902 -0.024806 0.24955 0.14335 0.22002 0.3741
283 series 0.22805 0.56135 0.15447 -0.60487 -0.24785 0.61739 -1.1351 0.29129 0.13176 -1.0554 0.53971 -0.12311 -0.24304 -0.13305 0.04723 0.23786 -0.
284 play -0.36011 0.61678 -0.48175 -0.066795 -0.0252 0.18144 -0.27854 0.31219 0.065527 -0.82089 0.15012 0.091332 -0.11295 -0.19711 0.096786 -0.2584
285 office -0.36322 0.049952 -0.31563 -0.039297 0.10458 0.027566 -0.26305 -0.12784 -0.40336 -1.7768 -0.15068 0.28388 0.085887 -0.10529 0.0066189 0.
286 following 0.074723 -0.040008 0.44226 -0.12438 0.26759 -0.05802 -0.40286 -0.00021706 0.23837 -1.5204 -0.19551 0.31701 -0.062581 -0.065533 0.2294

```

Abbildung 2: Ausschnitt aus einer Worteinbettungsdatei. Jede Zeile ist ein Embedding für ein Wort. Jedes Embedding hat 300 Zahlen (Dimensionen), von denen hier nur die ersten fünfzehn zu sehen sind.

Eine Worteinbettung „kodiert“ in gewisser Weise ein Wort inklusive seiner Beziehungen mit und Abhängigkeiten von anderen Wörtern als „Vektor“, das heißt als Zahlenfolge aus Dezimalzahlen.

Die Größe des Vektors spiegelt die Anzahl der Dimensionen der Worteinbettung wider. Jede einzelne dieser Dimensionen, so kann man sich das vorstellen, entspricht einem spezifischen Aspekt oder einer Schattierung der Wortbedeutung im Verhältnis zu anderen Wörtern. Im Ergebnis erscheinen Wörter, die in ähnlichen Zusammenhängen auftauchen und die ähnliche Bedeutungen haben, in Gruppen im Vektorraum, zumindest entlang einer oder mehrerer Dimensionen. Eine typische Größe für den Vektor ist in aktuellen KI-Systemen „300“ und wird vom Ingenieur festgelegt, der das KI-System trainiert.

PIZZA – ITALIEN + BERLIN = CURRYWURST

Worteinbettungen haben überraschende und nützliche – geradezu magische – Eigenschaften. Nicht nur erscheinen ähnliche Wörter in Gruppen: Wenn man Wortvektoren hinzufügt oder streicht, macht der sich daraus ergebende Vektor oft abstrakte linguistische Ideen oder sogenanntes „Weltwissen“ offenbar. Das folgende Beispiel verdeutlicht das:

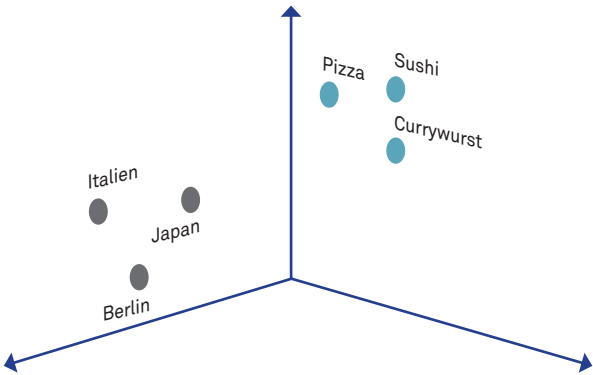


Abbildung 3: Word embeddings – ähnliche Wörter formen „Cluster.“

Diese Beispiele sind keine Spielerei, sondern echt. Wenn man eine große Sammlung von Worteinbettungen trainiert, erhält man Vektorrechnungen wie:

$$v_{\text{Pizza}} - v_{\text{Italien}} + v_{\text{Berlin}} = v_{\text{Currywurst}}$$

Das Ergebnis beruht darauf, dass der Vektor zwischen Pizza und Italien derselbe Vektor ist (mit derselben Richtung und Länge) wie der Vektor zwischen Berlin und Currywurst. Der Vektor scheint also die Vorstellung „typisches Essen“ darzustellen.

Ebenso ist der Vektor zwischen Toyota und Japan derselbe wie der zwischen BMW und München. Dieser Vektor scheint also „wichtigster Autohersteller“ zu beinhalten.

$$v_{\text{Toyota}} - v_{\text{Japan}} + v_{\text{München}} = v_{\text{BMW}}$$

WORTEINBETTUNGEN TRAINIEREN – ODER „THE MASKED LANGUAGE MODEL“

Natürlich werden Worteinbettungen nicht von Hand eingegeben oder einfach erfunden und niedergeschrieben. Sie sind das Ergebnis von maschinellem Lernen, in der Regel eines neuronalen Netzes, das eine große Sprachsammlung aus Millionen von Sätzen nutzt, beispielsweise alle Einträge in Wikipedia. Der Trainingsprozess lässt sich als eine „Black Box“ verstehen, die mit Text gefüttert wird und für jedes einzelne Wort darin einen Wortvektor erstellt.

Bevor man Worteinbettungen mit einem neuronalen Netz trainiert, muss man erst wissen, wofür man es trainiert. Das bedeutet, dass man

1. dem Netz eine bestimmte Aufgabe stellen muss,
2. es mit den für die Lösung dieser Aufgabe relevanten Daten füttern und dann
3. die Fortschritte hin zu einer erfolgreichen Lösung der Aufgabe messen sowie es
4. fortlaufend entsprechend anpassen lassen muss.

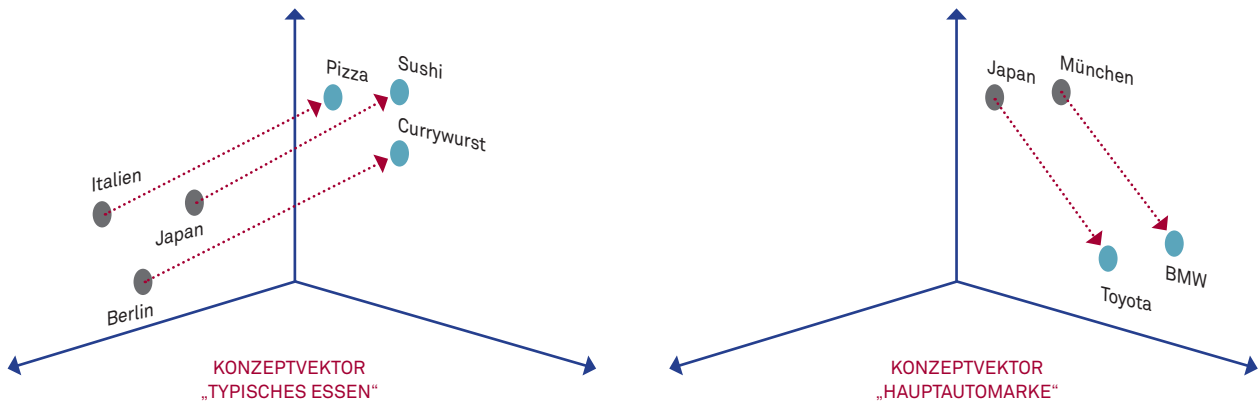


Abbildung 4: Konzeptvektoren in Worteinbettungen

EIN RÄTSEL ZEIGT, WIE ES GEHT

Im Fall des neuronalen Netzes für Worteinbettungen ist die gestellte Aufgabe ein einfaches Rätsel: „Finde das Wort heraus, das im Satz fehlt.“

Dieses Rätsel ist so einfach wie mächtig. Um zu verstehen, wieso diese einfache Aufgabe ausreicht, um anspruchsvolle Worteinbettungen („Sprachmodelle“ oder „Language Model“) zu trainieren, können wir dieses Spiel selber spielen.

Lesen Sie bitte den folgenden Satz:

„Auch im dritten Anlauf hat die britische Regierungschefin Theresa May das mit der Europäischen Union _____ Ausstiegsabkommen nicht über die Ziellinie gebracht.“

Ein Wort fehlt. Um welches Wort handelt es sich? Denken Sie kurz darüber nach, bevor Sie eine Auswahl aus den folgenden

Begriffen treffen: a) *abgelehnte* b) *komplizierte* c) *vereinbartem* d) *vereinbarte*

Wir können davon ausgehen, dass alle, die der deutschen Sprache mächtig sind, das gleiche Wort gewählt haben. Welches haben Sie gewählt? Und wieso? Jetzt halten Sie bitte kurz inne und versuchen, Ihre eigene Sprachkompetenz unter die Lupe zu nehmen. Wieso haben Sie das richtige Wort ausgewählt und nicht eines der anderen?

Und jetzt lesen Sie folgenden Satz (Achtung, es ist nicht genau der gleiche Satz wie oben, also aufmerksam lesen!)

„Auch im dritten Anlauf hat die britische Regierungschefin Theresa May das von der Europäischen Union _____ Ausstiegsabkommen nicht über die Ziellinie gebracht.“

Um welches Wort handelt es sich? a) *vorgeschlagene* b) *komplizierte* c) *vereinbartem* d) *vereinbarte*

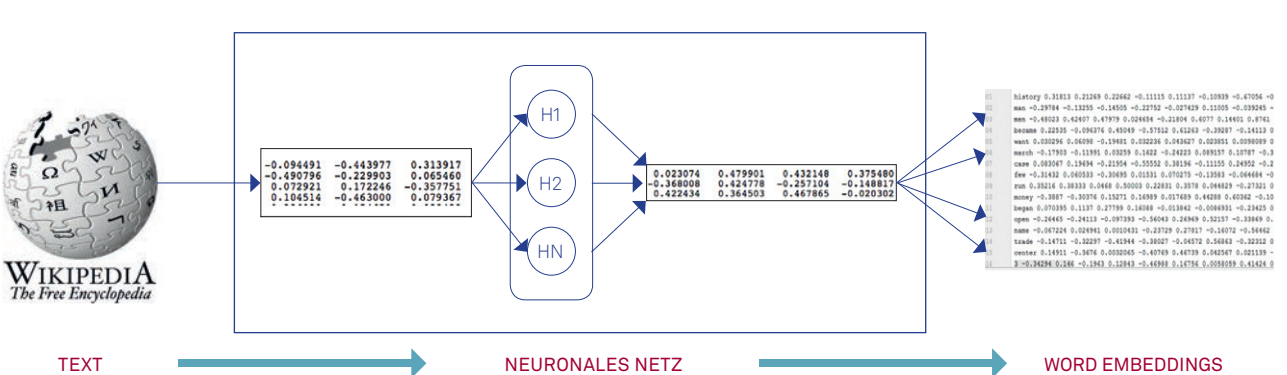


Abbildung 5: Durch ein neuronales Netz erzeugte Worteinbettungen der gesammelten Wikipedia-Texte

Der Encoder „kodiert“ Wörter in Embeddings und gibt die an Decoder weiter

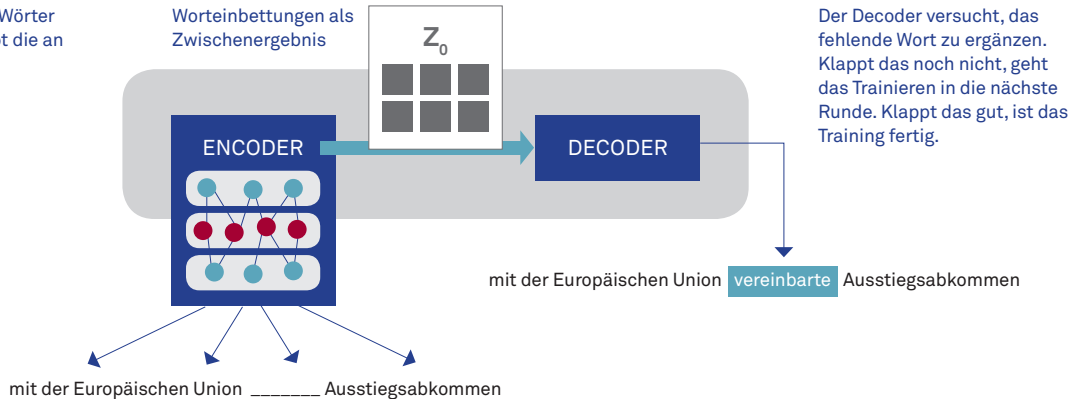


Abbildung 6: Schematische Darstellung des Zusammenspiels von Encoder und Decoder in einem Transformer

Haben Sie das gleiche Wort gewählt? Oder ein anderes? Und wieso? Wir stellen fest, dass wir auf sehr viele Informationen aus dem Gesamtkontext des Satzes gleichzeitig zurückgreifen, um das fehlende Wort richtig zu ergänzen. Weltwissen, Satz- und Wort-Bedeutung (Semantik), Satzbau und Wortformen (Grammatik, Syntax) – all das fließt in unsere Entscheidung ein, und zwar ohne, dass es uns bewusst ist.

Ebenso, wie wir sehr viel Kontext brauchen, um das fehlende Wort richtig zu ergänzen, ist im Umkehrschluss die Bedeutung von Wörtern extrem kontextabhängig (wie oben bereits erläutert). Worteinbettungen sind dafür die gängige NLP-Methode, und um Worteinbettungen zu erzeugen, trainiert man ein neuronales Netz darauf, genau dieses Rätsel, „Finde das Wort heraus, das im Satz fehlt“, zu lösen.

Um das zu erreichen, benötigt man sehr viele Sätze – in der Regel mehrere Millionen – und „maskiert“ oder „versteckt“ beispielsweise fünfzehn Prozent der Wörter in den Sätzen. Anschließend stellt man einem neuronalen Netz die Aufgabe, diese fehlenden 15 Prozent der Wörter richtig zu ergänzen.

Eine moderne und gängige Netzwerkarchitektur, die dazu verwendet wird, State-of-the-Art-Ergebnisse beim Lösen dieses Rätsels zu erzielen, heißt Transformer. Ein Transformer besteht, einfach ausgedrückt, aus zwei Teilen: einem Encoder und einem Decoder. Der Encoder nimmt den Input – in diesem Fall Trainingssätze mit fünfzehn Prozent versteckten Wörtern – und erzeugt ein Zwischenergebnis, das an den Decoder weitergereicht wird. Der Decoder nimmt das Zwischenergebnis entgegen und versucht, die gestellte Aufgabe zu lösen – in unserem Fall: „Finde das Wort heraus, das im Satz fehlt.“

Das Zwischenergebnis, das der Encoder in diesem Szenario erzeugt, sind genau die oben skizzierten Worteinbettungen (Embeddings). Diese Einbettungen werden im Encoder so lange „trainiert“, bis es dem Decoder gelingt, anhand dieser Einbettungen das richtige Wort zu ergänzen. Sobald der Decoder eine sehr hohe Genauigkeit beim Lösen der Aufgabe aufweist, kann das Training beendet werden. Die Worteinbettungen, die der Encoder erzeugt hat, sind sozusagen ein Nebenprodukt des Trainingsvorgangs.

Genauso wie wir Menschen auf unsere sehr stark ausgeprägte Sprachkompetenz zurückgreifen müssen, um das Rätsel des fehlenden Wortes zu lösen, haben auch neuronale Netze, die dieses Rätsel ähnlich gut wie Menschen lösen, eine sehr stark ausgeprägte Sprachkompetenz – oder simulieren sie wenigstens. Man spricht hier von „General Language Understanding“.

Das Elegante an diesem Ansatz ist, dass ein neuronales Netz, das so wie oben beschrieben trainiert wird, mit leichten Anpassungen verwendet werden kann, um eine Vielzahl von NLP-Aufgaben zu lösen – nicht nur die Aufgabe, fehlende Worte zu ergänzen. Solche NLP-Systeme auf Basis von Embeddings können auch für maschinelle Übersetzung, Frage-Antwort-Systeme, automatische Zusammenfassungssysteme und vieles mehr angepasst und verwendet werden. ●

1 „Language is, at its core, a system that is both digital and infinite. To my knowledge, there is no other biological system with these properties.“ Noam Chomsky (1991): *Linguistics and Cognitive Science: Problems and Mysteries*; in Asa Kasher (ed.), *The Chomskyan Turn*. Oxford: Blackwell, pp. 26-53, p. 50.
 2 <https://de.wikipedia.org/wiki/Turing-Test> (abgerufen am 11.03.2021).
 3 Ludwig Wittgenstein (1953): *Philosophische Untersuchungen*, §4
 4 John Rupert Firth (1957): „A synopsis of linguistic theory 1930-1955“ und (1935) „The Technique of Semantics“

KI, ENTSCHEIDE DU MAL!

| von HANS-ULRICH STRIFFLER
und DR. ANDREAS ZAMPERONI

Transparenz und Nachvollziehbarkeit automatisierter Entscheidungen



In unserem Alltag lassen wir uns wie selbstverständlich von IT-Anwendungen unterstützen. Wir vertrauen den dort getroffenen Berechnungen und Entscheidungen mehr oder weniger, da wir davon ausgehen, dass die in einer IT-Anwendung programmierte Logik korrekt und deterministisch ist, das heißt bei gleichem Input immer die gleichen Resultate erzielt. Wir akzeptieren die „Entscheidungen“, weil sie dann für uns erklärbar oder zumindest nachvollziehbar sind. Unser Anspruch an IT-Systeme ist jedoch auch, dass sie die Beantwortung immer komplexerer Fragestellungen unterstützen, immer „smarter“ werden. Der Einsatz von künstlicher Intelligenz (KI) soll diesen Anspruch erfüllen. Dennoch sollen die Ergebnisse dieser immer komplexeren IT-Systeme, die getroffenen Entscheidungen für uns Menschen erklärbar und nachvollziehbar bleiben – was angesichts der Komplexität und Abstraktion der Lösungstechnologie KI besonders schwierig erscheint.

Ein wesentlicher Nachteil der bisher verwendeten KI-Technologien ist die fehlende Erklärungskomponente, die den Menschen die Entscheidungswege einfach und verständlich nachvollziehen lässt. Dies mag bei Anwendungen wie Routenplanung und Sprachübersetzung nicht relevant erscheinen. Sollten wir bei der Einreise in ein anderes Land von einer KI-Anwendung aber beispielsweise als „Terrorist“ klassifiziert werden, würde sich die Relevanz für uns schlagartig ändern.

Automatisierte personenbezogene Entscheidungen können wesentliche Auswirkungen auf den Einzelnen haben. Der Gesetzgeber trägt diesem Sachverhalt durch den Artikel 22 der DSGVO¹ Rechnung. Aktuell erweitert die EU-Kommission die Gesetzgebung noch. EU-Kommissarin Margrethe Vestager fordert für KI-Anwendungen in der öffentlichen Verwaltung, dass sie „besonders hohe Standards in Bezug auf Rechenschaftspflicht und Transparenz erfüllen ...“ müssen. Das „Weißbuch zur Künstlichen Intelligenz“ der Kommission, das in einer ersten Version am 19. Februar 2020 veröffentlicht wurde, geht auf diesen Sachverhalt ein.²

Einer für Geschäftslogik relevante Sprachstandard – „Decision Modelling Notation“ (DMN)³ – könnte hier Abhilfe schaffen. Er erlaubt es, Geschäftslogik mit KI zu verknüpfen. Im Folgenden stellen wir Möglichkeiten dar, wie auf die Forderungen der Bürger nach Transparenz und Nachvollziehbarkeit eingegangen werden kann.

KÜNSTLICHE INTELLIGENZ KOMMT VON KUNST – ODER?

Der Begriff der künstlichen Intelligenz (KI) ist nicht eindeutig festgelegt. In den 1950er-Jahren als interdisziplinäre Forschungsrichtung entwickelt, hat sie sich in ihrer Auslegung den

stetig wachsenden technischen Fähigkeiten angepasst. Das wesentliche Merkmal vieler KI-Definitionen zu allen Zeiten ist, dass die IT-Anwendung ein intelligentes Verhalten, vergleichbar mit dem des Menschen, aufweist.

Getrieben durch die Digitalisierung und den Innovationswillen sind die seit über sechzig Jahren erforschten Verfahren der KI in der heutigen Welt weiterentwickelt und zum Einsatz gebracht worden. Begünstigt wurde dies in den letzten Jahrzehnten durch die wachsenden Verfügbarkeiten großer Datenbestände, hoher Rechenleistung und erheblicher Fortschritte in der Methodik.

Es gibt verschiedene Einsatzgebiete, in denen ein intelligentes Verhalten bei IT-Anwendungen von hohem Nutzen ist. Dieser Artikel konzentriert sich auf den Einsatz von KI für automatisierte Entscheidungen. Wichtige KI-Techniken und -Methoden dafür sind maschinelles Lernen und Big Data.

Mittels maschinellem Lernen entwickelt eine KI-Anwendung selbständig Modelle, die das Wissen und die Erfahrung über spezifische Aufgaben abbilden. Die KI-Anwendung lernt anhand ausgewählter Lernbeispiele und entwickelt so Modelle, anhand derer sie (für sie bislang) unbekannte Daten bewertet oder Prognosen für die Zukunft ableitet.

So kann eine KI-Anwendung anhand von ihr vorgelegten Bildern erlernen, woran auf einem Bild bspw. ein Pferd zu erkennen ist. Anschließend kann die Anwendung für alle zukünftig vorgelegten Bilder entscheiden, ob auf ihm ein Pferd abgebildet ist oder nicht.

Maschinelles Lernen bietet sich immer dann an, wenn hinreichend viele Beispiele in Form von Daten, wie Bilder, Texte oder Datensätze, vorliegen – besonders dann, wenn die Sachverhalte zu komplex oder vielfältig erscheinen, um sie explizit zu beschreiben.

Der Begriff „Big Data“ bedeutet in diesem Zusammenhang die Analyse und Verarbeitung sehr großer Datenmengen zum Training der KI-Anwendung und ihres Modells.

DER „EXPERTE“ ENTSCHEIDET

Eine der Hauptanwendungsgebiete für KI ist es, fachliche Entscheidungen zu treffen, auch wenn keine dafür geeigneten – menschlichen – Experten zur Verfügung stehen, um bei der Entscheidung zu unterstützen. Als sogenannte „Expertensysteme“ wurden fachliche Entscheidungssysteme in den 1980er- und 1990er-Jahren erstmals implementiert.

Mit den heutigen, deutlich erweiterten technischen Möglichkeiten können hervorragende Entscheidungsunterstützungssysteme entwickelt werden. Gerade in der öffentlichen Verwaltung bietet sich aufgrund des rechtsnahen Entscheidungswissens die Assistenz oder sogar die Automatisierung von Entscheidungen an vielen Stellen geradezu an.

NATUR DER KI UND DIE TRANSPARENZFRAGE

Grundsätzlich sollten von IT-Anwendungen erzeugte Ergebnisse für Menschen nachvollziehbar und transparent sein. Dies gilt im Besonderen für KI-basierte Entscheidungen. Deren Qualität und Korrektheit zu vermitteln, ist jedoch eine schwierige Aufgabe. Einerseits werden KI-Anwendungen gerade in den fachlichen Sachverhalten eingesetzt, in denen unter Umständen keine eindeutige, korrekte – gar beweisbare – Lösung gefunden werden kann. Andererseits besitzen KI-Anwendungen eine Lernfähigkeit, die es ihnen ermöglicht, sich mit jeder getroffenen Entscheidung zu verbessern. Damit wird, im Gegensatz zur klassischen Algorithmik, fortwährend das Modell, das der Entscheidung zugrunde liegt, verändert. Die Nachvollziehbarkeit einer Einzelfallentscheidung wäre nur dann gegeben, wenn das zum Zeitpunkt der Entscheidung aktuelle Modell mit Bezug zu dem Einzelfall archiviert würde. Menschliche Experten können die Funktionsweise des Modells zum Zeitpunkt später nicht mehr nachvollziehen und überprüfen. Dies führt dazu, dass fachliche Fehler in den Modellen sehr schwer zu identifizieren – und damit auch nur schwer systematisch zu beseitigen – sind.

ENTSCHEIDUNGEN, DIE PERSONEN BETREFFEN

KI-Anwendungen werden zum Beispiel auch zur Klassifizierung oder Profilierung von Personen eingesetzt. Damit treffen sie direkt oder indirekt Entscheidungen oder geben Handlungsempfehlungen, so zum Beispiel zur Vorbewertung von Bewerbern oder zur Einschätzung der Kreditwürdigkeit eines Bankkunden.

Die der KI inhärente fehlende Nachvollziehbarkeit und Transparenz der Entscheidungsfindung führte tatsächlich schon zu verdeckter – möglicherweise sogar gewollter – Ausgrenzung, Diskriminierung und Ungleichbehandlung von Menschen, mit unter Umständen weitreichenden Folgen für deren Lebensumstände.⁴ Um solche Diskriminierung zu vermeiden, werden derzeit rechtliche und ethische Anforderungen auf europäischer Ebene aufgestellt.

REAKTION DER POLITIK

Die EU hat 2018 mit der DSGVO eine umfangreiche Gesetzgebung zum Datenschutz für personenbezogene Daten beschlossen. In dieser werden explizite Regelungen für automatisierte „personenrelevante Entscheidungen“ festgelegt. 2020 hat sich die EU-Kommission in ihrem Weißbuch zur künstlichen Intelligenz⁵ auf einen ersten Entwurf der Grundprinzipien für die Regulierung von KI-Anwendungen verständigt.

DSVGO

Die Regelungen der DSGVO besagen, dass Betroffene einen Anspruch auf Auskunft über den Einsatz von automatisierten Entscheidungen und deren wesentliche Arbeitsweise sowie das Hinzuziehen eines menschlichen Entscheiders haben. Damit sind die Voraussetzungen gegeben, eine automatisierte Entscheidung überprüfen und gegebenenfalls korrigieren zu lassen. „Ein darüber hinaus gehender Anspruch auf Einsicht in die KI-Anwendung besteht nicht.“⁶

WEISSBUCH DER EU-KOMMISSION / „VERTRAUENSWÜRDIGE KÜNSTLICHE INTELLIGENZ“

Laut Weißbuch sollen KI-Anwendungen in Einsatzbereichen mit hohem Risiko wie Gesundheit, Polizei oder Teilen der öffentlichen Verwaltung (wie Asyl, Migration, Grenzkontrollen und Justizwesen, soziale Sicherheit und Arbeitsverwaltungen) transparent und nachvollziehbar sein. Die Technik müsse stets von Menschen beaufsichtigt werden können. Behörden sollen die von Algorithmen genutzten (Trainings-)Daten ebenso prüfen und zertifizieren können, wie sie es bei Kosmetika, Autos und Spielzeug bereits tun.⁷

TRANSPARENT GENUG ODER DARF ES EIN BISSCHEN MEHR SEIN?

Für die menschliche Akzeptanz einer Entscheidung ist deren Nachvollziehbarkeit entscheidend. Voraussetzungen dafür sind die Verständlichkeit und die Lesbarkeit des der Entscheidung zugrunde liegenden Modells.

Doch wie transparent muss ein Entscheidungsmodell dokumentiert sein, um eine ausreichende Akzeptanz bei Anwendern und Fachexperten zu erreichen?

Eine vollständige Transparenz wäre die Veröffentlichung des gesamten Modells und der KI-Anwendung. Doch wenn die zu-

grunde liegenden Algorithmen Wissen beinhalten, das sensibel ist und daher vertraulich behandelt werden muss, ist eine vollständige Offenlegung nicht möglich.

Eine Lösung ist daher, die möglichst vollständige Transparenz des Modells zu bieten. Es geht nicht darum, die Entscheidung bis ins kleinste Detail zu erklären, sondern nur darum, die wesentlichen Schritte darzulegen.

WIE ERKLÄRE ICH ES ICH MEINEM MENSCHEN?

Die Modelle mit ihrem hohen Anteil logischer und mathematischer Aussagen lassen sich oft gar nicht in natürlicher Sprache verständlich beschreiben.

Eine Lösung verspricht der Sprachstandard „Decision Model and Notation“ (DMN)⁸. Er wurde als eigenständige Ergänzung zur bekannten Prozessmodellierungssprache „Business Process Model and Notation“ (BPMN) von der Object Management Group (OMG) entwickelt.



Der DMN-Standard ist so gestaltet, dass mit entsprechender IT-Unterstützung die in diesem Standard spezifizierten Entscheidungsmodelle direkt ausgeführt werden können. Es gibt eine Reihe von den DMN-Standard unterstützenden und zertifizierten „Entscheidungsmaschinen“⁹ (zum Beispiel das Open-Source-Produkt Drools von Red Hat). Bekannte Unternehmen und Behörden nutzen die DMN zur Modellierung ihrer Entscheidungen und zur direkten Ausführung im Produktivbetrieb oder stehen kurz vor der Inbetriebnahme im Produktivbetrieb.

Eines der Ziele der DMN ist es, Entscheidungen für Geschäftsanwender, Business-Analysten und technische Entwickler verständlich und eindeutig zu formulieren. Die DMN unterteilt sich in zwei Ebenen (Level). Die Ebene „Requirements Level“

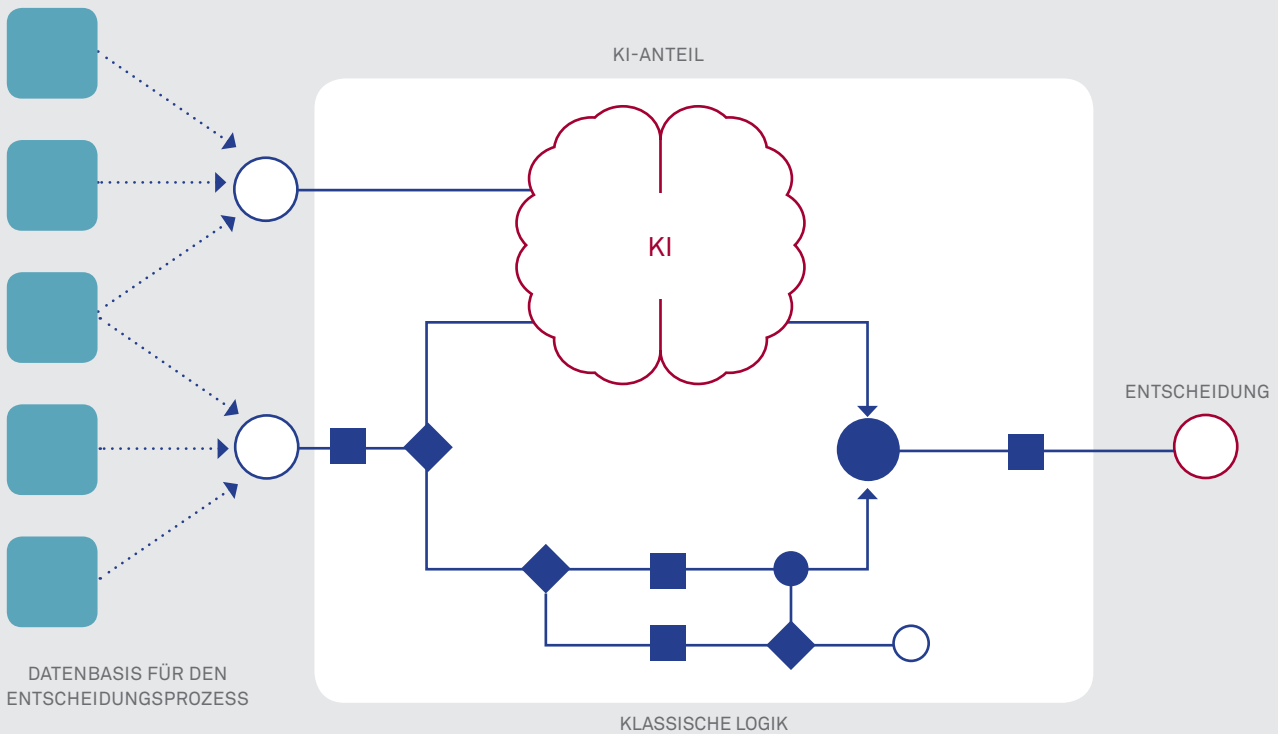


Abbildung 1: Entscheidungsprozess mit klassischer Logik und KI

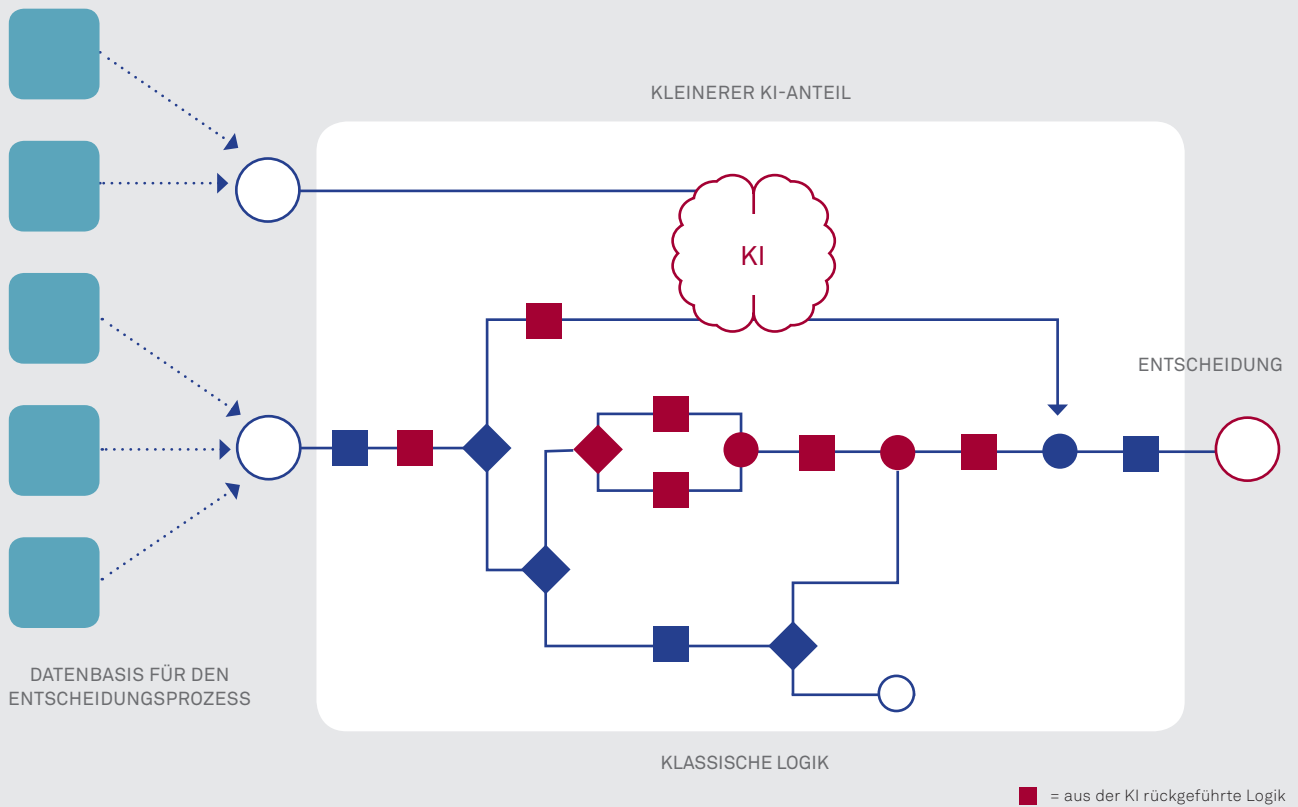


Abbildung 2: Rückführung von Teilen von KI in klassische Logik

gibt dem Leser eine grobe grafische Übersicht des Entscheidungsprozesses. Die zweite Ebene, „Logic Level“, beschreibt die konkrete Entscheidungslogik.

Der DMN-Standard erlaubt eine Einbindung von KI-Methoden in die Darstellung der Entscheidungsmodelle. Damit können Modelle explizit dargestellt und trotzdem die Details der KI-Anwendungen verborgen bleiben.

Der Sprachstandard DMN findet zunehmend Verbreitung in der öffentlichen Verwaltung. Bekannte Einsatzbeispiele sind das ATLAS-Projekt des deutschen Zolls¹⁰ und das OZG¹¹-Projekt „Föderales Informationsmanagement (FIM)“¹².

WAS ES ZU ENTSCHEIDEN GILT ODER: WIE WAR GLEICH NOCH MAL IHRE FRAGE?

KI-Anwendungen und Big-Data-Analysen sind keine Allheilmittel. Die Fragestellung an die KI muss eindeutig definiert sein. Ohne konkrete Vorgabe, was eigentlich zu entscheiden ist,

kann KI nicht zielgerichtet lernen oder nach Antworten suchen. KI kann dies nicht leisten, denn sie entwickelt Modelle weiter, ohne deren inhärente semantische Bedeutung zu „verstehen“. Zuerst muss die Fachlichkeit geklärt und der Entscheidungsraum festgelegt werden. Oftmals bedeutet diese fachliche Konkretisierung einen erheblichen Klärungs- und Konsolidierungsbedarf. Mit allen Beteiligten muss ein einheitliches, detailliertes, fachliches Verständnis herbeigeführt werden. Prozessanalysen sind ein guter Ansatz, um die fachliche Klärung durchzuführen. Mit der Analyse der Prozesse werden deren notwendigen Entscheidungspunkte und Entscheidungen ermittelt. Die Entsprechung zu DMN ist für die Prozesssicht die Modellierungssprache/Prozessnotation BPMN.

WENIGER (KI) IST MEHR (TRANSPARENZ)

Wie oben skizziert, liegt es dem Prinzip, wie KI-Systeme funktionieren, zugrunde, dass es schwierig bis unmöglich ist, die gewünschte Transparenz zu bieten (siehe Abbildung 1). Betrachtet man, wie ein (erwachsener) Mensch entscheidet, ist es auch so



gut wie unmöglich nachzuvollziehen, was und wie er in seiner Vergangenheit gelernt hat, um heute zu genau dieser Entscheidung zu kommen. Wie also die einzigartigen Fähigkeiten der KI nutzen, ohne die notwendige Transparenz aufzugeben?

Eine Möglichkeit ist, in einer Anwendung den KI-Anteil so gering wie möglich zu halten und stattdessen so viel wie möglich explizit durch klassische, nachvollziehbare Logik zu modellieren. An die Grenzen der expliziten Logikmodellierung stößt man dann, wenn es sich um sehr dynamisch veränderliche Sachverhalte handelt. Diese können besser durch die potenzielle Dynamik von KI gehandhabt werden. Die Intransparenz ist dann der Preis für diese nicht anders zu beherrschende Dynamik.

Oftmals verbirgt sich jedoch in einem Entscheidungsbestandteil hoher – wahrgenommener – Dynamik ein sehr komplexes, aber doch invariantes Konzept. Dieser Bestandteil lässt sich, einmal durchdrungen, sehr gut mittels klassischer Logik darstellen. Der tatsächliche Umfang des „harten Kerns“ der dynamischen Logik einer Anwendung ist oft wesentlich kleiner als zunächst angenommen.

Eine Strategie, den nachvollziehbaren und transparenten Teil eines KI-unterstützten Entscheidungssystems zu erhöhen, ist daher, dynamische, komplexe Anteile, die zum Beispiel

auf einer großen Datenbasis beruhen, durch eine KI-basierte (Vor-)Analyse zu erschließen und anschließend die nicht nachvollziehbaren, aber wiederholbaren Entscheidungen durch klassische Logik „nachzumodellieren“ und die KI-Komponente zu ersetzen, das heißt in eine Komponente klassischer Logik zurückzuführen (siehe Abbildung 2).

FAZIT

Unsere Gesellschaft und damit auch die Gesetzgeber haben die grundlegende Anforderung an KI erkannt, Entscheidungen nachvollziehbar und transparent zu machen und damit Bürger – aber auch Behördenmitarbeiter – einer echten oder gefühlten „Willkür“ auszusetzen. Regelungen mit entsprechenden Forderungen nach Transparenz und Nachvollziehbarkeit liegen vor und werden weiter vorangetrieben, um Akzeptanz für automatisierte Entscheidungen zu schaffen.

Dafür muss nicht zwingend eine vollständige Offenlegung des Entscheidungsmodells erfolgen. Eine Einbindung von nicht-transparenten KI-Anteilen ist möglich. DMN ist ein Standard, der sich für die Beschreibung von Entscheidungsprozessen gut eignet. Der Markt hat dies erkannt, und die Nutzung dieser Notation nimmt zu.

Um KI sinnvoll einzusetzen und zu beherrschen, ist es essenziell, vor dem KI-Einsatz die zu beantwortenden Fragestellungen im Rahmen einer Prozessentwicklung, zum Beispiel mittels BPMN, präzise zu ermitteln und abzustimmen.

Zur Erhöhung der Transparenz und Nachvollziehbarkeit von Entscheidungen auf Basis von KI-Systemen sollte immer darauf geachtet und geprüft werden, ob sich KI-Anteile einer Entscheidung durch nachträgliche Modellierung in klassische Logik rückführen lassen. ●

1 <https://dsgvo-gesetz.de/art-22-dsgvo/> (abgerufen am 25.02.2021).
 2 https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf (abgerufen am 17.03.2021).
 3 Decision Model and Notation™ (DMN™) | Object Management Group (omg.org) (abgerufen am 15.03.2021).
 4 <https://www.sueddeutsche.de/digital/microsoft-gesichtserkennung-rassismus-1.4934730> (abgerufen am 18.03.2021).
 5 Siehe 2.
 6 https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/BSt_DSGVOundADM_dt.pdf (abgerufen am 15.03.2021).
 7 https://ec.europa.eu/germany/news/20200219digitale-zukunft-europas-eu-kommission-stellt-strategien-fuer-daten-und-kuenstliche-intelligenz_de (abgerufen am 25.02.2021).
 8 <https://www.omg.org/dmn/> (abgerufen am 15.03.2021).
 9 Decision Model and Notation TCK (dmn-tck.github.io) (abgerufen am 15.03.2021).
 10 <https://www.itzbund.de/DE/itloesungen/fachverfahren/atlas/atlas.html> (abgerufen am 17.03.2021).
 11 <https://www.bmi.bund.de/DE/themen/moderne-verwaltung/verwaltungsmodernisierung/onlinezugangsgesetz/onlinezugangsgesetz-node.html> (abgerufen am 17.03.2021).
 12 <https://www.bmi.bund.de/DE/themen/moderne-verwaltung/verwaltungsmodernisierung/foerderales-informationsmanagement/foerderales-informationsmanagement-artikel.html> (abgerufen am 17.03.2021).

IS IT A MATCH?



Dublettenerkennung in der öffentlichen Verwaltung

| von NEDISLAV NEDYALKOV und DOMINIK MÜLLER

Manchmal erinnert unsere Arbeit an die Aufforderung eines etwas ruppigen Hauptkommissars an seinen Mitarbeiter in einem Vorabendkrimi: „Hier, gleichen Sie mal die Daten der Rentenkassen mit diesen Meldedaten ab!“ Nur landet in unserem Fall kein Aktenstapel auf dem Schreibtisch, sondern der Zugang zu einer äußerst umfangreichen Datenbank – so wie im Projektbeispiel bei der Bundesagentur für Arbeit, das wir in diesem Beitrag vorstellen.

Doch was hat das mit *Data Science* zu tun? Könnten wir nicht einfach jedes mögliche Rentenkassen-Meldedaten-Paar ansehen und prüfen, ob es sich

hier um dasselbe Objekt handelt? Leider nicht. Es würde schlichtweg viel zu lange dauern. Im Beispiel der Arbeitsagentur geht es um zehn Millionen Datensätze der Rentenkassen und drei Millionen Datensätze aus Meldedaten. Bei einem vollständigen Abgleich kämen wir auf dreißig Billionen Paarungen. Selbst wenn jeder automatische Abgleich nur eine Millisekunde dauern würde, entspräche das einer Rechenzeit von etwa 950 Jahren – ohne Pause, ohne Ausfälle. So lange möchte niemand warten – weder wir noch die Arbeitsagentur. Gefragt ist also eine weitaus schnellere Lösung. Und sie muss (zumindest fast) so gut sein wie ein vollständiger Abgleich. Damit stehen wir

also vor nicht weniger als der Herausforderung, sogenannte *Dubletten* zuverlässig erkennen zu können.

Herausforderung angenommen: Im Folgenden stellen wir eine Lösung für dieses Problem vor, ebenso wie einen Prototyp für die Implementierung. Keine Sorge: Auf mathematische Definitionen und Modellbeschreibungen verzichten wir. Stattdessen liefern wir eine anschauliche und praxisnahe Beschreibung. Für diejenigen Leser, die gerne mehr erfahren möchten, haben wir Schlüsselbegriffe *kursiv* gekennzeichnet und die wichtigsten in Infoboxen näher beschrieben.

PIPELINE-ÜBERBLICK

Zum Zweck der Dublettenerkennung wollen wir jeweils zwei Datensätze miteinander vergleichen und damit möglichst schnell und präzise herausfinden, ob sie dasselbe Objekt in der realen Welt beschreiben.

- Zunächst legen wir fest, welche Eigenschaften (*Features*) der Datensätze wir betrachten wollen. Das heißt, wir müssen eine Auswahl der Eigenschaften (*Feature Selection*) vornehmen, um Rechenzeit zu sparen und unser Ergebnis zu verbessern.
- Danach extrahieren wir Kandidatenpaare für die weitere Analyse. Ein Kandidatenpaar liegt vor, wenn mindestens eine bestimmte Anzahl an Features der beiden Datensätze zueinander passen.
- Als Nächstes trainieren wir einen sogenannten *Klassifikator*, der uns später ermöglicht, die Datenpaare in „Dubletten“ und „Nicht-Dubletten“ einzuteilen. Dafür benötigen wir Testdaten, die von Hand mit der entsprechenden Kategorie annotiert sind.
- Für die tatsächliche Überprüfung der Kandidatenpaare bilden wir aus jedem Paar einen sogenannten *Ähnlichkeitsvektor*. Er zeigt an, wie ähnlich sich die beiden Datensätze sind.

- Zuletzt analysieren wir diese Vektoren mittels des Klassifikators, um endgültig festzustellen, welche Paare dasselbe reale Objekt verkörpern, also echte Dubletten sind.

Im Folgenden unterziehen wir jeden dieser Teilschritte (siehe Abbildung 1) einer genaueren Betrachtung.

FEATURE SELECTION

In diesem Schritt suchen wir jene Eigenschaften unserer Datensätze heraus, die möglichst viele nützliche Informationen für unsere Zwecke enthalten. Das sind Features, die bei unterschiedlichen Objekten in der realen Welt möglichst häufig unterschiedlich sind. Oft ist hier Expertenwissen oder ein genauer Blick auf die Daten gefragt – es gibt jedoch auch statistische Methoden.

Um den Nutzen von Features messen zu können, ist es hilfreich, einen annotierten Testdatenbestand zu haben. Das heißt, jedem Testdatensatz ist eine Kategorie zugeordnet – etwa „Dublette“ oder „keine Dublette“. So lässt sich herausfinden, welche Features für unser weiteres Vorgehen interessant sind und welche nicht. Zum Beispiel können wir betrachten, welche Features bei

Nicht-Dubletten häufig gleich sind. Features wie Geschlecht oder Staatsangehörigkeit helfen verständlicherweise wenig, da sehr viele Personen dieselbe Ausprägung dieser Merkmale aufweisen. Deutlich seltener hingegen tritt bei unterschiedlichen Personen etwa eine identische Kombination von Name und Vorname auf.

Nachdem wir in diesem Schritt unsere Datensätze auf relevante Features getrimmt haben, sind wir bereit, mögliche Kandidatenpaare zu suchen. Da wir jedoch nicht jede mögliche Kombination einzeln analysieren können, nutzen wir *Locality-sensitive Hashing* (LSH).



LOCALITY-SENSITIVE HASHING

Locality-sensitive Hashing (LSH) benutzt Hashfunktionen, um die Datensätze einzuteilen. Eine Hashfunktion verwandelt große oder lange Eingabewerte in kleinere Zielwerte, mit denen man leichter arbeiten kann. LSH benutzt hierbei Hashfunktionen, die für ähnliche Ausgangswerte wieder ähnliche Zielwerte ergeben.

Auf jedes Feature werden mehrere solcher Hashfunktionen angewendet. Die Ergebnisse werden in sogenannte bins (Eimer) eingeteilt. In jedem bin sind die Werte zueinander ähnlich, ein Wert kann aber auch in mehreren bins vorkommen. Zwei Datensätze werden dann als Kandidatenpaar angesehen, wenn ihre Features sich eine Mindestanzahl an bins teilen. Diese Mindestanzahl ist abhängig von der Gestalt der Daten und muss durch Tests auf annotierten Testdaten bestimmt und optimiert werden.

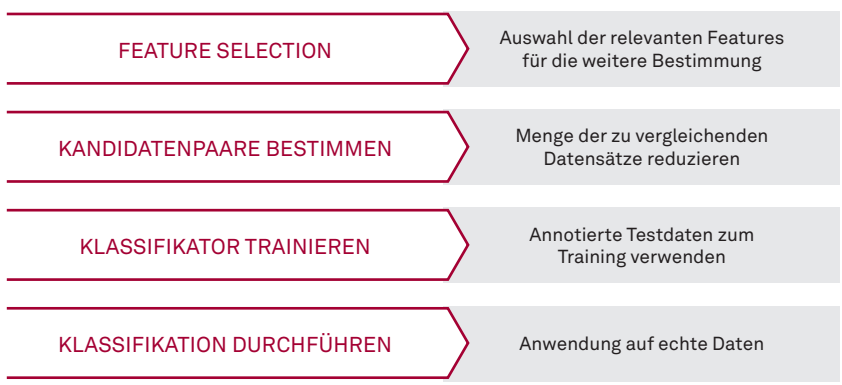


Abbildung 1: Schematische Darstellung eines möglichen Arbeitsablaufs bei der Dublettenerkennung

KANDIDATENPAARE BESTIMMEN

In diesem Schritt reduzieren wir drastisch die Menge an tatsächlichen Vergleichen, die wir explizit durchführen müssen. Wir versuchen also, aus allen möglichen Paarungen der Daten diejenigen zu finden, die am wahrscheinlichsten eine Dublette bilden.

Das LSH ordnet jeden Datensatz auf Basis seiner Features sogenannten *bins* (Eimern) zu.

Bis zu diesem Punkt haben wir unsere Features ausgewählt und eine stark reduzierte Anzahl an Kandidatenpaaren erzeugt. Diese müssen wir jetzt genauer untersuchen.

KLASSIFIKATOR TRAINIEREN UND ÄHNLICHKEITSVEKTOREN ERMITTELN

In diesem Abschnitt betrachten wir den letzten Schritt – die genaue Analyse der Kandidatenpaare. Dazu bauen wir für jedes Paar zunächst einen Ähnlichkeitsvektor. Er gibt an, wie ähnlich sich die beiden Datensätze sind. Danach füttern wir einen Klassifikator mit unseren Ähnlichkeitsvektoren. So finden wir schlussendlich heraus, bei welchen der Paare es sich tatsächlich um Dubletten handelt.

Die Ähnlichkeit wird einzeln für jedes Feature berechnet und unterscheidet sich je nach Datentyp. So bietet sich etwa für Zahlen der Abstand voneinander als Ähnlichkeitskriterium an und für Features mit eingeschränkter Wertemenge (zum Beispiel vier bis zehn mögliche Ausprägungen beim Familienstand je nach Anwendung) schlicht deren Übereinstimmung. Für Zeichenketten (etwa bei Namen) verwenden wir hingegen die sogenannte *Cosine Similarity*.

Diese Ähnlichkeitsbestimmung führen wir für alle Features unserer Kandidatenpaare durch und bekommen so eine Menge Ähnlichkeitsvektoren. Im letzten Schritt nutzen wir nun diese Vektoren, um herauszufinden, welche Paare wirklich Dubletten darstellen.

KLASSIFIKATION DER KANDIDATENPAARE DURCHFÜHREN

Wir haben jetzt für jedes Kandidatenpaar einen sogenannten „Ähnlichkeitsvektor“. Nun werden die Vektoren noch in „gleich“ und „nicht gleich“ eingeteilt. Dafür bedienen wir uns eines sogenannten „Klassifikators“, hier einer *Support Vector Machine* (SVM).

Ein Klassifikator ist ein Konzept, mit dem man Daten in Kategorien einteilen, also klassifizieren kann. Das Vorgehen unterteilt sich hierbei allgemein in zwei Schritte: Training und Anwendung.

- Beim **Training** eines Klassifikators werden die Parameter festgelegt, anhand deren der Klassifikator später die Klassenzugehörigkeit der Datensätze bestimmt. Dies geschieht üblicherweise auf Basis von Trainingsdaten, also bereits vorher annotierter Daten, für die die korrekte Klassenzugehörigkeit bekannt ist.



COSINE SIMILARITY

Die *Cosine Similarity* (Kosinus-Ähnlichkeit) vergleicht zwei Vektoren basierend auf ihrem Winkel zueinander. Um unsere Wörter aber in Vektorform zu bringen, zerlegen wir sie zunächst in 3-Gramme. Das sind Teile eines Worts (umliegende Leerzeichen oder Ähnliches eingeschlossen), die drei Zeichen lang sind. Diese Zerlegungen werden in Vektorform gebracht und dann verglichen, um eine Ähnlichkeit der Wörter zu ermitteln.

Das bedeutet, wir konstruieren aus allen möglichen 3-Grammen in den Wörtern unseres Features eine einzelne Menge, quasi eine sehr lange Liste. Diese Liste eines Wortes ist genau so lang, wie das Wort 3-Gramme enthält, besteht aber nur aus „0“ und „1“. Eine „0“ steht an einer Stelle, wo das entsprechende 3-Gramm nicht im Wort auftaucht. Wenn das 3-Gramm Teil des Wortes ist, so steht an dieser Stelle eine „1“.

Diese Vektorrepräsentationen werden nun anhand des Winkels zwischen den beiden Vektoren verglichen.¹

Vergleichen wir zum Beispiel die Wörter „Ende“ und „Ente“. Die 3-Gramme der beiden Wörter werden in Tabelle 1 gezeigt. Nun transformieren wir diese in Vektoren. Die Grundmenge der Einträge ist die Anzahl aller unterschiedlichen 3-Gramme, also genau die Menge $M = \{“_En”, “End”, “nde”, “de_”, “Ent”, “nte”, “te_”\}$. Die Vektorrepräsentationen sind ebenfalls in Tabelle 1 eingetragen.

Wort	3-Gramme	Vektor auf {“_En”, “End”, “nde”, “de_”, “Ent”, “nte”, “te_”}
Ende	“_En”, “End”, “nde”, “de_”	(1, 1, 1, 0, 0, 0)
Ente	“_En”, “Ent”, “nte”, “te_”	(1, 0, 0, 0, 1, 1)

Tabelle 1: Aufteilung von Wörtern in 3-Gramme und Vektorrepräsentation



SUPPORT VECTOR MACHINE

Eine Support Vector Machine (SVM) unterteilt Datensätze in zwei Kategorien. Dies geschieht mittels einer sogenannten „Hyperebene“, die den gesamten Wertebereich der Datensätze aufspannt. Datensätze, die auf der einen Seite der Hyperebene liegen, bekommen dann die Klasse A (in diesem Fall Dubletten) zugewiesen. Die Datensätze auf der anderen Seite der Hyperebene sind Teil der Klasse B (in diesem Fall Nicht-Dubletten).

Die Lage der Hyperebene wird durch die Trainingsdaten bestimmt. Aus der Positionierung der annotierten Trainingsdaten ergibt sich eine optimale Hyperebene. Optimal heißt: Die Klassifikation der Testdaten in die Klassen A und B ist möglichst korrekt (das heißt, die Testdaten liegen auf der richtigen Seite der Hyperebene), und sie ist möglichst eindeutig (das heißt, der Abstand zwischen den beiden Klassen A und B ist möglichst groß).

In der trainierten Anwendung werden neue Datensätze dann einfach durch Positionierung im Hyperraum auf Basis ihrer Kriterien den Klassen A (Dubletten) oder B (Nicht-Dubletten) zugeordnet.

- Bei der **Anwendung** wird eine Funktion auf die Datensätze angewendet und so entschieden, welcher Klasse jeder Datensatz angehört.

In unserem Fall möchten wir unsere Ähnlichkeitsvektoren danach einteilen, ob sie eine Dublette darstellen oder nicht. Unser Klassifikator soll bei der Anwendung also genau das entscheiden.

BEISPIELHAFTE IMPLEMENTIERUNG

Bis hierher haben wir ein Vorgehen zur Dublettenerkennung beschrieben, das so auch beim Projekt der Bundesagentur für Arbeit zum Einsatz kam. Im Folgenden stellen wir eine prototypische Implementierung vor.

Die Dublettenerkennung bei Personendaten aus unterschiedlichen Datenbanken im Behördenumfeld ist ein konkretes und regelmäßig wiederkehrendes Matching-Problem. Deshalb haben wir auf dieser Grundlage einen Proof-

of-Concept erstellt – also einen Software-Prototyp, mit dem die prinzipielle Durchführbarkeit des KI-basierten Vorgehens belegt werden kann.²

PROGRAMMIERSPRACHE UND BIBLIOTHEKEN

Bei unserer prototypischen Implementierung haben wir auf die in KI-Kreisen gängige Programmiersprache *python* gesetzt. Neben *python* selbst haben wir uns zusätzlich die freie Software-Bibliothek *scikit-learn* und das *Python Record Linkage Toolkit* (*recordlinkage*) zum maschinellen Lernen für die Programmiersprache *python* zunutze gemacht.

DIE DATEN

Jedes KI-Projekt braucht eine große Menge Daten. Nur mit ihnen können die Algorithmen lernen, und nur mit ihnen können wir im Anschluss die Güte des antrainierten Algorithmus bestimmen. Im konkreten Fall haben wir Daten aus

der Bibliothek *Febrl* (Freely Extensible Biomedical Record Linkage) genutzt. Dabei handelt es sich um Personendaten, die für KI-Zwecke frei verfügbar sind und beispielsweise Merkmale wie Vor- und Nachname aufweisen.

SOFTWARE-PROTOTYP

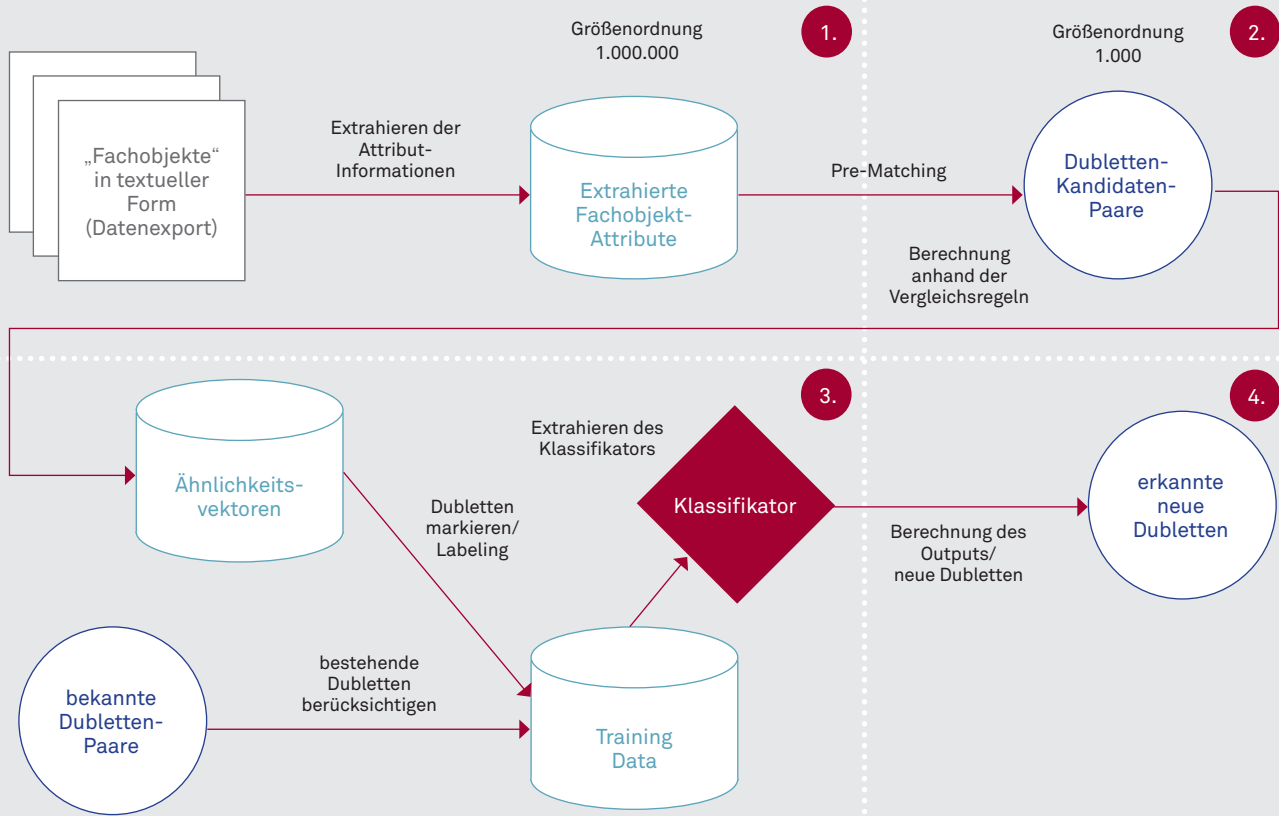
Basierend auf den bisher besprochenen theoretischen Grundlagen und Software-Tools haben wir die Implementierung in vier Schritten durchgeführt (siehe Abbildung 2).

Schritt 1 – Auswahl der Kandidatenpaare

Zuerst haben wir die zu vergleichenden Paare von Objekten aus der Menge aller möglichen Paare gesammelt. Für die Bestimmung dieser Dubletten-Kandidatenpaare haben wir das sogenannte Index-Konstrukt aus der Bibliothek *recordlinkage* ausgewählt. Dieses ermöglicht es vergleichsweise schnell, Paare zu identifizieren, die mit einer recht hohen Wahrscheinlichkeit Dubletten sein könnten. Somit haben wir die Anzahl der zu vergleichenden Paare erheblich reduziert, um die Menge an zu vergleichenden Objekten mit vertretbarem Aufwand bewältigen zu können. Im vorliegenden Fall konnten wir auf das Konzept der *Feature Selection* verzichten, weil die vorliegenden Trainingsdaten eine überschaubare Anzahl an Merkmalen (*Features*) aufweisen und eine weitere Auswahl somit nicht notwendig ist.

Schritt 2 – Vergleichsregeln definieren

Im zweiten Schritt werden für die vorher bestimmten Dubletten-Kandidatenpaare mit aufwendigeren und viel präziseren Zeichenkettenvergleichen Vergleichsregeln definiert. Beispielsweise werden die Vor- und Nachnamen auf Ähnlichkeit geprüft. Für diesen Schritt bietet sich unter anderem das Verfahren aus dem Abschnitt „Kandidatenpaare bestimmen“ an.



1. Auswahl der Kandidatenpaare 2. Vergleichsregeln definieren 3. Wahl des Klassifikators und Lernen 4. Dublettenerkennung anwenden

Abbildung 2: Der Prozess der Dublettenerkennung durch KI

Mit den einzelnen Ergebnissen dieser Vergleiche werden die sogenannten Ähnlichkeitsvektoren aufgestellt, die dann der Input für den nächsten Schritt sind.

Schritt 3 – Wahl des Klassifikators und Lernen

Für den Kern der Lösung wählen wir einen Klassifikator und trainieren ihn anhand der Trainingsdaten. Es wird das sogenannte überwachte Lernen angewendet, bei dem bereits klassifizierte Daten die Grundlage bilden. Für unseren Algorithmus haben wir die *Logistic Regression* ausgewählt, ein etab-

liertes mathematisches Verfahren in der Statistik. Mit dieser Methode kann die Wahrscheinlichkeit bestimmt werden, mit der Fragen etwa mit *Ja* oder *Nein*, *Gewinn* oder *Verlust*, *Gesund* oder *Krank* beantwortet werden. Die Parameter unseres mathematischen Modells wurden anhand bekannter Daten bestimmt.

Schritt 4 – Dublettenerkennung anwenden

Erst jetzt, nach dem Trainieren der KI, können wir die gelernten Verfahren auf neue, unbekannte Daten anwenden, um die Datenpaare korrekt in *Dubletten* oder *Nicht-Dubletten* zu unterteilen. In diesem

Schritt wenden wir deshalb die antrainierte *Logistic Regression* auf neue Datensätze an. Dabei sind wir in zwei Durchläufen folgendermaßen vorgegangen:

Im ersten Durchlauf haben wir die *Logistic Regression* mit dem in Schritt 3 antrainierten Parametern durchgeführt. Gefunden wurden 490 Dubletten bei tatsächlich 500 vorhandenen, was ein sehr gutes Ergebnis ist.

Um eine mögliche Veränderung der Ergebnisse zu analysieren, haben wir in einem zweiten Durchgang die Parameter

Anzahl Testdatensätze	1.000
Alle möglichen Paare	ca. 1.000.000
Anzahl Dubletten-Kandidatenpaare nach der Bestimmung der Dubletten-Kandidatenpaare (Schritt 1)	3.636
Anzahl als Dubletten erkannte Kandidaten (Schritt 4)	490 (Durchlauf 1) 415 (Durchlauf 2)
Tatsächliche Dubletten	500

Tabelle 2: Ergebnisse der Dublettenprüfung des Prototyps

für die Bestimmung der Dubletten-Kandidatenpaare geändert. Das Ergebnis: Ein Wegfall des Vergleichs der Sozialversicherungsnummer führt zu einer Verschlechterung der Ergebnisse (siehe Tabelle 2). Hier finden wir nur noch 415 Dubletten bei 500 tatsächlich vorhandenen. Dementsprechend haben wir nun zusätzlich belegt, dass die Sozialversicherungsnummer eine zentrale Bedeutung für die Qualität der Lösung hat.

Durch diese schrittweise Verfeinerung können die Ergebnisse besser verstan-

den und optimiert werden, bis die Ergebnisse für die reale Nutzung geeignet sind.

FAZIT

In diesem Artikel haben wir das reale und regelmäßig wiederkehrende Problem der Dublettenerkennung vorgestellt, Lösungsansätze skizziert und einen Proof-of-Concept beschrieben, der mit realen Testdaten funktioniert. Dies sind wichtige Grundlagen für die Orientierung und den Einstieg in den Themenkomplex.

In der Wirklichkeit existieren zusätzliche Hürden. Eine davon ist die Qualität der zugrunde liegenden Daten: Datenbestände werden zumeist von Menschen gepflegt, und Menschen machen Fehler. Zum Beispiel Tippfehler oder unterschiedliche Konvertierung von Sonderzeichen. Im Projekt der Bundesagentur für Arbeit war zum Beispiel nicht immer ersichtlich, welches Feld eines Datensatzes den Vornamen und welches den Nachnamen enthielt. Die gute Nachricht: Auch dafür gibt es Lösungen. Im konkreten Fall wurden beide Features als eine einzige Zeichenkette betrachtet.

Das Beispiel zeigt, wie die Arbeit von Data Scientists, Data Analysts und KI-Experten für die öffentliche Verwaltung immer wichtiger wird. Denn die öffentliche Verwaltung steht in nahezu allen Bereichen vor demselben Problem: Schnell wachsende Datenbestände machen eine manuelle Verarbeitung schon heute so gut wie unmöglich. Die Lösung für solche Probleme kann bereits gut durch automatisierte KI-Verfahren unterstützt werden. ●

¹ Für die genaue Formel verweisen wir auf <https://arxiv.org/pdf/1910.09129.pdf> (abgerufen am 04.02.2021).

² Der Prototyp basiert auf den konzeptionellen und technologischen Erkenntnissen und Lösungen aus: Shu Rong, Xing Niu, Evan Wei Xiang, Haofen Wang, Qiang Yang and Yong Yu, A Machine Learning Approach for Instance Matching Based on Similarity Metrics (https://link.springer.com/content/pdf/10.1007%2F978-3-642-35176-1_29.pdf, abgerufen am 30.03.2021) sowie Bilenko et al., Learnable Similarity Functions and their application to data linkage and clustering (<https://www.cs.utexas.edu/~ml/papers/martin-aaaidc-04.pdf>, abgerufen am 30.03.2021).

Für die technologische Auswahl siehe: <https://www.python.org> (abgerufen am 01.02.2021), <https://sourceforge.net/projects/febrl/> (abgerufen am 01.02.2021), <http://recordlinkage.readthedocs.io/en/latest/ref-datasets.html> (abgerufen am 01.02.2021), <https://scikit-learn.org/stable/index.html> (abgerufen am 01.02.2021), http://recordlinkage.readthedocs.io/en/latest/notebooks/data_deduplication.html (abgerufen am 01.02.2021), <https://scikit-learn.org/stable/> (abgerufen am 29.03.2021).

„Sprachinteraktion ist beim Endkunden angekommen und akzeptiert. Der Einsatz von Chatbots und konversationellen Benutzerschnittstellen im Bereich E-Government kann und muss jetzt begonnen werden, um diese zeitnah einsatzbereit zu haben. Im E-Government bietet sprachbasierte Interaktion immense Potentiale.“

Dr. Stefan Schaffer, DFKI GmbH (Quelle: NEGZ STANDPUNKT, NR. 6 – SPRACHSTEUERUNG VON E-GOVERNMENT DIENSTEN, 2020)

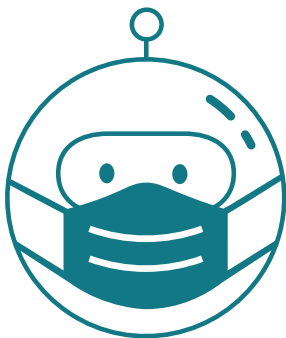
Hierfür wird im Rahmen der Dienstekonsolidierung Bund die zentrale Entwicklung einer einheitlichen digitalen Lösung angestrebt. Der „Basisdienst Chatbot“ soll gleichartige Aufgaben in den verschiedenen Behörden erfüllen können. Die Gesamtverantwortung für die Einführung trägt das Bundesinnenministerium. Fachliche Anforderungen werden über eine fachverantwortliche Stelle in der Generalzolldirektion ressortübergreifend gebündelt. Für den Betrieb ist das Informationstechnikzentrum Bund (ITZ Bund) zuständig, der zentrale IT-Dienstleister der Bundesverwaltung.

DIE CHATBOTS

C-19 ist das erste öffentliche Pilotprojekt in diesem Kontext. Vom Gesundheitsministerium bis zur Generalzolldirektion haben bislang zehn Ministerien und Behörden ihre Informationen zum Coronavirus in die Chatbot-Datenbank eingespeist, aus der C-19 seine Antworten bezieht. Weitere Bundesbehörden werden als Nutzer von C-19 erwartet.

Internetnutzer verlassen sich zunehmend auf C-19. Mehrere hunderttausend Zugriffe zählt der Chatbot monatlich. Davon waren rund zwanzig Prozent durchgeführte Gespräche, in denen die Nutzerinnen und Nutzer durchschnittlich etwa zehn Fragen an den Chatbot stellten. Im Sommer standen vorrangig Fragen zum Reisen im Vordergrund. Mit dem Wiederanstieg der Infektionszahlen und aufgrund der erweiterten Antworten in C-19 bezieht sich ein großer Teil der Anfragen auf die aktuellen Fallzahlen und Regelungen. In den Auswertungen werden ausschließlich anonyme Gesprächsverläufe ohne jeglichen Bezug auf konkrete Personen verarbeitet – damit können sich die Behörden auch einen guten Überblick verschaffen, was die Menschen in der Pandemie aktuell beschäftigt.

Die Generalzolldirektion (GZD) stand vor einer anderen Herausforderung: Sie wollte verlässliche Auskünfte zum Thema Brexit liefern. Zu diesem Zweck hat sie seit Ende 2020 einen Chatbot konzipiert und mit Unterstützung des ITZBund realisiert. Der Brexit-Bot stand pünktlich zum Brexit zur Verfügung und gibt seitdem zweisprachig Auskunft zum Thema.



C-19 ist der Chatbot der Bundesverwaltung zu Themen rund um Corona
<https://chatbot.it.bund.de/>



Der Brexit-Bot ist der Chatbot der Zollverwaltung unter Beteiligung des Bundesministeriums des Innern, für Bau und Heimat (BMI) und des Bundeszentralamtes für Steuern (BZSt) zu Fragen rund um den Brexit.
<https://brexit.it.bund.de/>



Der Chatbot „LinA“ ist der Chatbot der Zollverwaltung zu den Themen Kraftfahrzeugsteuer und Internet-Verbrauch- und Verkehrsteuer-Anwendung (IVVA).
<https://lina-zoll.it.bund.de/>

Abbildung 1: Beispiele für Chatbots der Bundesverwaltung

Ebenfalls von der GZD stammt LinA, ein Chatbot, der Fragen rund um die Kraftfahrzeugsteuer und zur Internet-Verbrauch- und Verkehrsteuer-Anwendung beantwortet. Im Fokus steht hier, interne Aufwände im eigenen Servicecenter zu reduzieren, indem die Topthemen von Anfragenden gezielt aufgegriffen und beantwortet werden.

CHATBOTS – WOHIN GEHT DIE REISE?

Wie die Darstellung der möglichen Entwicklungsstufen von Chatbots zeigt, befinden wir uns immer noch in der Anfangsphase der Entwicklung von Assistenten mit künstlicher Intelligenz (siehe Abbildung 2).

Heute können Chatbots in der Verwaltung häufig vor allem einfache Verbindungen zwischen Nutzerfragen und statischen Antworten herstellen, die zuvor von einer Redaktion erstellt wurden. Diese Verbindung basierte in der Vergangenheit auf klaren Regeln, die in der Programmierung festgeschrieben waren. Aktuelle Chatbot-Plattformen lernen mithilfe künstlicher Intelligenz immer besser, was ein Nutzer, eine Nutzerin eigentlich meint. Sie können auch den Gesprächskontext besser deuten, sich Informationen merken und dann im Kontext fast wie ein menschlicher Gesprächspartner reagieren.

In Zukunft werden immer mehr Chatbots in der öffentlichen Verwaltung für viele unterschiedliche Behörden Fragen beant-

worten. Daher wird es immer wichtiger, allgemeine Informationen für alle behördenspezifischen Bots zentral zur Verfügung zu stellen. Der nächste größere Schritt wird dann in der Vermittlung von Bot zu Bot liegen. Chatbots in der öffentlichen Verwaltung müssen also lernen, in Zukunft untereinander vernetzt Informationen auszugeben. Nutzerinnen und Nutzer können sich dann zum Beispiel in einem Chatbot-Dialog über Coronaschutzimpfungen informieren und sich an selber Stelle direkt nach den Öffnungszeiten und der Adresse der örtlichen Impfstelle erkundigen.

In Zusammenarbeit mit weiteren Basisdiensten wie etwa dem Nutzerkonto Bund¹, können Chatbots der nächsten Entwicklungsstufe im Auftrag der Nutzerin oder des Nutzers sogar Verwaltungsleistungen beantragen. Die notwendigen Formulare wären dann bereits vorausgefüllt, weil die Nutzerin oder der Nutzer über ihr oder sein Nutzerkonto Bund authentifiziert ist.

Andere Chatbot-Konzepte sehen zukünftig etwa eine Live-Begleitung und die Übergabe an Beschäftigte in Callcentern vor – nachdem der Chatbot im Vorfeld bereits Standardaufgaben wie die Erfassung von Adressinformationen übernommen hat.

Ein weiterer Schlüssel zum Erfolg dieser Systeme ist deren möglicher Einsatz auf den unterschiedlichsten Kommunikationskanälen, wie beispielsweise in Messengern und in vorhandenen Chats oder Onlineforen.

Kundenservice / Bot Stufe	Regelbasierte Dialoge	Gespräche im Kontext	Merken und Einsetzen von Information	Handeln im Auftrag des Nutzers (Bot zu Bot)	Als Repräsentant des Nutzers mit anderen Bots verhandeln
Chatbot (seit 10 Jahren)	✓				
Chatbot Stufe 2, FAQ Bots (jetzt)	✓	✓			
Chatbot Stufe 3 (Potenzial für die nächsten 2 Jahre)	✓	✓	✓		
Chatbot Stufe 4, persönlicher Assistent (in 2-5 Jahren)	✓	✓	✓	✓	✓

Abbildung 2: Übersicht über die Entwicklungsstufen von Chatbots

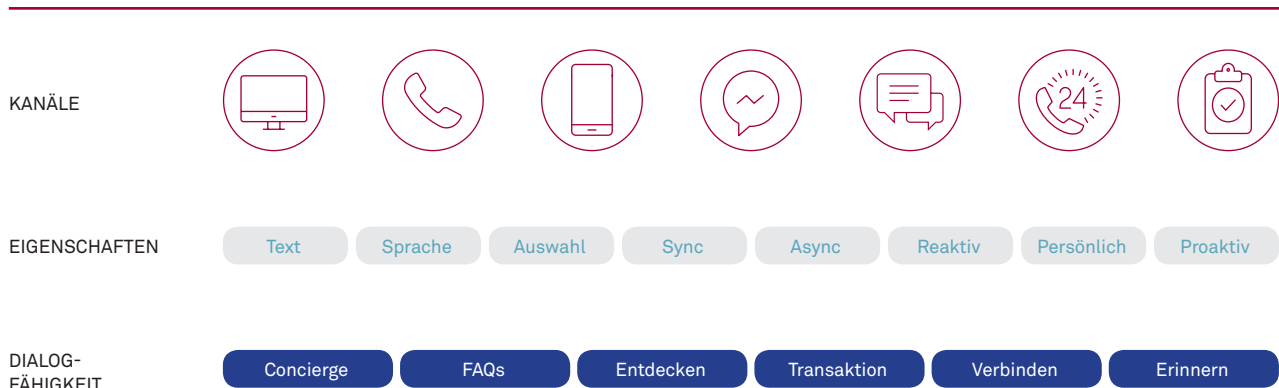


Abbildung 3: Die unterschiedlichen Kanäle und ihre Eigenschaften für dialoggesteuerte Systeme (Quelle: ITZBund)

VIELFÄLTIG EINSETZBAR – DIE BANDBREITE VON DIALOGSYSTEMEN

Das Chatbot-Prinzip muss nicht ausschließlich auf Webseiten zum Einsatz kommen, sondern kann andere Ausprägungen annehmen. Man spricht im weiteren Sinne von dialoggesteuerten und nicht-textbasierten Systemen. Beispielsweise können „Voicebots“ am Telefon bereits ohne menschliche Hilfe viele Fragen von Bürgerinnen und Bürgern beantworten – deutlich besser als bisherige automatisierte Telefonservices.

Ebenso ist das Smartphone ein Kanal, auf dem schon heute eigene Chatbot-Apps die Nutzerinnen und Nutzer unterstützen. Auch Messenger-Dienste werden immer häufiger Chatbots beinhalten, die gezielt im synchronen Dialog Auskunft geben, bis hin zum persönlichen „Agenten“, der aufgrund seiner Kenntnisse über den Nutzer konkrete Aufgaben selbstständig und proaktiv übernehmen kann.

CHATBOTS IN DER VERWALTUNG

Der Basisdienst Chatbot hat bereits einige mögliche zukünftige Einsatzszenarios entwickelt. Da jedoch mehr und mehr Fachstellen ihr Interesse bekunden, wird die Liste an neuen Anforderungen immer länger. Eine der zukünftigen Kernaufgaben des Basisdienstes wird somit die zentrale Bereitstellung von standardisierten Chatbots sein, die auf Knopfdruck installiert und möglichst einfach konfiguriert, trainiert und betrieben werden können.

Chatbots können für die unterschiedlichsten Themenfelder und Ministerien sinnvoll sein, so zum Beispiel ein Fach-Chatbot für Familienfragen oder eine spezielle Version für Migrationsfragen.

Eine besondere Rolle nehmen in diesem Zusammenhang auch Voicebots ein, da sie helfen, das Anrufaufkommen in Behörden zu verringern. Zusätzlich sind Chatbots auch für die internen Abläufe in der Verwaltung ein Thema, um beispielsweise den technischen Helpdesk zu unterstützen.

ACHT GRÜNDE, WARUM DIE ÖFFENTLICHE VERWALTUNG CHATBOTS NUTZEN SOLLTE

Chatbots sprechen die Sprache der Menschen

Bürgerinnen und Bürger beherrschen selten das sogenannte Behördendeutsch, sondern drücken sich meistens umgangssprachlich, alltagsorientiert oder gar im Dialekt aus. Chatbots haben das Potenzial, die Sprache der Bürgerinnen und Bürger zu verstehen und komplexe Sachverhalte einfach und verständlich zu formulieren. Dank Künstlicher Intelligenz lernen sie außerdem aus allen Gesprächen und Fragen von Nutzerinnen und Nutzer immer besser, was ihre Anfragen bedeuten und welche Antwort dazu passt.

Chatbots sagen die Wahrheit

Chatbots im Dienst der Verwaltung können ein Mittel gegen Fake News sein, da sie sich nur aus zuverlässigen offiziellen Quellen bedienen („Single Source of Truth“) und die Informationen personalisieren. Bürgerinnen und Bürger wollen Vertrauen zu einer zentralen Quelle aufbauen können, die zudem noch einfach zugänglich und leicht verständlich ist.

Überall zu Hause

Chatbots erreichen einen hohen Anteil von Nutzerinnen und Nutzern in ihren unterschiedlichen Lebenssituationen, da sie

in vielen Kommunikationskanälen parallel zu Hause sein können. Immer mehr Endgeräte sind dialogfähig, man denke nur an Siri von Apple auf dem Smartphone oder an Amazons Alexa im heimischen Wohnzimmer.

So hat etwa die World Health Organisation einen Chatbot etablieren können, der über den Messenger WhatsApp Millionen von Nutzerinnen und Nutzer auf ihren Smartphones erreicht. In Lockdown-Zeiten hat die Nutzung von Messenger-Diensten wie WhatsApp um durchschnittlich fünfzig Prozent zugenommen. Damit bieten diese Kanäle auch Behörden eine weitere, zunehmend relevante Dialogmöglichkeit mit den Bürgerinnen und Bürgern.

Chatbots können auch sprachgesteuerte Kanäle bedienen. So bietet etwa das Bayerische Staatsministerium für Gesundheit und Pflege einen so genannten „Alexa-Skill“. Dank ihm kann man Alexa zum Beispiel auffordern: „Alexa, frage Coronavirus Bayern nach den Fallzahlen in München.“²

Hohe Qualitätsstandards durch einen zentralen Dienst

Chatbots bedeuten heute noch einen gewissen Aufwand: Ihre inhaltliche, fachliche Ausgestaltung ist zeitintensiv. Hinzu kom-

men die weiterführende Pflege und Optimierung des Chatbots. Das erfordert spezialisiertes Personal mit entsprechenden technischen und fachlichen Kenntnissen. Kaum leistbar für einen einzelnen, kleineren Themenbereich in der Verwaltung. Daher ist es wichtig, einen zentralen Dienst zu schaffen. Eine für die Bundesverwaltung verfügbare zentrale On-Premise-Infrastruktur für konversationelle Dienste – also im weitesten Sinne Chatbots – ist bereits heute im ITZBund definiert. Sie schafft die Grundlage für eine umfassende Qualitätssicherung und reduziert inhaltliche und technische Betriebskosten.

Chatbots wissen mehr

Bürgerinnen und Bürger erwarten von einem Chatbot umfassende Informationen zu einem Thema, und zwar über Behördengrenzen hinweg – die sie nicht kennen und die für sie keine Rolle spielen. An dieser Stelle können Chatbots eine thematische Klammer bilden. Es gibt dann nicht den Chatbot des Bundesministeriums für Verkehr und digitale Infrastruktur, sondern beispielsweise den Kfz-Chatbot. Der kennt auch Themen zur An- und Abmeldung eines Autos, zu Verkehrsregeln und zur Kfz-Steuer und kann dann etwa auch in Zukunft die örtliche Zulassungsstelle ausfindig machen.

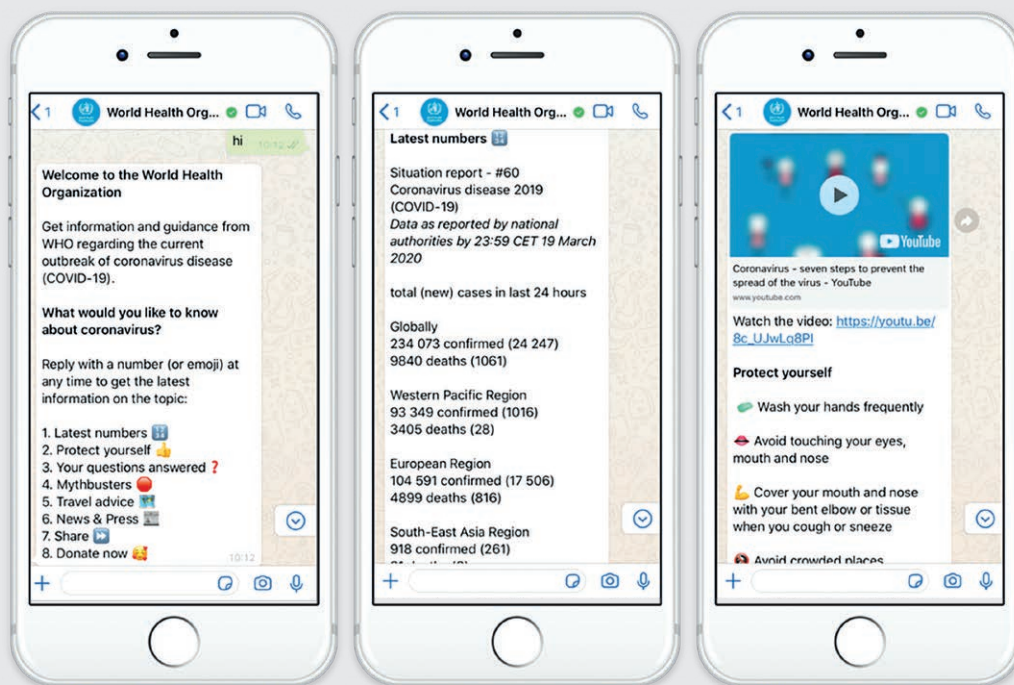


Abbildung 4: Der COVID-19-Chatbot der WHO



WAS IST EIN ALEXA-SKILL?

Amazon bietet einen Store für sogenannte Skills, die von Drittanbietern erstellt werden können. Skills werden vom Benutzer aktiviert und bieten zusätzliche Funktionen wie das Steuern von Smart-Home-Geräten, Spiele, Nachrichten oder die Kommunikation mit einer Bank.⁴

Chatbots können Informationen aus vielfältigen Quellen zu einem fachlichen Thema vereinen und personalisieren. Damit gehen sie weit über eine behördenspezifische Webseite hinaus und bieten ein umfassendes Informationsmedium.

Chatbots können mehr

Aktuelle Chatbots in der öffentlichen Verwaltung funktionieren heute häufig wie typische Frage-Antwort-Listen (FAQ): Man kann sie zu einem Thema befragen und erhält eine fachlich passende Antwort. Einige Chatbots sind jedoch längst nicht mehr darauf beschränkt. Sie können etwa auf Anfrage direkt die fachlich relevanten Formulare anbieten, vorausfüllen oder für den Nutzer sogar einen Termin bei einer Behörde vereinbaren. Erste technische Lösungen werden bereits im Rahmen des Basisdienstes Chatbot prototypisch entwickelt. Jedoch befinden sich die Entwickler noch am Anfang. Und das hat nicht nur technische Gründe: Das Zusammenspiel aus persönlichem Austausch mit dem Chatbot und der Nutzung gespeicherter Informationen für ein Formular muss in puncto Datenschutz besonders sorgfältig bewertet werden.

Chatbots lernen voneinander

In der Welt der Behörden gibt es viele fachspezifische Inhalte. Daneben existieren aber noch viel mehr grundlegende Informationen, die für Bürgerinnen und Bürger von Bedeutung sind. Dazu gehören etwa die Öffnungszeiten und Standorte örtlicher Behörden sowie Zuständigkeitslisten. Solche übergreifend wichtigen Informationen aufzubauen und gesammelt verfügbar zu machen, ist ein wichtiges Ziel eines zentralen Services. Gleichzeitig werden die Erkenntnisse aus weiteren spezifischen Chatbot-Projekten laufend an den Basisdienst übertragen. So können zukünftige Projekte sie nutzen oder aus ihnen lernen. In Zukunft sollen übergreifende Informationen in jeden individuellen Chatbot einfach integriert werden können und sich autonom aktualisieren, um den projektspezifischen Erstellungsaufwand zu minimieren.

Behörden können von Chatbots lernen

Chatbots sind keine Einbahnstraße der Informationsvermittlung. Sie bieten auch Einblicke in die Interessen ihrer Nutzerinnen und Nutzer, in ihre Fragen und in die Themen, die sie interessieren. Die Dialoge des Chatbots werden ohne Bezug auf ihre Autoren protokolliert und statistisch ausgewertet. Sich verändernde Interessen der Nutzerinnen und Nutzer können somit nachverfolgt und das Informationsangebot durch entsprechende Erweiterungen fortlaufend verbessert werden. So sind auch Auswertungen zu falsch oder unbefriedigend beantworteten Fragen für eine stetige Qualitätsverbesserung des Chatbots wichtig. Chatbots bieten somit einen wichtigen Zugang zu den Interessen ihrer Nutzerinnen und Nutzer, der weit über Zugriffstatistiken von Webseiten hinausgeht.

WAS SIND DIE LEHREN AUS CHATBOT-PROJEKTEN IN DER ÖFFENTLICHEN VERWALTUNG?

Im ersten Jahr, in dem die ersten Prototypen von Chatbots für Bundesbehörden entwickelt wurden und im Testbetrieb liefen, wurden bereits viele Erfahrungen gesammelt. Sie helfen, zukünftige Anforderungen, aber auch typische Probleme bei der Konzeption und beim Einsatz von Chatbots zu identifizieren und zu lösen. Gleichzeitig zeigen sich sowohl neue Herausforderungen, denen sich der Basisdienst Chatbot stellen muss, als auch Möglichkeiten, die es zu nutzen gilt.

ERWARTUNGEN MANAGEN – CHATBOTS KÖNNEN NICHT ALLES

Ein äußerst wichtiger Aspekt in einem Chatbot-Projekt ist es, die Erwartungen der beauftragenden Behörde und der zukünftigen Nutzerinnen und Nutzer richtig einzuschätzen und unter Umständen auf sie einzuwirken.

Die Behörde erhofft sich natürlich einen möglichst hohen Nutzen von ihrem Chatbot. Am besten soll er Inhalte aus bestehenden Informationsquellen wie Webseiten extrahieren und ohne weiteren Aufwand dauerhaft einsetzbar sein. Bringen die Entwickler dann Themen wie die Wartung und Pflege des Chatbots ins Spiel, müssen sie sensibel vorgehen, da gerade hierfür oft die erforderlichen Ressourcen fehlen. Und dennoch: Die langfristige Entlastung der Behörde durch ihren Chatbot funktioniert nur, wenn er auch gepflegt wird.

Nutzerinnen und Nutzer hingegen erwarten umfassendes Wissen von Chatbots. Sobald sie merken, dass ein Chatbot nichts zu einem Thema sagen kann, setzt Ernüchterung ein. Wenn „er“

dann noch Fragen falsch versteht, wird daraus schnell Ablehnung – der Chatbot wird als dumm empfunden. Dies geschieht auch dann, wenn er redundante Information bereithält, die der Nutzer viel leichter über eine einfache Google-Suche finden könnte.

Bevor ein Chatbot-Projekt initiiert wird, sollten die Erwartungen an den Chatbot von allen Seiten klar definiert und erfasst sein. Was erwartet der Auftraggeber, wer gehört zur Zielgruppe, welche Inhalte, welchen Funktionsumfang und welche Reichweite soll der Bot haben? Sind die Informationen des Chatbots eine nützliche Zusatzinformation zu den bereits publizierten Inhalten?

Neben den inhaltlichen Aspekten ist es auch wichtig, projektspezifische Erwartungen zu deckeln. Chatbots entwickeln ihr Sprachverständnis ähnlich wie heranwachsende Menschen. Sie „hören zu“ und lernen aus ihren Fehlern. Dieses Verständnis wird laufend trainiert. Ebenso werden Inhalte immer wieder angepasst und aktualisiert. Es ist daher empfehlenswert, diese Aspekte des zukünftigen Betriebs bereits während der Konzeption mitzudenken, denn sie wirken sich direkt auf den Funktionsumfang aus.

CHATBOTS BRAUCHEN NAMEN

Nutzerinnen und Nutzer wollen wissen, mit wem sie es zu tun haben. Sprechen sie mit einer Maschine oder mit einem Menschen? Wie sieht ihr Gegenüber aus? Wie heißt er oder sie? Die Verwendung menschlicher Namen (wie BITZ oder LinA) verleiht Chatbots eine Identität. Ebenso ist die Verwendung von Bildern und Avataren hilfreich (siehe Abbildung 1). Sie schaffen Vertrauen aufseiten der Nutzerinnen und Nutzer. Dennoch ist es bei aller Personalisierung wichtig, dass Chatbots immer auch als solche erkennbar bleiben.

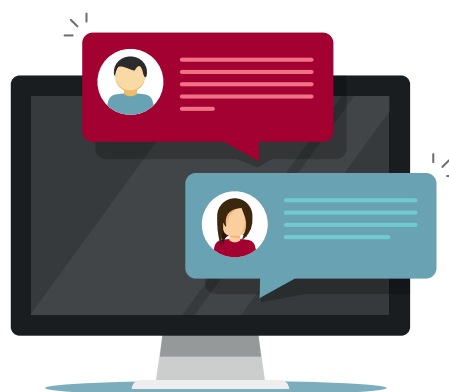
DER RICHTIGE TON MACHT'S

Für die Gestaltung der Dialoge benötigt das Chatbot-Redaktionsteam einen Anhaltspunkt, an dem sie sich orientieren können. Das Entwickeln der Dialoge des Chatbots geschieht mithilfe einer Persona, das heißt einer fiktiven Person, die für eine bestimmte Zielgruppe steht. Ähnlich werden auch die zukünftigen Nutzerinnen und Nutzer eine Art Persona in den Chatbot projizieren. Aber nicht nur die Ansprache ist in diesem Zusammenhang wichtig. Auch im Dialog sollten Chatbots relevante und äquivalente Antworten sicherstellen und spannende Unterhaltungen ermöglichen. Während einseitige, starre Dialoge mit falschen

und nicht passenden Antworten zu Frustration bei den Nutzerinnen und Nutzern führen, können verschiedene Varianten von Antworten die Akzeptanz erhöhen. Unterhaltsame Aussagen und empathische Antworten verbessern die Wahrnehmung bei (frustrierten) Nutzerinnen und Nutzer, gestalten Unterhaltungen sympathisch und natürlich. Eine kurze Beschreibung des Funktionsumfangs und des Zuständigkeits- oder Themenbereichs des Bots hilft, Nutzererwartungen proaktiv zu gestalten und direkt zu Dialogbeginn mit der Begrüßung transparent zu machen.

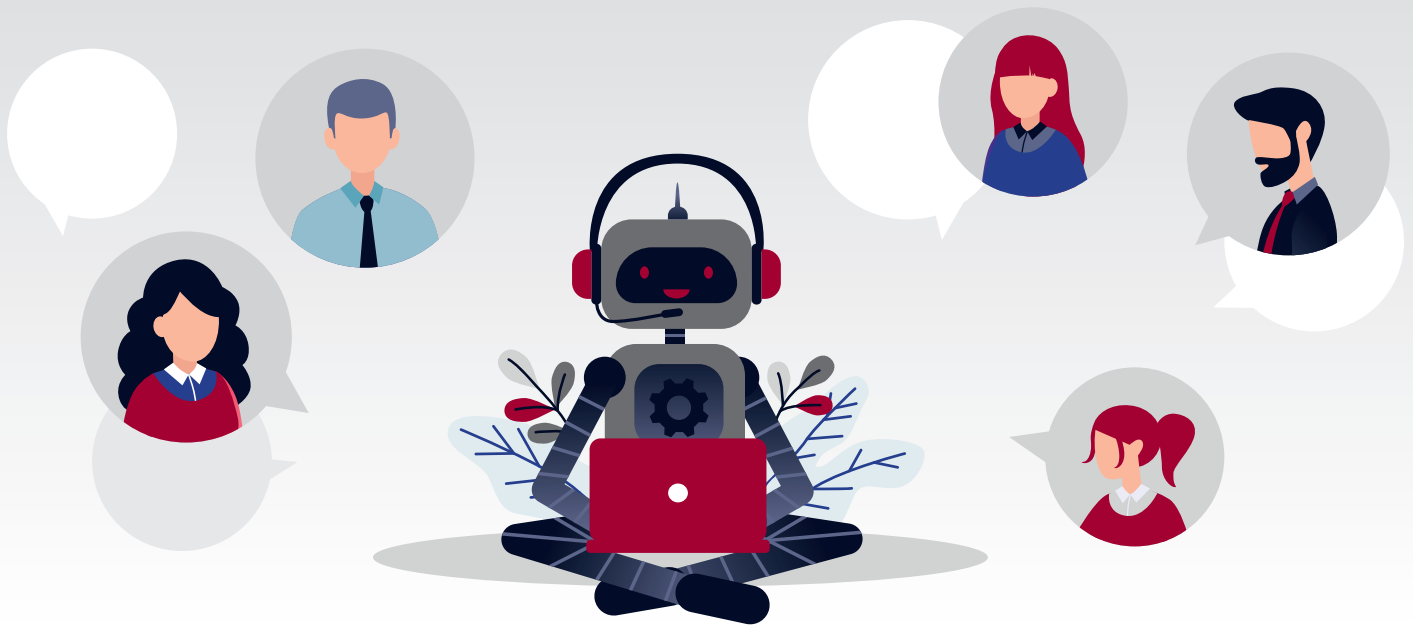
NUTZER WOLLEN MEHR ALS NUR INFORMATION

Durch den Einsatz der ersten Chatbots hat das Projekt festgestellt, dass es für die Nutzerinnen und Nutzer wichtig ist, die Antworten des Chatbots bewerten zu können. Das Ergebnis: Zu lange und zu umfangreiche Antworten in einer zu formellen Sprache werden häufig schlechter bewertet als eine kurze einfache und vor allem treffende Information. Dieses Nutzerfeedback kann die Behörde dazu nutzen, etwa die Informationen auf ihren Webseiten ständig weiter anzureichern.



CHATBOTS SIND UND BLEIBEN EINE HERAUSFORDERUNG

Als „dumm“ wahrgenommene Chatbots schädigen das Image ihrer Absender im selben Maße, wie schlaue Bots es aufzubauen helfen. Im Vergleich zu einer Webseite werden Chatbots als Persönlichkeiten wahrgenommen und nicht als neutrales Informationsmedium. Das kann im positiven Fall zu einer Identifikation führen, im schlechten aber auch zur Ablehnung. Der Dialog zwischen Mensch und Maschine hat immer auch einen emotionalen Aspekt. Deshalb ist es wichtig, den Chatbot immer auf einem aktuellen Stand zu halten. Zudem ist es wichtig, dass ein Chatbot nicht nur fachliche Informationen bereithält, sondern auch „Small Talk“ mit seinen Nutzerinnen und Nutzern führen kann – zumindest in begrenztem Umfang.



Um eben nicht „dumm“ zu wirken und um noch mehr Nutzen zu bieten, entwickeln sich Chatbots immer weiter. Ein Beispiel ist die „Natural Language Generation“. Dieser Begriff bezeichnet neuartige Methoden, die es einem Chatbot ermöglichen, selbst eine KI-gesteuerte Antwort zu formulieren, ohne dass Menschen diese Inhalte im Vorfeld manuell erstellen müssen. Das macht den Chatbot wesentlich flexibler und verleiht ihm mehr Eloquenz. Ebenso haben zukünftige Chatbots eine Art Gedächtnis. Hier speichern sie jeden Kontext und alle Informationen, die ihnen Nutzerinnen und Nutzer anvertrauen, um sie als Grundlage zur besseren Beantwortung weiterer Fragen anderer zurate ziehen können. Dieses Gedächtnis wird auch die Grundlage für Rückfragen bilden, die ein Chatbot den Nutzerinnen und Nutzern stellen kann, wenn ihm zur richtigen Beantwortung einer Frage noch weitere Angaben fehlen.

Auch sind miteinander vernetzte Chatbots, die bei speziellen Fragen der Nutzer an den jeweiligen Experten weiterleiten, ein wichtiges Forschungsfeld, das im Rahmen des Basisdienstes Chatbot angegangen wird. Vernetzte Chatbots greifen zusätzlich auf diverse externe Datenquellen zu, um zum Beispiel Fragen mit Ortsbezug anhand einer Geodatenbank beantworten zu können.

Die letzte große Herausforderung bei Chatbots stellt die heute noch kaum genutzte Fähigkeit dar, logische Rückschlüsse aus dem Gespräch mit ihren Nutzern ziehen zu können. Dazu werden Ontologien und Graphen sowie Subjekt-Prädikat-Objekt-Beziehungen genutzt, um mit dem sogenannten „Reasoning“ logische Rückschlüsse ziehen zu können. Ein einfaches Beispiel für

solch einen Rückschluss ist, dass eine männliche Person, die Kinder hat, automatisch auch ein Vater ist. So einfach dieser Rückschluss auf den ersten Blick wirkt, ist er doch sehr wirkmächtig im Gesprächskontext eines Chatbots.

FAZIT

In der öffentlichen Verwaltung eingesetzt, sind Chatbots Chance und Risiko gleichzeitig. Einerseits können sie deutlich in der Kommunikation mit den Bürgerinnen und Bürgern helfen. Andererseits besteht aus den ausgeführten Gründen die Gefahr, als nicht kompetent genug wahrgenommen zu werden. Wichtig ist es, die Schlüsselfaktoren zu kennen und umzusetzen: passendes Dialogdesign, redaktionelle Pflege und Vermittlung von wirklich relevanter Information im Kontext der Nutzerfragen.

Gleichzeitig zeigt sich, dass sich die öffentliche Verwaltung dieses Themas intensiv annimmt, sich den Herausforderungen stellt und über die Zeit eine wichtige und hilfreiche Erweiterung der Bürgerkommunikation etablieren wird. Die ersten Pilotprojekte zeigen bereits heute einen Ausschnitt der Möglichkeiten, die Chatbots zukünftig bieten können. Und sie machen neugierig auf mehr. ●



KONTAKT

Anfragen zum Thema „Basisdienst Chatbots“ bitte an chatbot@itzbund.de richten.

1 <https://www.id.bund.de/de/eservice/konto> (abgerufen am 24.02.2021).
2 Siehe: <https://www.stmgp.bayern.de/coronavirus/> (abgerufen am 24.02.2021).
3 <https://www.messengerpeople.com/de/corona-whatsapp-chatbot/> (abgerufen am 24.02.2021).
4 https://de.wikipedia.org/wiki/Amazon_Echo#Skills-Funktionserweiterungen (abgerufen am 24.02.2021).

BARRIEREFREIHEIT leichter umsetzen und **DOKUMENTIEREN**

Die Umsetzung von Barrierefreiheit bei allen Webseiten und Dokumenten, wie sie das Gesetz vorschreibt, erfordert entwicklungsbegleitend Spezialwissen und ein standardisiertes Ergebnisdokument. msg.BALM ist die Softwarelösung für Prüfung, Test und Dokumentation von Barrierefreiheit. Davon profitieren Sie im Rahmen eines Beratungsmandats mit msg. Oder Sie erwerben das Tool für eine langfristige Nutzung.



msg.BALM auf einen Blick:

- Alle gesetzlichen Anforderungen an einem Ort „ready-to-use“
- Status jederzeit transparent
- Alle Tests jederzeit wiederholbar
- Gesicherte Erfüllung der Richtlinien
- Einfache Bedienung
- Evaluierungsbericht inkl. Testergebnissen und Vorschlägen zur Steigerung der Qualität
- Höhere Qualität der Anwendungen

IHR ANSPRECHPARTNER



RICHARD PIELCZYK
Principal Project Manager
und Abteilungsleiter

0152 54679104

richard.pielczyk@msg.group



| von DR. CHRISTIAN KIEHLE

REGISTERMODERNISIERUNG – ZIEHT DEUTSCHLAND NACH?

Ein Kommentar zum rechtlichen Status quo der Registermodernisierung

Durch Umsetzung des Onlinezugangsgesetzes (OZG) müssen bis 2022 die 575 OZG-Leistungen online verfügbar sein. Eine vollständige Digitalisierung der OZG-Leistungen ist allerdings noch nicht erfolgt. Für diesen nächsten Digitalisierungsschritt liefert das lange geforderte Registermodernisierungsgesetz nun die (gesetzliche) Grundlage. Kritik gibt es vor allem aus Gründen des Datenschutzes und einer möglichen Verletzung von Persönlichkeitsrechten.

REGISTERMODERNISIERUNG – NOTWENDIG, ABER KOMPLEX

Die deutsche Verwaltung betreibt über alle föderalen Ebenen hinweg mehr als 200 Register. Diese dezentral entstandenen Verfahren sind in Art, Umfang, Qualität und Technologie von einer

ausgesprochenen Heterogenität gekennzeichnet. Schon seit vielen Jahren wird dieser Zustand kritisiert und eine umfassende Modernisierung angemahnt. Der Nationale Normenkontrollrat hat 2017 in einer umfangreichen Studie beschrieben, wie eine Registermodernisierung zur Digitalisierung der Verwaltung beitragen kann.¹ Dr. Johannes Ludwig (Vorsitzender des Nationalen Normenkontrollrats) bringt es in einer

Pressemitteilung auf den Punkt: „Ohne Registermodernisierung ist kein digitaler Staat zu machen.“² Die E-Government-Vorzeigegeneration Estland ist hier schon lange weiter, hat es jedoch auch mit einer nicht annähernd so komplexen und gewachsenen Registerlandschaft wie in Deutschland zu tun. Näher liegt da der Blick nach Österreich, wo



REGISTER: Datenbestand der öffentlichen Verwaltung, der notwendig ist, um Verwaltungsdienstleistungen zu erbringen. Darüber hinaus können sie für die amtliche Statistik (beispielsweise Anzahl neu gemeldeter Kraftfahrzeuge pro Jahr und Bundesland) oder den Zensus genutzt werden. Sie bieten oftmals die Grundlage, um politische oder administrative Entscheidungen treffen zu können. Bekannte Register für Bürger sind das Grundbuch, das örtliche Melderegister und das zentrale Fahrzeugregister.

das Thema Registermodernisierung bereits erfolgreich und datenschutzkonform umgesetzt wurde.³

Ein Register ist im Kern nichts anderes als ein behördlich geführter Datenbestand. Je nach zugrunde liegender Fachlichkeit (Ausländerzentralregister, Melderegister, Außenhandelsregister, Waffenregister etc.) ist er entsprechend sensibel und komplex aufgebaut. Dabei

spielt die Datenqualität eine entscheidende Rolle insbesondere, wenn Datensätze aus unterschiedlichen Registern (zum Beispiel Melderegister und zentrales Fahrzeugregister) miteinander verknüpft werden sollen. Hier gibt es vielfältige Fehlerquellen in den Basisdaten (Name, Geburtsort, Datumsangaben etc.), die durch Transkriptionsfehler, Formatfehler oder weitere Ungenauigkeiten bei der Datenübernahme entstehen können.

Im Zuge der fortschreitenden Digitalisierung ist eine automatisierte und digitale Übermittlung von Daten in Register und zwischen Registern wünschenswert. Erst diese digitale Übermittlung ermöglicht es, Geschäftsprozesse vollständig zu digitalisieren. Vor allem an der Schnittstelle zum Bürger und zur Bürgerin soll nach Möglichkeit das Once-Only-Prinzip – das heißt die Erhebung der Daten an nur einer Stelle – angewendet werden. Durch ein übergreifendes Identitätsmanagement, für das eine zentrale und eindeutige Identifizierung von Personendaten notwendig ist, können die Daten dann durch weitere Register (bei berechtigtem Interesse und unter Berücksichtigung des Datenschutzes) nachgenutzt werden. So können typische Anwendungsfälle wie Elterngeldbeantragung oder Ummeldungen bei Wohnortänderungen vollkommen digital umgesetzt werden und bedürfen keiner wiederholten Bereitstellung von Unterlagen durch die Bürgerinnen und Bürger.

Im Sinne der Digitalisierung behördlicher Leistungen ist eine Modernisierung der deutschen Registerlandschaft also zwingend notwendig. Aufgrund der dezentralen und über lange Jahre gewachsenen, heterogenen Struktur der Register ist diese jedoch als äußerst komplex einzustufen.

REGMOG LIEGT VOR – ZAHLREICHE BEDENKEN EBENFALLS

Seit September 2020 liegt nun der Gesetzesentwurf der Bundesregierung zum Registermodernisierungsgesetz (Reg-



ONCE-ONLY-PRINZIP: Zielt darauf ab, dass Bürger der Verwaltung und den Behörden Daten nur ein einziges Mal zur Verfügung stellen. Für die Bürger bedeutet dies eine Erleichterung, da sie Daten nicht bei verschiedenen Verwaltungsvorgängen wiederholt eingeben müssen. Für die Verwaltung ist die Zielsetzung, dass die Daten, die einmalig erhoben werden, auch nur einmalig gespeichert werden, unabhängig davon, in welchem Kontext die Bürger diese eingegeben haben.



MoG)⁴ – in voller Länge: „Entwurf eines Gesetzes zur Einführung und Verwendung einer Identifikationsnummer in der öffentlichen Verwaltung und zur Änderung weiterer Gesetze (Registermodernisierungsgesetz – RegMoG)“ – vor. In 22 Artikeln – wovon sich neunzehn Artikel mit Änderungen bestehender Gesetze befassen – wird die Modernisierung der

Register geregelt. Von besonderem Interesse ist dabei Artikel 1, der das Gesetz zur Einführung und Verwendung einer Identifikationsnummer in der öffentlichen Verwaltung beschreibt (Identifikationsnummerngesetz – IDNRG, siehe Infokasten). Um der Wichtigkeit des Themas Ausdruck zu verleihen, wird die Einrichtung einer Registermodernisierungsbehörde geregelt, die mit der übergeordneten Steuerung der Registermodernisierung betraut ist. Hierfür ist das Bundesverwaltungsamt vorgesehen, das hierzu eine Referatsgruppe „Registermodernisierung und Nachfragemanagementorganisation“ innerhalb der Abteilung „Digitalisierung, Registermodernisierung, Nachfragemanagementorganisation“ einrichtet.⁵

Kern – und an vielen Stellen Stein des Anstoßes – ist die Definition der Steuer-Identifikationsnummer als zentrales Ordnungsmerkmal zur einheitlichen Identifikation von Personen in allen Registern. Während der generelle Vorstoß eines Registermodernisierungsgesetzes von zahlreichen Branchenverbänden und sonstigen Stakeholdern begrüßt wird, so stößt der Einsatz der



IDENTIFIKATIONSNUMMERNGESETZ (IDNRG): Gesetz zur Einführung und Verwendung einer Identifikationsnummer in der öffentlichen Verwaltung definiert in siebzehn Paragraphen die Ziele des Gesetzes, beschreibt Aufgaben und Einrichtung einer Registermodernisierungsbehörde und detailliert Art, Umfang und Nutzung der ID. Es ist der eigentliche Kern des RegMoG und definiert die Einführung einer Identifikationsnummer (natürlicher Personen) als das Ordnungsmerkmal, das registerübergreifend genutzt werden kann, um Registerinformationen abzufragen. Die zugehörige Anlage definiert Register nach § 1 des Gesetzes (derzeit: 56).





TRANSPARENZ- UND SCHUTZMASSNAHMEN GREIFEN (VERMUTLICH) ZU KURZ

4-CORNER-MODELL: Architekturmodell zur kontrollierten Datenübermittlung über Vermittlungsstellen. Der Austausch von Daten zwischen einer öffentlichen Stelle und einer anderen öffentlichen Stelle geschieht nicht direkt (2-Corner-Modell) oder über eine zentrale Stelle (3-Corner-Modell), sondern indirekt über eine Infrastruktur, die durch zahlreiche Serviceanbieter zur Verfügung gestellt wird. Eine Nachricht (Corner 1) kann somit über Zugriffspunkte (Corner 2) versendet werden. Ein weiterer Zugriffspunkt (Corner 3) empfängt diese Nachricht und stellt diese einem Empfänger (Corner 4) zur Verfügung. Die Zugriffspunkte haben dabei keine Einsicht in die Nachricht, können aber prüfen, ob der gewünschte Datenaustausch ermöglicht oder abgelehnt werden soll, und protokollieren den Nachrichtenversand. Siehe hierzu auch European Commission 2015.⁸

Das Registermodernisierungsgesetz sieht bereits diverse Maßnahmen vor, um Grund- und Persönlichkeitsrechte nicht zu gefährden. Allen voran ist das 4-Corner-Modell zu nennen (siehe Infokasten) sowie eine Ende-zu-Ende-Verschlüsselung der Kommunikation zwischen Behörden. Darüber hinaus erweist sich die vermeintliche Schwäche verteilter Register aus Sicht der Informationssicherheit als potenzielle Stärke: Eine Zusammenführung umfassender Profile über die zentrale Steuer-ID würde die Abfrage zahlreicher Register erfordern, was als nicht praktikabel angesehen wird.

Positiv zu bewerten ist der § 17 RegMoG (Strafvorschriften, siehe Infokasten), der bereits auf gesetzlicher Ebene den Missbrauch der Identifikation unter Strafe stellt.

Darüber hinaus ist ein Datenschutzcockpit (siehe Infokasten) vorgesehen, das den Bürgern Transparenz zur Verwendung ihrer Daten bieten soll.

Steuer-ID als zentrales Ordnungs- und Identifikationsmerkmal auf heftige Kritik. Zentrale Kritikpunkte sind:

- Die Steuer-ID wurde einzig und allein für steuerliche Zwecke eingeführt. Das Grundrecht auf Datenschutz wird als gefährdet angesehen, wenn dies nun nicht mehr der Fall ist.
- Mit einer zentralen ID auf zum Teil sehr sensible Daten in staatlicher Verwahrung zugreifen zu können, löst verständlicherweise Bedenken bei Datenschützern aus. Der Verweis auf das Volkszählungsurteil des Bundesverfassungsgerichts von 1983 (BVerfG 1983⁹) fokussiert „die Befugnis des Einzelnen, grundsätzlich selbst über die Preisgabe und Verwendung seiner persönlichen Daten zu bestimmen“.
- Mit der Steuer-ID werden dem Missbrauch Tür und Tor geöffnet, um Personenprofile anlegen zu können oder unberechtigte Datenzugriffe durchführen zu können.

Ein Blick zu unseren österreichischen Nachbarn legt ein weiteres interessantes Modell offen. Österreich hat schon seit mehr als fünfzehn Jahren das Thema Registermodernisierung ernst genommen und vorangetrieben. Hier generiert die zentrale Datenschutzbehörde mittels kryptografischer Verfahren eine geheime Stammzahl aus der Zahl des Zentralen Melderegisters, die als Ordnungs- und Identifikationsmerkmal genutzt wird. Diese kann zu bereichsspezifischen Identitätskennzeichen (zum Beispiel für den Bereich Soziales) umgerechnet werden, um so Informationen aus Registern abzufragen. Die Übersetzung von Kennzahlen unterschiedlicher Bereiche ist ausschließlich über die Datenschutzbehörde möglich. Eine einfache Profilbildung durch Verknüpfung einer einzelnen ID ist bereichsübergreifend nicht möglich. Ein Missbrauch ist

daher nicht so leicht möglich, wie dies bei der im deutschen Modell vorgesehenen zentralen Steuer-ID theoretisch der Fall wäre. Es ist allerdings strittig, ob das österreichische Modell ohne erhebliche Gesetzesänderungen auf die Bundesrepublik Deutschland übertragbar wäre.⁷

Des Weiteren könnten durch eine Verknüpfung von Registern die einst streng anlassbezogen erhobenen Daten nun in anderen Kontexten wiederverwendet werden, vorausgesetzt, es werden keine technischen oder organisatorischen Regelungen eingebracht.



DATENSCHUTZCOCKPIT: Das Datencockpit soll laut RegMoG aufgebaut werden, um Transparenz zu vorgenommenen Datenübermittlungen zwischen Behörden zu ermöglichen. Damit soll über eine zentrale Stelle für Bürgerinnen und Bürger ersichtlich werden, wie die Identifikationsnummer genutzt wurde. Es wird über eine Änderung des Onlinezugangsgesetzes (OZG) in den §§ 10 und 11 definiert und soll als IT-Komponente im Portalverbund realisiert werden.

AUSBLICK

Die Registermodernisierung ist insbesondere im Kontext des Onlinezugangsgesetzes positiv zu sehen und stellt einen wichtigen nächsten Schritt auf dem Weg in eine digitalisierte, vernetzte Verwaltung dar. Die digitale Verknüpfung von Registern bietet einen echten Mehrwert, um Verwaltungsvorgänge beschleunigt und digital durchführen zu können.

Im europäischen Vergleich ist die Modernisierung der deutschen Registerlandschaft überfällig. Der Normenkontrollrat konstatiert, dass die „führenden E-Government-Nationen in und außerhalb Europas [...] diesen Weg bereits gegangen [sind]. Sie investieren schon seit mehr als zehn Jahren in die Modernisierung und Konsolidierung ihrer Registerlandschaft.“¹ Mit dem Registermodernisierungsgesetz liegt nun auch in Deutschland die gesetzliche Grundlage vor, dieses wichtige und komplexe Thema anzugehen. Gleichzeitig werden mit der Einrich-



tung einer Registermodernisierungsbehörde die Weichen gestellt, die Thematik und planvoll aufzusetzen. Um das Thema nachhaltig voranzutreiben, ist eine übergeordnete Steuerung ebenso notwendig wie ein entsprechender Finanzierungsrahmen. Gleichwohl gilt es, die nach wie vor bestehenden Vorbehalte gegen den Missbrauch der Daten auszuräumen und geeignete technische und organisatorische Maßnahmen zu definieren, um diesen Missbrauch von vornherein weitmöglichst zu erschweren. Die dezentrale Struktur der

deutschen Registerlandschaft, die Nutzung des 4-Corner-Modells, die Einrichtung eines Datencockpits zur Herstellung von Transparenz gegenüber dem Bürger sind gut überlegte, vertrauensbildende Maßnahmen. Dennoch gibt es kritische Stimmen, die davon ausgehen, dass die Nutzung der zentralen Steuer-Identifikationsnummer in der derzeit vorgesehenen Form keinen Bestand haben wird.⁷ Dies zu beurteilen wird jedoch die Aufgabe der Rechtsprechung in naher Zukunft sein. ●

- 1 Nationaler Normenkontrollrat (Hrsg.) (2017): Mehr Leistung für Bürger und Unternehmen: Verwaltung digitalisieren. Register modernisieren. Online: <https://www.normenkontrollrat.bund.de/resource/blob/72494/476004/12c91fffb877685f4771f34b9a5e08fd/2017-10-06-download-nkr-gutachten-2017-data.pdf> (abgerufen am 22.01.2021).
- 2 <https://www.normenkontrollrat.bund.de/nkr-de/aktuelles/registermodernisierung-in-angriff-genommen-bundesregierung-beschliesst-registermodernisierungsgesetz-17903022> (abgerufen am 22.01.2021).
- 3 Parycek, Pl.; Huber, V.; Hunt, S.; Novak, A.-S.; Thapa, B. (2020): Analyse der rechtlich-technischen Gesamtarchitektur des Entwurfs des Registermodernisierungsgesetzes. Online: <https://oeffentliche-it.de/documents/10181/14412/Analyse+der+rechtlich-technischen+Gesamtarchitektur+des+Entwurfs+des+Registermodernisierungsgesetzes> (abgerufen am 22.01.2021).
- 4 Deutscher Bundestag (2020): Entwurf eines Gesetzes zur Einführung und Verwendung einer Identifikationsnummer in der öffentlichen Verwaltung und zur Änderung weiterer Gesetze (Registermodernisierungsgesetz – RegMoG). Online: <https://dip21.bundestag.de/dip21/btd/19/242/1924226.pdf> (abgerufen am 22.01.2021).
- 5 https://www.bva.bund.de/DE/Das-BVA/Organisation/organisation_node.html (abgerufen am 22.01.2021).
- 6 BVerfG, Urteil des Ersten Senats vom 15. Dezember 1983 – 1 BvR 209/83 –, Rn. 1-215.
- 7 Kelber, U. (2020): Zentrales Personenkennzeichen: Angst vor dem allwissenden Staat. Online: <https://background.tagesspiegel.de/digitalisierung/zentrales-personenkennzeichen-angst-vor-dem-allwissenden-staat> (abgerufen am 22.01.2021).
- 8 European Commission (2015): Introduction to the Connecting Europe Facility – eDelivery building block. Online: https://ec.europa.eu/inea/sites/inea/files/building_block_dsi-introdocument_edelivery_v1_00.pdf (abgerufen am 22.01.2021).

Der verlässliche Weg zu

NACHWEISLICH MEHR SOFTWAREQUALITÄT

Wenn Sie in Ihrem Entwicklungsprojekt sicher sein wollen, dass die Software auch die qualitativen Anforderungen nach dem internationalen Standard der ISO/IEC 25010 erfüllt: msg.TEN liefert auf Grundlage von Fragenkatalogen, Checklisten und Workshops einen Katalog von nichtfunktionalen Anforderungen mit zugehörigen Metriken und Testverfahren.



msg.TEN

- ist ein erprobtes Verfahren, das von msg-Beratern effektiv und effizient geführt wird,
- bietet durch die Anlehnung an die Norm die Gewissheit, nichts vergessen zu haben,
- und enthält eine durchgehende Dokumentation.

Und wenn Sie den Nutzen unseres Ansatzes skalieren wollen, können Sie msg.TEN auch erwerben und selbst für all Ihre Projekte einsetzen.

Selbstverständlich können wir Sie auch durch den ebenfalls an die ISO/IEC 25010 angelehnten Audit-Prozess zur Sicherstellung der Qualitätsanforderungen begleiten.

SPRECHEN SIE UNS AN!



RICHARD PIELCZYK
Principal Project Manager
und Abteilungsleiter

0152 54679104

richard.pielczyk@msg.group



YOUCONNECT – EIN IT-SYSTEM FÜR BEHÖRDENÜBERGREIFENDEN DATENAUSTAUSCH

| von MARTIN MEYER, Leiter des Projekts YouConnect, Bundesagentur für Arbeit

Zur IT-Unterstützung des Datenaustauschs im Rahmen der Kooperationen der Rechtskreise SGB II, III und VIII zur Unterstützung von jungen Menschen am Übergang von Schule zu Beruf hat die Bundesagentur für Arbeit (BA) die IT-Lösung YouConnect erstellt. Das Projekt hat mit der IT-Lösung für einen behördenübergreifenden Datenaustausch in vielerlei Hinsicht Neuland betreten und musste unterschiedlichste Herausforderungen bewältigen.

KOOPERATIONEN ZUR UNTERSTÜTZUNG JUNGER MENSCHEN

In Deutschland gibt es einen sehr hohen Anteil an jungen Menschen unter 25 Jahren, die keine Berufsausbildung abgeschlossen haben und aktuell arbeitslos sind. Des Weiteren brechen viele junge Menschen jedes Jahr ihre Ausbildung oder ihr Studium ab. Es besteht die Gefahr, dass sie den Übergang von der Ausbildung ins Erwerbsleben nicht schaffen und folglich dort nicht Fuß fassen. Dies zieht sich dann in vielen Fällen durch die gesamte Erwerbsbiografie: Sie sind häufiger arbeitslos und von Langzeitarbeitslosigkeit bedroht. Im Umkehrschluss: Wenn der Übergang von der Schule ins Erwerbsleben erfolgreich gemeistert wird, ist das eine gute Grundlage für eine positive Erwerbsbiografie.

Mit dem Ziel, den Anteil der Ungelernten zu verringern, sind in den vergangenen Jahren zunehmend Jugendberufsagenturen gegründet worden. Dabei handelt es sich um Kooperationen der zuständigen Behörden und Einrichtungen für Grundsicherung, Arbeitsförderung und Jugendhilfe, die in unterschiedlichen Rechtskreisen geregelt sind, nämlich im zweiten (= Grundsicherung), dritten (= Arbeitsförderung) und achten (= Jugendhilfe) Sozialgesetzbuch. Denn wenn die Unterstützung eines jungen Menschen mit verschiedenen Maßnahmen (der unterschiedlichen Träger) individuell zugeschnitten und aufeinander abgestimmt erfolgt, ist der Erfolg wesentlich wahrscheinlicher. So reduziert die Kooperation und Betreuung eines Falles „Hand in Hand“ beispielsweise das Risiko, sie oder ihn auf dem Weg von Stelle A zu Stelle B zu verlieren.

Das erfordert allerdings, dass alle beteiligten Stellen von den – häufig mehrfachen – Problemen des jungen Menschen und von den Hilfestellungen wissen. Die entsprechenden Informationen müssen weitergegeben werden beziehungsweise den jeweiligen Bearbeitern vorliegen. Das bedeutet die Notwendigkeit eines Datenaustauschs über die Grenzen einzelner Behörden und Rechtskreise hinweg. Und dieser kann IT-gestützt vollständiger, frei von Verzögerungen und sicherer erfolgen als auf herkömmlichen Wegen. Das ist die Idee hinter YouConnect: ein gemeinsames IT-System, das die prozessuale und inhaltliche Zusammenarbeit erleichtert und zur qualitativen Weiterentwicklung der Jugendberufsagenturen beiträgt. Die Aufgabe, dieses System zu entwickeln und bis um 1. Januar 2021 bereitzustellen, hat die Bundesagentur für Arbeit (BA) übernommen.¹

DIE HERAUSFORDERUNGEN

Die Entwicklung einer solchen IT-Lösung für den rechtskreis- und behördenübergreifenden Datenaustausch stand vor mehreren Herausforderungen:

1. Beteiligung verschiedener Behörden und Einrichtungen

YouConnect soll nicht nur von der Bundesagentur für Arbeit (das heißt von den Arbeitsagenturen und den als gemeinsamen Einrichtungen organisierten Jobcentern) genutzt werden, sondern auch von den kommunalen Partnern, nämlich sowohl den kommunalen Jobcentern als auch den Institutionen des SGB VIII, der Jugendhilfe. Das ist nicht nur rechtlich Neuland, denn die Beteiligten haben jeweils eigene Anforderungen an ein solches System.



LISA – EIN FALLBEISPIEL

Lisa ist siebzehn Jahre alt, Schülerin vor dem Mittleren Schulabschluss, und schwanger. Weil sie außerdem in ihrem Elternhaus von Gewalt bedroht ist, geht sie zum Jugendamt. Im Gespräch mit ihrer Betreuerin kommen beide überein, dass sie zu Hause ausziehen sollte. Natürlich verfügt sie jetzt und auf absehbare Zeit nicht über die Mittel, um eine eigene Wohnung zu finanzieren. Wichtiger ist auch, dass sie die Schule und später eine Ausbildung abschließt. Das Jugendamt nimmt daher Kontakt mit dem zuständigen Jobcenter auf und hilft gegebenenfalls auch bei der Antragstellung.

Mit dem Zugriff auf ein gemeinsames Informationssystem weiß die Bearbeiterin oder der Bearbeiter im Jobcenter von Lisa von ihrer Problemlage, vom Kontakt mit dem und der Empfehlung des Jugendamtes, ohne dass Lisa ihre gesamte Geschichte noch einmal erzählen muss oder die nötigen Unterlagen erst schriftlich angefordert werden müssen.²

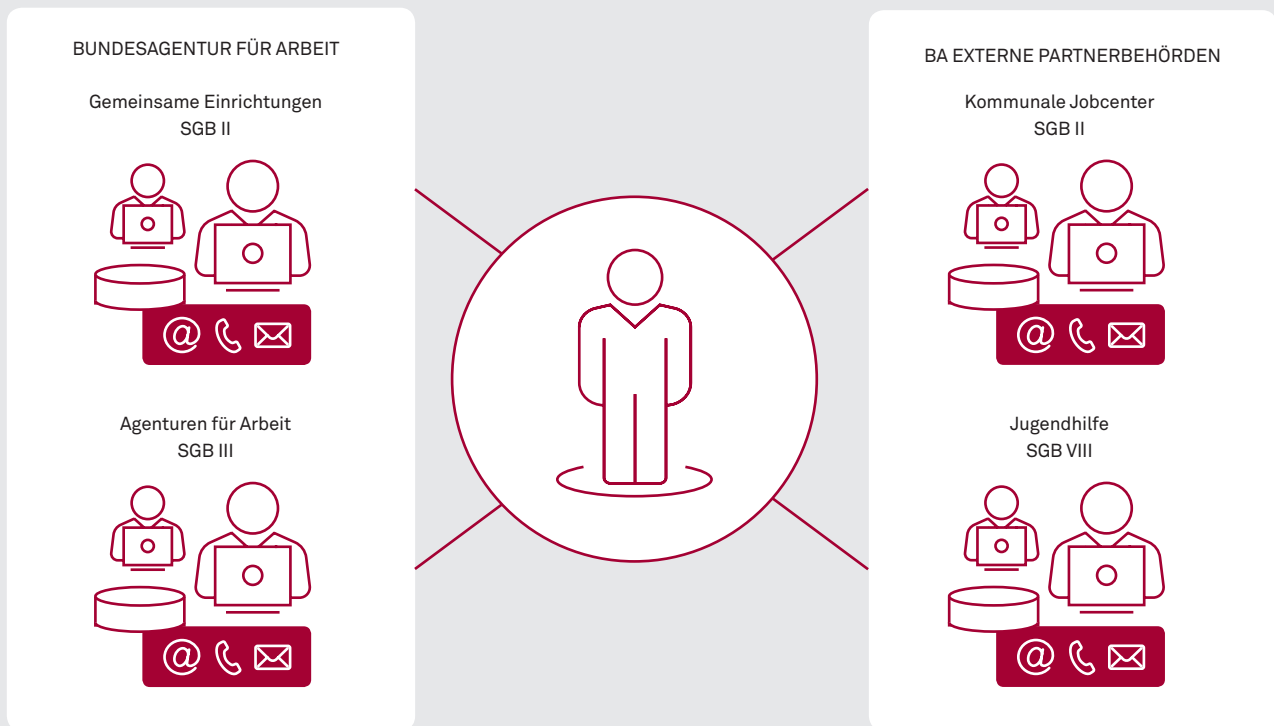


Abbildung 1: Um junge Menschen bedarfsgerecht auf ihrem Weg zu begleiten, ist häufig die Zusammenarbeit mehrerer Behörden erforderlich.

Die Bundesagentur für Arbeit ebenso wie andere Behörden haben keine Erfahrungen mit vergleichbaren IT-Projekten zu einer derart weitreichenden und verteilten behördenübergreifenden Zusammenarbeit.

2. Unterschiedliche Voraussetzungen bei den Beteiligten

Die Bedürfnisse und Erwartungen der unterschiedlichen Behörden und Rechtskreise müssen abgestimmt und berücksichtigt werden. Die Rahmenbedingungen und Vorgehensweisen unterscheiden sich teils stark. Es gibt einerseits Jugendberufsagenturen/Kooperationen, die alle an einem Ort, teilweise sogar in einem Gebäude untergebracht sind (One Stop Government), andererseits – vor allem in Flächenländern – sehr verteilte Standorte haben. Auch innerhalb der verschiedenen Institutionen gibt es sehr unterschiedliche Aufteilungen von Zuständigkeiten und Zusammenarbeit. Zentrale und allgemeingültige Vorgaben gibt es kaum. Alle Ausprägungen der Kooperationen sollen durch das IT-System unterstützt werden.

3. Datenschutz und Informationssicherheit

Es bestehen besonders hohe Anforderungen an Datenschutz und rechtliche Vorgaben. Insbesondere für die Jugendämter, in deren Kontext häufig sensible Daten aus dem persönlichen Lebensbe-

reich ausgetauscht werden, gelten Spezialvorschriften aus dem SGB VIII. Insgesamt müssen durch das neue IT-System sämtliche Aspekte für Datenschutz, Datensicherheit und Betroffenenrechte im rechtskreisübergreifenden Datenaustausch entsprechend der neuen DSGVO für alle beteiligten Rechtskreise erfüllt werden.

ANSÄTZE ZUR LÖSUNG

Um den Interessen der beteiligten Rechtskreise gerecht zu werden, wurde im ersten Schritt ein Beirat ins Leben gerufen. Er setzt sich aus Vertretern des Bundesministeriums für Arbeit und Soziales, des Bundesministeriums für Familie, Senioren, Frauen und Jugend, des Deutschen Landkreistages, des Deutschen Städtetages zusammen. Weiterhin für die Länder aus den Vorsitzenden der Arbeitsgruppen zentrale IT (Hamburg) und Eingliederung (Niedersachsen) sowie aus Teilnehmern vonseiten der Bundesagentur für Arbeit. Der Beirat hat wesentlich dazu beigetragen, dass die Sichtweisen und spezifischen Bedarfe der verschiedenen Rechtskreise berücksichtigt werden konnten. Er hat Partner für die Entwicklung und Pilotierung identifiziert, deren Rückmeldungen in die agile Entwicklung der Lösung eingeflossen sind und die die Erstellung von Schulungs- und Informationsmaterialien unterstützen.

Um die dem IT-Systemhaus der BA, das YouConnect entwickelt, wenig bekannten Arbeitsabläufe in den kommunalen Jobcentern und vor allem der Jugendhilfe zu berücksichtigen, wurden die Anforderungen der aus BA-Sicht externen Partner im Rahmen von gemeinsamen Anwender-Workshops, Vor-Ort-Besuchen und Hospitationen in Jugendberufsagenturen erhoben beziehungsweise mit den zukünftigen Anwenderinnen und Anwendern gemeinsam erarbeitet. Auch während der Umsetzungsphase wurden die Anwenderinnen und Anwender aller Rechtskreise weiter aktiv eingebunden. In Präsenz-Workshops – und von Beginn an zur Vermeidung von Reisekosten und -zeiten auch in Online-Sessions – wurde der aktuelle Stand wiederholt vorgestellt und das Feedback der Anwenderinnen und Anwender eingeholt. Immer wieder wurden aufgrund der Rückmeldungen der Anwenderinnen und Anwender Anpassungen vorgenommen.

Um den diversen technischen Voraussetzungen gerecht zu werden, wird YouConnect technisch in zwei unterschiedlichen Ausprägungen angelegt: Während das System für die Anwenderinnen und Anwender der BA medienbruchfrei in einem vorhandenen Fachverfahren (VERBIS) erreichbar ist, steht YouConnect den BA-externen Anwenderinnen und Anwendern über eine Web-Anwendung mit neutralem Design zur Verfügung. Innerhalb von

YouConnect wird eine gemeinsame Fallakte geführt, auf die die beteiligten Institutionen im Rahmen ihres jeweiligen gesetzlichen Auftrags und mit entsprechend eingeschränkter Sicht zugreifen können. Dazu können auf BA-Seite die Daten aus der Fachanwendung über eine Schnittstelle direkt übernommen werden. Die anderen Einrichtungen müssen die Daten, die in der Zusammenarbeit benötigt werden, in YouConnect erfassen, ebenso wie den Status ihrer jeweiligen Aufgabe. YouConnect stellt eine Standard-schnittstelle zur Verfügung, über die die Primärsysteme der externen Partner angebunden werden können.

In der Verwendung von YouConnect sind alle Beteiligten gleichberechtigte Partner. Jegliche Institution kann einen Fall anlegen, den Prozess der Zusammenarbeit starten und einen Fall beziehungsweise eine Aufgabe einer anderen zuständigen Institution zuweisen, die automatisch darüber benachrichtigt wird. Ebenso können die BA-externen Partnerinstitutionen über eine webbasierte Partner- und Mitarbeiterverwaltung die für YouConnect zugelassenen Anwenderinnen und Anwender administrieren, also anlegen, Passwort zukommen lassen, aktualisieren oder löschen. Darüber hinaus erlaubt das System beispielsweise die Erfassung und Verwaltung von Personen- und Falldaten, die Verwaltung gesetzlicher Vertreter und anderer Beteiligter, eine Personensuche,

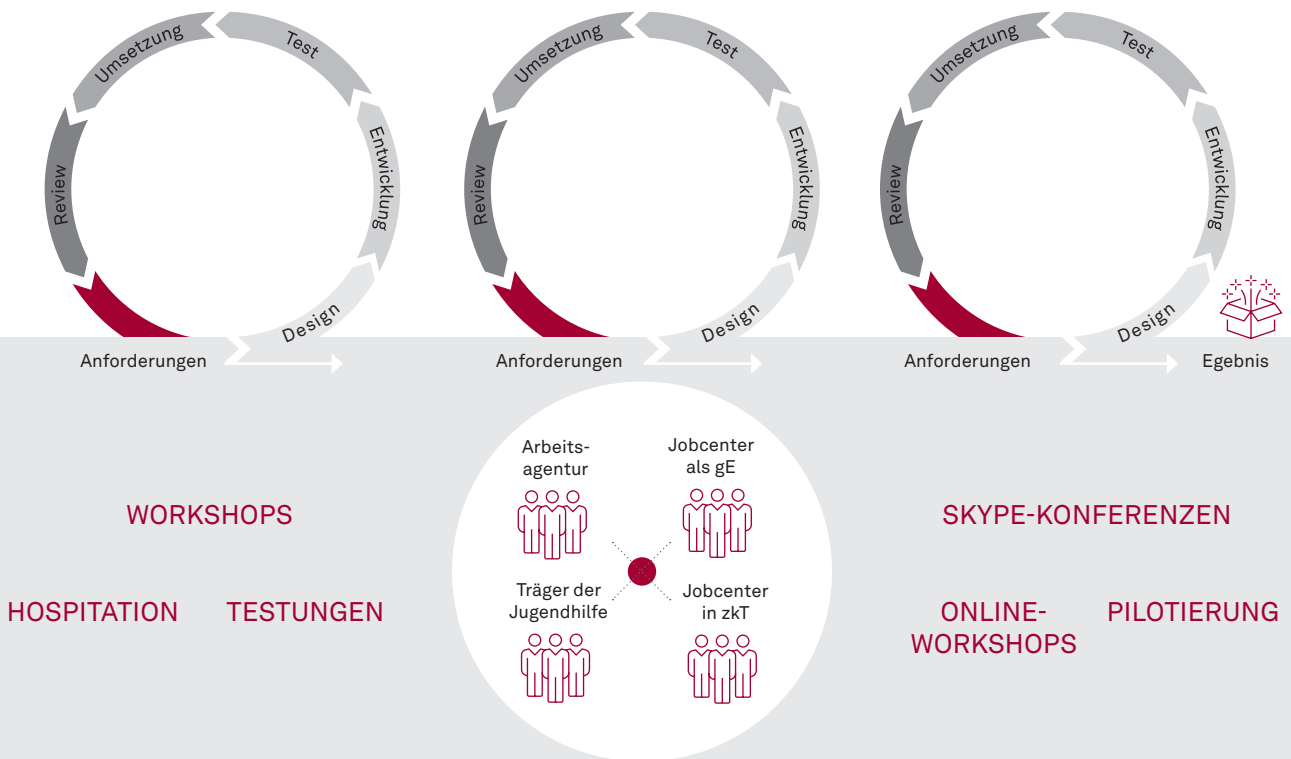


Abbildung 2: Immer wieder wurden aufgrund der Rückmeldungen der Anwenderinnen und Anwender Anpassungen vorgenommen.

die Erfassung biografischer Informationen und von Informationen über die beteiligten Institutionen. In der Fallkommunikation werden die Kommunikationsdaten der Beteiligten sowie der Status der Bearbeitung angezeigt und E-Mails etwa bei Statusänderungen automatisiert versendet.

Im Ergebnis ist mit YouConnect ein einfach zu bedienendes, intuitives IT-System entstanden, das die Bedürfnisse aller Partner berücksichtigt und folglich eine hohe Akzeptanz bei den Institutionen und den Anwenderinnen und Anwendern hat. Neue Anforderungen oder Anpassungen werden fortlaufend integriert.

Außerdem werden im Laufe der Pilotierung sukzessive weitere Funktionalitäten entwickelt und für den Echtbetrieb freigeschaltet. Dazu gehört beispielsweise die Umsetzung der Option, eine Fallbesprechung anonym, ohne Nennung von Namen durchzuführen, um die Inhalte aus den Bereichen der persönlichen Lebensführung zu schützen. Durch diese Funktionalität können mögliche Unterstützungsleistungen vorgeklärt werden, ohne dass personenbezogene Daten ins Spiel kommen.



PILOTIERUNG MIT HINDERNISSEN – DER CORONA-EFFEKT

Seit 1. September 2020 wird YouConnect – nach coronabedingten Verzögerungen – von deutschlandweit sechzehn Kooperationen im Echtbetrieb erprobt. Die pilotierenden Standorte bestehen aus einem Mix aus Kooperationen in Großstädten, in Kleinstädten und in Flächenländern, mit unterschiedlich hoher Arbeitslosigkeit und damit unterschiedlich kritischen Situationen auch für die jungen Menschen. Der Informationsaustausch zwischen den beteiligten Partnern erfolgt nun im Echtbetrieb über das IT-System YouConnect.

Die Pilotierung dient dem Zweck, die Funktionalität zu erproben, Fehler zu identifizieren und die korrekte Abbildung der fachlichen Prozesse vor Ort zu verifizieren. Regelmäßige Rückmeldungen erfolgen in telefonischen Interviews und in Form von schriftlichen Statusberichten. Die Ergebnisse aus den Feedbacks werden im Rahmen der agilen Entwicklung priorisiert und größtenteils bereits während der viermonatigen Pilotierung im IT-System angepasst.

Abbildung 3: YouConnect ist ein intuitives, leicht bedienbares IT-System.

Bis zum Start der Pilotierung wurde jedoch einiges umgeplant. Der Shutdown im März 2020 machte die eigentlich für April bis Juni 2020 vorgesehene Erprobung unmöglich, nicht nur weil die Vorbereitungen und die Betreuung der Pilotierung selbst den Pandemiebedingungen angepasst werden mussten, sondern vor allem, weil das Tagesgeschäft der pilotierenden Kooperationen nicht wie gewohnt und im vollen Umfang lief, beziehungsweise läuft und laufen wird.

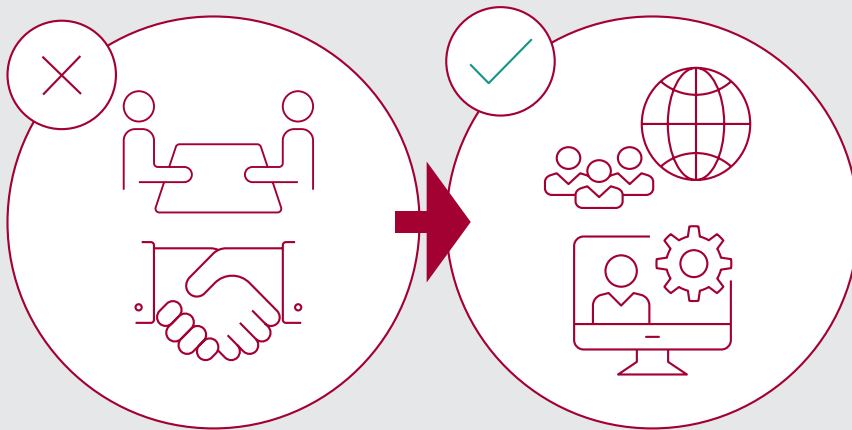


Abbildung 4: Durch die pandemiebedingten Kontaktbeschränkungen wurden Anpassungen am Vorgehen der Pilotierung und auch am IT-System erforderlich.

Datenschutz der BA und dem BMAS wurde eine digitale Option entwickelt und integriert: Über einen Weblink und mit einer eindeutigen Identifizierung können die jungen Menschen beziehungsweise ihre gesetzlichen Vertreter oder Vertreterin dem Datenaustausch zustimmen. Für eine schnelle Umsetzung dieser Funktionalität wurden die bereitzustellenden Leistungsmerkmale (Features) im Rahmen des agilen Vorgehens umpriorisiert.

Die zur Befähigung der Beschäftigten geplanten Vor-Ort-Termine werden durch Online-Workshops ersetzt, bei denen die zukünftigen Anwenderinnen und Anwender nach einer Eingangspräsentation Übungsaufgaben lösen, die in einem Schulungssystem mit Web-Oberfläche zur Verfügung stehen. Und statt der geplanten Betreuung der Pilotstandorte vor Ort werden sie mittels elektronischer Abfragen der Situation, Telefoninterviews und Statusberichten betreut.

Darüber hinaus gibt es coronabedingt Anpassungen am IT-System, die noch im Laufe der Pilotierung umgesetzt werden: Um sensible Informationen aus dem Bereich der persönlichen Lebensführung auszutauschen, bedarf es einer schriftlichen Einwilligung der jungen Menschen beziehungsweise deren gesetzlicher Vertreter oder Vertreterin. Ursprünglich hat das IT-System vorgesehen, die Einwilligungserklärung auszudrucken und dann unterschrieben in Papierform abzulegen. Innerhalb von YouConnect sollte vermerkt werden, dass und wo die Einwilligung vorliegt. Der Publikumsverkehr ist in den Institutionen jedoch eingeschränkt. Um die jungen Menschen, deren Situation sich durch die Auswirkungen der Pandemie im Allgemeinen eher verschlechtert hatte, weiter sinnvoll betreuen zu können, muss nun eine Alternative gefunden werden. In Zusammenarbeit mit der Stabstelle

FAZIT

Die Erprobung des Systems ist trotz der pandemiebedingten erschwerten Rahmenbedingungen ein Erfolg, nicht zuletzt, weil es gelungen ist, mit YouConnect ein intuitives, leicht bedienbares IT-System umzusetzen. Das System wird von den Pilotierungsteilnehmerinnen und -teilnehmer positiv bewertet. Es erleichtert ihnen die Arbeit in den Kooperationen. YouConnect trägt zu einer Weiterentwicklung der gemeinsamen Prozesse der Jugendberufsagenturen bei und erhöht die Transparenz bei rechtskreisübergreifender Fallarbeit. Die Qualität der Zusammenarbeit erhöht sich. Davon profitieren Institutionen und junge Menschen. Durch das hohe Engagement der Beschäftigten in den pilotierenden Kooperationen und im Projekt war und ist es möglich, trotz der Einschränkungen sowohl die Befähigung in Online-Workshops als auch die Fernbetreuung während der Pilotierung erfolgreich zu meistern. ●

- 1 Grundlage dafür ist die Regelung in § 368 Absatz 2a SGB III.
- 2 <http://www.arbeitsagentur.de/youconnect> (abgerufen am 30.03.2021).

ZUR PERSON



Martin Meyer ist in der Zentrale der Bundesagentur für Arbeit (BA) in Nürnberg tätig. Er verfügt über langjährige Erfahrung im Geschäftsbereich Arbeitsmarkt (Integration und Beratung) mit dem Schwerpunkt Jobcenter und ist hier für die fachliche Weiterentwicklung von Geschäftsprozessen und IT-Fachverfahren verantwortlich.





PROJECT OFFICE AS A SERVICE – EIN LANGER WEG SEIT GANTT

Projektunterstützung
maßgeschneidert

| von **BERND HAHN** und **ULRICH BERGHAUS**

Projektmanagement hat eine lange Geschichte. Instrumente wie das Gantt-Chart¹ sind über hundert Jahre alt, und obwohl der Kern des Projektmanagements – die Vereinbarung des magischen Dreiecks „Lieferobjekte, Termine und Kosten“ – über die Jahrzehnte gleichgeblieben ist, verändert sich das Projektmanagement stetig.

NEUE RAHMENBEDINGUNGEN ERMÖGLICHEN NEUE LÖSUNGEN

Digitalisierung, Agilität oder Managementansätze wie Servant Leadership² haben ebenso einen Einfluss auf das Projektmanagement wie auf die Rolle des Projektmanagers.

Hinzu kommen neue Konzepte, die den Projektmanager dabei unterstützen sollen, die vereinbarten Liefergegenstände bei gleichzeitiger Einhaltung der geplanten Kosten, der Zeit und unter Erreichung der gewünschten Qualität bereitzustellen. Eines dieser Konzepte ist das „Project Office as a Service“ (POaaS).

PROJECT OFFICE VS. PROJECT MANAGEMENT OFFICE

In den verschiedenen Projektmanagementstandards (PMI, PRINCE2, IPMA/GPM) sind Project Office (PO) oder Project Management Office (PMO) nicht eindeutig definiert. In diesem Artikel wird die Definition gemäß GPM verwendet.

Entsprechend der GPM „unterstützt [das Project Office] das Projekt bzw. den Projektmanager in administrativen Tätigkeiten für die Dauer der Projektlaufzeit. Dieses Büro ist dem Projektmanager zugeordnet und wird dediziert für das betreffende Projekt angelegt“.³

Das PO ist nicht mit dem PMO zu verwechseln. Das PMO ist gemäß GPM „eine permanente Einheit, die für die Etablierung, Implementierung und Fortentwicklung des Projektmanagementsystems zuständig ist“. Damit ist das PMO für die Strategie im Projektmanagement eines Unternehmens oder einer Behörde zuständig.

DAS PROJECT OFFICE

Das Project Office (PO) ist eine dem Projektmanager untergeordnete projektinterne Stabsstelle oder Projekteinheit. Sie unterstützt den Projektmanager bei der Ausübung seiner Tätigkeiten. Im Regelfall übernimmt das PO primär administrative Tätigkeiten, wie zum Beispiel die Pflege des Projekthandbuchs, On- und Offboarding der Projektmitarbeiter, Dokumenten- oder Meeting-Management. Die steuernden Tätigkeiten verbleiben im Kompetenzbereich des Projektmanagers. Er verantwortet weiterhin die Projektzielerreichung.

In kleineren Projekten übernimmt der Projektleiter in der Regel alle Projektmanagementtätigkeiten selbst. In größeren Projekten oder Programmen – wenn die Gesamtausübung der Projektmanagementtätigkeiten die Kapazität des Projektleiters überschreitet – wird häufig ein festes Project Office etabliert. Dabei werden die benötigten Ressourcen für das Project Office von Beginn an fest eingeplant, davon ausgehend, dass die Aufwände gleichbleibend über die Projektlaufzeit benötigt werden.

SCHWANKENDER PM-AUFWAND FÜHRT ZU RESSOURCEN-PROBLEMEN

Jedoch kann der Aufwand für das Managen eines Projekts über die Dauer der Projektlaufzeit schwanken (siehe Abbildung 1). Die Gründe liegen in den zu erledigenden Aufgaben der unterschiedlichen Projektphasen. So sind für den Projektmanager die Phasen der Projektinitialisierung und des Projektabschlusses (Abnahme von Liefergegenständen) meist aufwendiger als andere Phasen. Hinzu kommt, dass es in der Realität immer wieder zu Abweichungen vom Projektplan kommt. Unerwartete, unplanbare Ereignisse erfordern hierbei die volle Aufmerksamkeit und das schnelle Handeln des Projektmanagers. In diesen Phasen führen Standardtätigkeiten, wie das Erstellen eines Statusberichts, schnell zur Überlastung des Projektleiters.

Der schwankende Projektmanagementaufwand hat einen direkten Einfluss auf die Auslastung eines Project Offices. In Phasen mit geringem Projektmanagementaufwand sind im Project Office

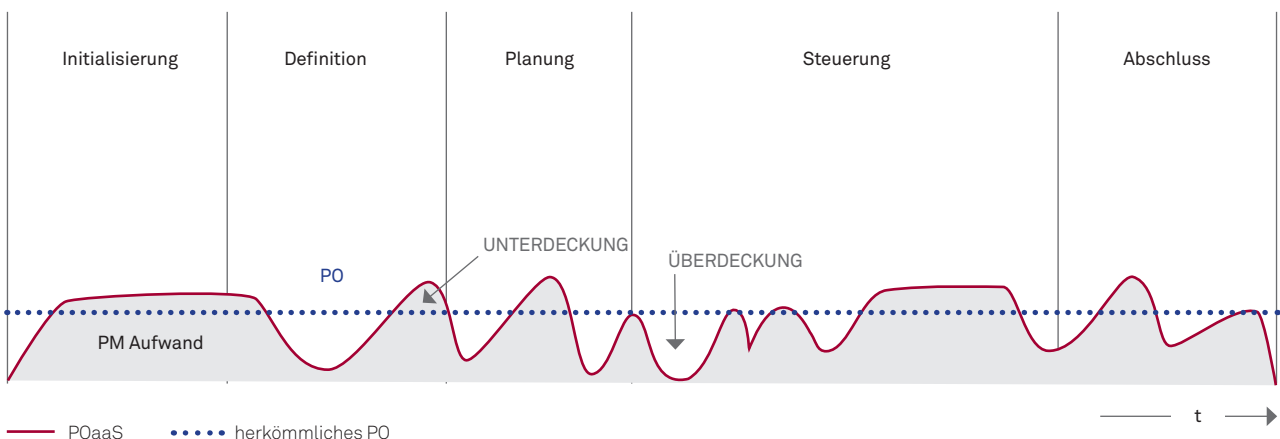


Abbildung 1: Ressourceneinsatz bei herkömmlichen PO und bei POaaS

Kapazitäten frei, seine Auslastung ist nicht gewährleistet. Das wirkt sich negativ auf die Projektkosten aus, da unproduktive Personalkosten anfallen. Dem stehen die Projektphasen mit hohem Managementaufwand gegenüber. In diesen Phasen ist das Project Office überlastet, was sich häufig auf die Qualität der Arbeit, das Einhalten von Terminen und auf die Zufriedenheit der Mitarbeiterinnen und Mitarbeiter auswirkt.

Hinzu kommt, dass in den Unternehmen und Behörden meist eine Vielzahl von Projekten gleichzeitig stattfindet – von kleinen Projekten bis hin zu großen Programmen. Je nach Projektgröße liegt der Schwerpunkt der PO-Tätigkeiten an anderer Stelle. Bei kleinen Projekten kommt es schnell zu einer Überlastung des Projektmanagers, da beispielsweise das Erstellen von Berichten „immer zum falschen Zeitpunkt“ kommt. In großen Projekten oder Programmen wiederum muss auf die Auslastung der Mitarbeiter geachtet werden. Nicht zu vergessen die Projekte, die zwischen Kleinprojekt und großem Programm liegen. Hier stellt sich dem Projektmanager häufig die Frage: „Etabliere ich ein PO oder nicht?“

DIE LÖSUNG: POaaS

Die Lösung für dieses Dilemma ist das POaaS (Project Office as a Service). Mit der Einführung eines POaaS kann das klassische PO im Projekt entfallen. Projektmanager müssen nicht mehr abwägen, ob sich die Etablierung eines PO rechnet und die Auslastung gewährleistet ist. Jeder Projektmanager kann flexibel, bedarfsgerecht und individuell PO-Leistungen als Service beziehen.

Um das zu gewährleisten, wird ein Project Office zentral als Stabstelle eingerichtet. Alternativ kann auf einen externen Dienstleister zurückgegriffen werden, der die entsprechenden Services bedarfsgerecht anbietet.

Das POaaS bietet alle Tätigkeiten eines Project Offices als Services, auf Basis fest definierter Leistungen an. Vom Projektmanager können die Leistungen flexibel gebucht werden. Um die Qualität der PO-Services zu gewährleisten, verfügt jeder Service über fest definierte Qualitäts- und Abnahmekriterien.

Diese Organisationsform ermöglicht es dem Projektmanager, bedarfsgerecht PO-Leistungen zu beziehen, und zwar dann, wenn sie tatsächlich benötigt werden. Damit werden die Schwankungen im Projektmanagement ausgeglichen. Es entsteht ein positiver Effekt bezüglich der Kosten und der Belastung der Projektmitarbeiterinnen und -mitarbeiter.

LEISTUNGEN EINES POaaS

Um die Organisationsform des POaaS effektiv zu nutzen, werden hier Leistungen zusammengefasst, die in Projekten meist nur temporär benötigt werden. Dazu gehören – neben den klassischen Tätigkeiten der Projektassistenzen zur Erstellung von Statusberichten, Protokollen etc. oder das On- und Offboarding von Mitarbeitern – auch Tätigkeiten wie das Risiko- und Stakeholder-Management oder im agilen Umfeld auch die Funktion eines Scrum-Masters. Um diese Tätigkeiten effektiv, effizient und qualitativ hochwertig durchzuführen, wird im POaaS entsprechend qualifiziertes Personal, unter anderem ausgebildete Projektmanager, eingesetzt. Die Erfahrung zeigt, dass PO-Mitarbeiter aufgrund ihrer Tätigkeit in der Regel gewissenhafter und weniger risikofreudig als Führungskräfte (wie zum Beispiel Projektleiter) sein müssen und sind.

POaaS ALS DIENSTLEISTER

Wie bereits beschrieben, wird das Project Office zentral als Stabstelle etabliert und ist somit eine projektexterne und projektübergreifende Organisation. Durch die Besetzung mit erfahrenen und gut ausgebildeten Mitarbeitern werden Projektmanager bei ihren Tätigkeiten professionell unterstützt. Mit entsprechenden Schulungen werden alle Mitarbeiter des POaaS ständig fortgebildet. So steht stets bestens geschultes Personal zur Verfügung.

Als zentraler Dienstleister im Unternehmen bietet das POaaS seine Tätigkeiten und sein Know-how vielen Projekten an. Dadurch entsteht auf der Seite des POaaS ein Skaleneffekt, der schwankende Bedarfe ausgleicht. Das führt zu einer gleichmäßigen Auslastung der Ressourcen im POaaS-Pool.

POaaS UND QUALITÄT

Im Idealfall arbeitet das POaaS eng mit dem PMO des Unternehmens zusammen und wendet die vorhandenen Projektmanagementstandards, zum Beispiel beim Erstellen des Projekthandbuchs, an. Damit unterstützt das POaaS nicht nur die Projekte, sondern auch das PMO bei der Implementierung und Etablierung von Projektmanagementstandards. Dies führt zu einer kontinuierlichen Qualitätssteigerung der Projektmanagementtätigkeiten über alle Projekte hinweg und ermöglicht es, einen unternehmensweiten Projektmanagementstandard einzuführen und zu verbessern (Erreichen von Projektmanagement Excellence).

Wird das POaaS entsprechend aufgebaut, verfügt es über eine Vielfalt an erfahrenen PO-Mitarbeitern. Innerhalb des POaaS können somit leichter „Good Practices“ für das Unternehmen beziehungsweise die Behörde entwickelt werden, die die Abwicklung von PO-Services kosten- und zeiteffizienter ermöglichen sowie qualitativ verbessern.

POAAS – EINE LÖSUNG FÜR ALLE?

Auf den ersten Blick bietet das POaaS nur Vorteile für alle Beteiligten. Allerdings drängt sich die Frage auf, wieso das POaaS bis heute keine weitverbreitete, standardisierte Organisationform ist. Dies lässt sich primär durch die folgenden vier zu meistern Herausforderungen erklären:

Digitalisierung

Das Konstrukt des POaaS erfordert eine IT-Infrastruktur, die eine schnelle, transparente Beauftragung sowie Abwicklung der PO-Services ermöglicht – im Idealfall unter Berücksichtigung der definierten Qualitäts- und Abnahmekriterien. Technische Lösungen für diese Anforderungen, wie zum Beispiel *Jira Service Desk*, gibt es erst seit Kurzem.

Reifegrad

Das POaaS erfordert einen gewissen, durchaus hohen Reifegrad der Organisation in Bezug auf Projekt- und Prozessmanagement. Eine POaaS-Organisation erfordert fest etablierte, standardisierte Prozesse und Projektmanagementstandards. Nur dann ist es dem POaaS möglich, die Projekte bestmöglich in hoher, gleichbleibender Qualität mit PO-Services zu versorgen, sowie kontinuierliche Qualitätsverbesserungen zu erzielen.

Initialer Aufwand

Der Wechsel vom klassischen PO zu einer POaaS-Organisation ist zunächst mit Aufwänden verbunden. Es müssen die IT-Infrastruktur geschaffen sowie Prozesse (zum Beispiel der Kontinuierliche Verbesserungsprozess (KVP)) und Strukturen (beispielsweise Kommunikationsstrukturen) entwickelt und implementiert werden. Zudem müssen ein Leistungskatalog sowie ein Vergütungsmodell erarbeitet werden. Dies ist zu Beginn aufwendig und zeitintensiv. Auch die Transitionsphase, in der das POaaS in die Organisation eingeführt wird, ist in der Praxis mit erhöhtem Aufwand verbunden, da Prozesse und Strukturen nachgeschärft werden müssen.



Mitarbeiterinnen und Mitarbeiter

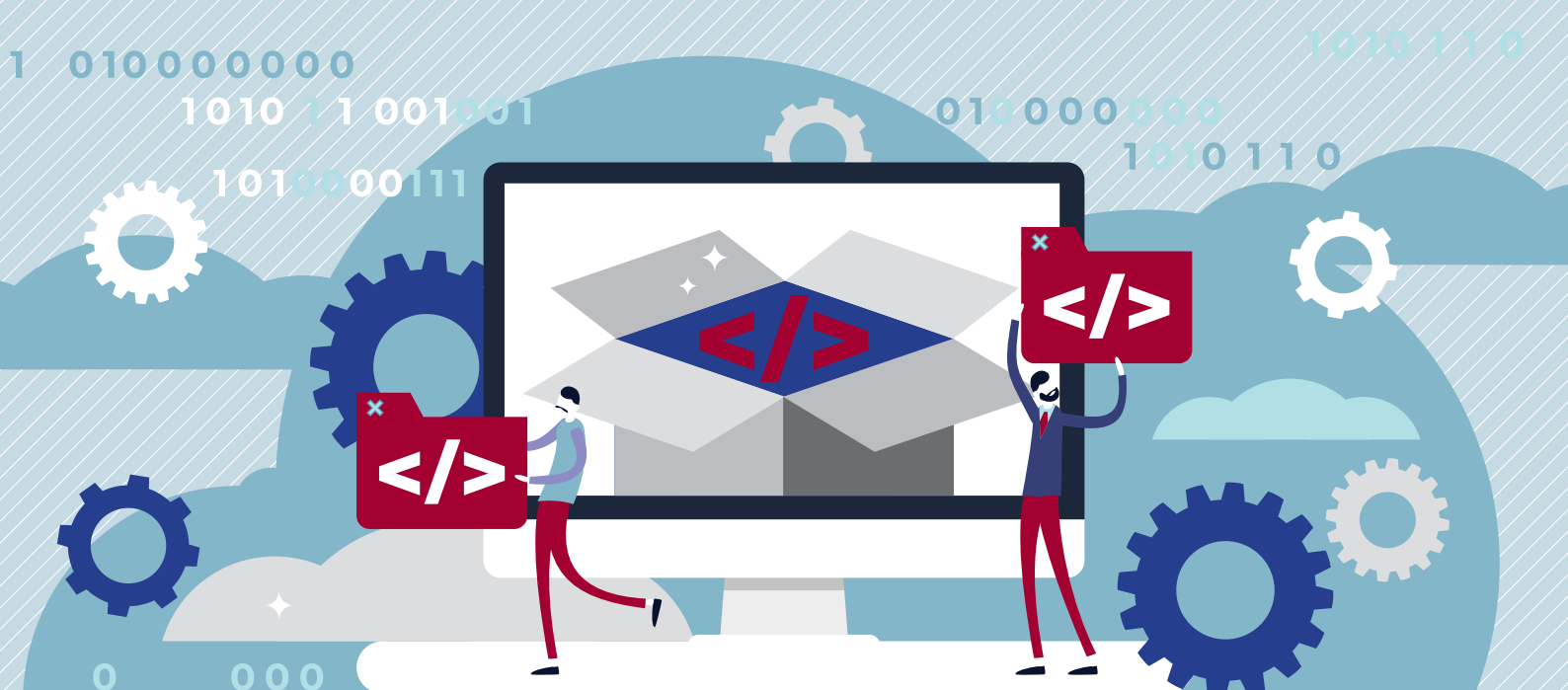
Wie bei jeder großen, strukturellen Veränderung braucht auch der Wechsel weg von dedizierten Project Offices, hin zu einer POaaS-Organisation die Unterstützung der Beteiligten, der Betroffenen und vor allem des Managements. Das POaaS kann nur erfolgreich sein, wenn sowohl entscheidende Instanzen als auch Projektmanager, Projektmitarbeiterinnen und -mitarbeiter diese Organisationsform unterstützen.

Es sind auch diese vier Herausforderungen, die den Einsatz externer POaaS-Services – zumindest in der Anfangsphase – nahelegen. POaaS-Anbieter haben diese vier Herausforderungen bei der Bereitstellung von PO-Leistungen in der Regel schon gemeistert. Sie können daher auch aufgrund ihrer Erfahrung bei der Einführung einer unternehmens- oder behördeninternen POaaS-Organisation unterstützen und beraten.

FAZIT

Ein POaaS ist eine zukunftsorientierte und innovative Organisationsform, die Unternehmen oder Behörden mit einer Vielzahl von Projekten und fortgeschrittenem Reifegrad große Vorteile bietet. Sie ermöglicht die bedarfsgerechte, flexible und individuelle Versorgung der Projekte mit PO-Services. Ist die Organisation einmal erfolgreich etabliert, ermöglicht das POaaS Unternehmen, Behörden und vor allem den Projekten, Kosten maßgeblich zu senken sowie hohe Qualitätsstandards zu etablieren und weiterzuentwickeln. ●

1 Henry Gantt (1861 bis 1919), US-amerikanischer Maschinenbauingenieur und Unternehmensberater.
2 https://de.wikipedia.org/wiki/Servant_Leadership (abgerufen am 16.03.2021).
3 GPM: Kompetenzbasiertes Projektmanagement (ICB 4).



HOCHWERTIGE ENTWICKLUNGSDOKUMENTATION AUTOMATISCH MIT ASCIIDOC

| von MARTIN SERNOW

Je größer ein Softwareprojekt, desto höher die Wahrscheinlichkeit, dass das Anpassen der Entwickler-, System- und Releasedokumentation von den Bearbeitern vergessen wird. Sei es, weil zu viel Dokumentation existiert oder weil sich das Projekt in einer kritischen Phase

befindet und die „goldenen Henkel“ ganz sicher später nachgeholt werden – so jedenfalls der gute Vorsatz.

Die Lösung für dieses Problem wäre, den Medienwechsel zwischen IntelliJ+Git und Word+SVN möglichst gering zu halten

und die Dokumentation dort abzulegen, wo sich die dokumentierte Problematik befindet. Hier kann die Auszeichnungssprache AsciiDoc Abhilfe schaffen. Wir verwenden es, um den Releaseletter, die Installationsanleitung, das Systemhandbuch und den Systementwurf durch AsciiDoc automatisiert generieren zu können. Durch die Umstellung von einem Worddokument auf mehrere AsciiDoc-Dokumente ergeben sich verschiedene Vorteile, die nachfolgend erläutert werden.

WAS IST ASCIIDOC?

AsciiDoc ist eine Auszeichnungssprache, mit der man schnell und einfach formatierte Dokumente erstellen kann. AsciiDoc kann mithilfe des Tools AsciiDoctor¹ direkt in den Softwareerstellungs- und Build-Prozess integriert werden. Aus der Code-Dokumentation werden Lieferartefakte erzeugt, die mit spezifischen Formatvorlagen für unterschiedlichste Ausgabeformate hinterlegt werden können. Dadurch entfällt die neben oder nach der Entwicklungs- und Releasephase anfallende zusätzliche Dokumentationsphase. Die erzeugte Dokumentation kann durch viele Werkzeuge (zum Beispiel IntelliJ, Eclipse, Visual Studio) geöffnet, erweitert und angesehen werden, ohne dass weitere Anwendungen gestartet werden müssen.



WELCHE PROBLEME BEI DER SYSTEM- UND RELEASEDOKUMENTATION WERDEN DURCH ASCIIDOC GELÖST?

Softwareentwickler nutzen für ihre Arbeit eine Entwicklungsumgebung (zum Beispiel IntelliJ) sowie gegebenenfalls weitere Entwicklungswerkzeuge, wie



Eine Auszeichnungssprache ist eine maschinenlesbare Sprache für die Gliederung und Formatierung von Texten. So kann man mit einfachen Mitteln wie ***Fett*** schreiben, mit **TIP:** ein Icon einfügen und per `image:icons/avatar.png[title="Avatar"]` sogar ein Bild ins Dokument einfügen. Ziel ist es, mit wenig Aufwand ein einheitliches und gut formatiertes Dokument zu schreiben.²

zum Beispiel Konsole, Datenbankviewer und Browser und verwalten ihre Arbeit in einer modernen Versionsverwaltung (zum Beispiel Git). Die Dokumentation dagegen findet über MS Word und die Verwaltung über eine „klassische“ Versionsverwaltung statt (zum Beispiel svn). Der an dieser Stelle entstehende Medienbruch und zusätzliche Aufwand führen häufig dazu, dass die Anpassung oder Aktualisierung der Dokumentation verschoben oder sogar ganz vergessen wird. Auch die Suche nach Informationen in der Dokumentation erweist sich oft als schwierig, da viele Entwicklungsumgebungen nicht sehr leistungsfähig in der separat abgelegten Dokumentation suchen können.

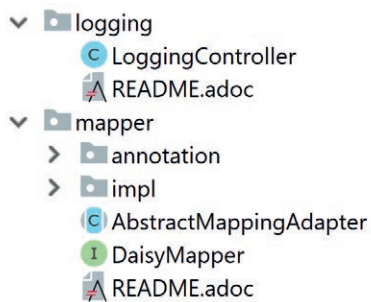


Abbildung 1: Projektstruktur mit integrierter ADOC-Datei

Dagegen liegen bei der Verwendung von AsciiDoc die Dateien für die Dokumentation direkt beim Code, im selben Verzeichnis wie die implementierte Klasse.

Andere Entwickler, die sich genau mit diesem Code befassen, können problemlos das entsprechende AsciiDoc-Dokument aufrufen, das entsprechende Informationen für die Arbeit an der Klasse bereitstellt. Zudem können Änderungen aufgrund von Change Requests oder Bugs sofort und an Ort und Stelle vorgenommen werden.

Code und Dokumentation können in einem gemeinsamen Review-Prozess der Entwicklung geprüft werden. Den Reviewern stehen beide Prüfobjekte gemeinsam zur Verfügung, und sie erkennen am Fehlen des AsciiDoc-Dokuments, dass eine Dokumentation noch aussteht.

Ein weiteres Problem ist das einheitliche Layout von MS-Word-Dokumenten. Da es sehr viele Möglichkeiten und Optionen gibt, ein Dokument „wunschgemäß“ zu formatieren, steigt die Vielfalt der „Layouts“ und Formatierungen, je länger ein Dokument bearbeitet wird und je mehr Personen daran arbeiten. Die Entwickler verlieren sich immer mehr darin, Design über Inhalte zu stellen.

AsciiDoc bietet zwar ebenfalls eine gute Auswahl an Dokumentdesign- und Formatierungsmöglichkeiten, allerdings wird das Design erst bei der Erzeugung des Artefakts (wie PDF, HTML, DOCX, PPT) auf Basis von Formatvorlagen (Templates) von AsciiDoctor über den Inhalt „gestülpt“. So müssen und können sich Entwickler vor allem den Inhalt konzent-

rieren. In vielen Fällen schreibt AsciiDoc sogar selbst vor, wie bestimmte Informationen formatiert sein müssen (zum Beispiel für Überschriften, Code Snippets, Bilder oder Tabellen).

Grafiken sind bei der System- und Entwicklerdokumentation oft ein wichtiges Mittel, um die Technik dahinter verständlich zu machen. Man möchte die Medien so ablegen, dass sie im generierten Dokument richtig angezeigt werden. Bei MS Word werden die Medien einfach per Drag & Drop in das Dokument „gezogen“. Allerdings ist danach das Aktualisieren der Bilder schwierig. Dafür bietet Word nur an, die Medien über eine Referenz zu ihrem Ablageort zu speichern. AsciiDoc ermöglicht das analog über die Referenzierung via URL-Notation des Ablagepfads. Dabei reicht auch eine relative Referenz mit Bezug zu einem zentralen Webserver aus, so dass man die Medien zentral bearbeiten, zum Download anbieten und Duplikate vermeiden kann.

Das Ausgabeformat von AsciiDoctor ist standardmäßig auf PDF und HTML5 beschränkt. Das Tool Pandoc³, das oft für Medienkonvertierungen über die Kommandozeile verwendet wird, bietet weitere Ausgabeformate an. Zusätzlich kann man Dokumente in mehreren Formaten exportieren: So kann man für Projektmitarbeiter die Dokumentation auf einem Siteserver hosten, den Beteiligten ein PDF schicken und gegebenenfalls für die nachträgliche Bearbeitung noch ein Docx bereitstellen. Um jedoch Kommentare von Nutzern zurück ins System zu spielen und dort weiterzuverarbeiten, wäre ein noch zu erstellendes Tool notwendig, das diese Kommentare aus dem dem Dokument extrahiert und zum Beispiel als To-dos an das entsprechende AsciiDoc-Dokument anhängt.

WIE INTEGRIERT MAN DEN BUILD-PROZESS IN SEIN PROJEKT?

Ein einfaches Beispiel zeigt, wie man eine Datei in einem Build Tool wie Maven integrieren kann, damit beim Build-Process das entsprechende Dokument mitgeneriert und an die Beteiligten ausgeliefert werden kann:

- Im ersten Schritt bindet man das asciidoctor-maven-Plugin ein, am besten im Build Process mit dem Profil asciidoc, damit man über dieses Profil steuern kann, wann man die Dokumente generiert.
- Als Nächstes legt man die verschiedenen Konvertierungen (executions) fest – in diesem Fall für den Systementwurf

```
<plugin>
  <groupId>org.asciidoctor</groupId>
  <artifactId>asciidoctor-maven-plugin</artifactId>
  <version>1.6.0</version>
  <executions>
    <execution>
      <id>generate-systementwurf-as-docbook</id>
      <phase>generate-resources</phase>
      <goals>
        <goal>process-asciidoc</goal>
      </goals>
      <configuration>
        <backend>docbook</backend>
        <doctype>book</doctype>
      </configuration>
    </execution>
  </executions>
</plugin>
<configuration>
  <sourceDirectory>.</sourceDirectory>
  <sourceDocumentName>SYSTEMENTWURF.adoc</sourceDocumentName>
  <outputDirectory>target/generated-docs</outputDirectory>
  <headerFooter>true</headerFooter>
  <logHandler>
    <outputToConsole>>false</outputToConsole>
    <failIf>
      <severity>DEBUG</severity>
    </failIf>
  </logHandler>
</configuration>
</plugin>
```

Abbildung 2: Projekt Object Model um ADOC Plugin erweitert

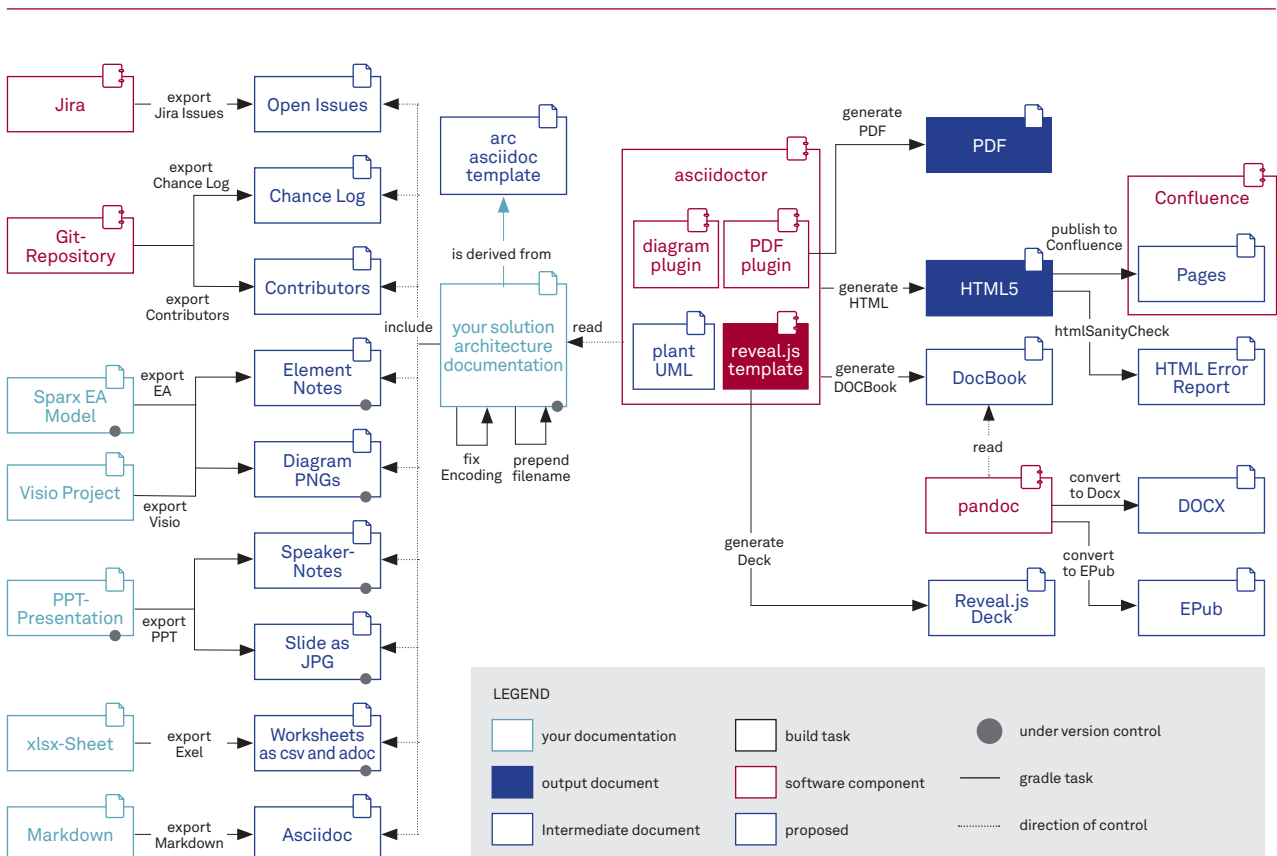


Abbildung 3: Übersicht der Möglichkeiten von DocToolchain®

DocBook (dies kann später zu dem Word-Dokumententyp docx konvertiert werden). In der Konfiguration sagt man mit dem Tag „backend“, dass man ein docbook wünscht. Für das DocBook-Dokument kann man html oder auch pdf verwenden, natürlich immer mit dem entsprechenden Doctype.

- Beliebig weitere executions kann man erzeugen, um entweder den Systementwurf in weitere Formate zu konvertieren oder auch andere Dokumente zu erzeugen.
- In der Konfiguration werden allgemeine, für alle executions gültige Werte übergeben, wie beispielsweise das Ausgabeverzeichnis (hier outputDirectory), welche Logeinträge geschrieben werden und woher die Quelldateien kommen.

DOCTOOLCHAIN

Doch was, wenn man mit mehreren Tools wie Sparx EA, Git, Jira oder Visio arbeitet und daraus die Inhalte in seinem AsciiDoc -Dokument anzeigen und aktuell halten möchte?

Das Werkzeug DocToolchain⁴ ermöglicht es, dynamisch noch mehr entwicklungsrelevante Informationen automatisiert in die AsciiDoc-Dokumente einzufügen. Ursprung der mittlerweile zahlreichen, automatisch Dokumentationsfunktionen unterstützenden DocToolchain war der Wunsch, Diagramme aus dem Enterprise Architekt⁵ in AsciiDoc-Dokumente einzubinden (siehe Abbildung 3).

Beispielszenario

Ein gutes Beispiel für „Documentation am Code“ aus der Projektpraxis ist der Einsatz von AsciiDoc im Ticketprozess. Änderungen im Programmcode aufgrund eines Tickets werden über ein spezielles Commit in Git abgebildet. Abschließende Commits wurden um das Kennzeichen „#releaseletter“ erweitert. DocToolchain sammelt automatisch für den Releaseletter alle Commits aus der Git-Historie und schreibt sie in den Releaseletter. Somit werden wichtige-Feature-Entwicklungen oder Fehlerkorrekturen im Releaseletter verzeichnet, ohne dass ein Release-Verantwortlicher die entsprechenden Tickets suchen und manuell hinzufügen muss.

ZUSAMMENFASSUNG

Mit dem Wechsel von MS Word auf eine Dokumentation direkt am Code durch AsciiDoc werden den Entwicklern manuelle Arbeiten abgenommen. Dies mindert das Fehlerpotenzial und versetzt Entwickler in die Lage, Code schneller zu verstehen und, neben JavaDoc⁷, besser zu dokumentieren. Der Release-Prozess wird durch die automatische Zusammenstellung von Release-Dokumentationen schlanker, die manuelle Arbeit entfällt zum großen Teil. Alles in allem „lebt“ die Entwicklungsdokumentation und wird stetig aktualisiert und publiziert. ●

1 <https://asciidoc.org/> (abgerufen am 26.10.2020).

2 Eine genaue Auflistung der AsciiDoc Formatierungen siehe hier: <https://asciidoc.org/docs/asciidoc-syntax-quick-reference/> (abgerufen am 30.03.2021).

3 <https://pandoc.org/> (abgerufen am 30.03.2021).

4 <https://doctoolchain.github.io/docToolchain/> (abgerufen am 26.10.2020).

5 <https://www.sparxsystems.de/> (abgerufen am 26.10.2020).

6 https://doctoolchain.github.io/docToolchain/#_overview_of_available_tasks (abgerufen am 26.10.2020).

7 <https://www.oracle.com/technical-resources/articles/java/javadoc-tool.html> (abgerufen am 26.10.2020).

AUTORENVERZEICHNIS



Werner Achtert ist Mitglied der Geschäftsleitung Public Sector der msg systems ag. Seine Schwerpunkte sind unter anderem die Analyse von Organisationsstrukturen und die Bewertung der Prozessreife mit Assessmentmethoden wie CMMI.



Ulrich Berghaus ist als Principal Project Manager für die msg systems ag tätig. Unter anderem als PRINCE2 Practitioner zertifiziert, verfügt er über langjährige Erfahrungen im agilen, klassischen und hybriden Projektumfeld in der öffentlichen Verwaltung.



Bernd Hahn ist als Lead Projektmanager bei der msg systems ag tätig und leitet das CoC Projektmanagement im Public Sector. Als IPMA- und PMI-zertifizierter Projektmanager verfügt er über eine langjährige Erfahrung im Managen von Projekten in der öffentlichen Verwaltung.



Dr. Christian Kiehle ist Abteilungsleiter im Public Sector der msg systems ag und verantwortet zahlreiche Projekte in der öffentlichen Verwaltung, in denen Architektur eine zentrale Rolle zur nachhaltigen Gestaltung der Behörden-IT spielt. Seine Interessenschwerpunkte liegen auf Enterprise Architecture Management in der öffentlichen Verwaltung sowie der IT-Standardisierung.



John Loutzenhiser ist Sprachwissenschaftler und Software-Ingenieur. Der geborene US-Amerikaner lebt seit 1997 in Deutschland und ist Experte für NLP (Natural Language Processing). Bei der msg systems ag setzt er als Principal IT Consultant und Leiter des Center of Competence „Data Science“ KI- und NLP-Projekte für die öffentliche Verwaltung um.



Christian Meyer ist seit mehreren Jahren als Principal Business Consultant und KI-Experte bei der msg systems ag im Public Sector tätig.



Dominik Müller ist bei der msg systems ag als IT Consultant im Public Sector tätig. Durch seine Arbeit im Projekt DAISY (Deutsche Auslands-Schularbeit-Informationen-System) verfügt er über umfangreiche Erfahrungen im Bereich Softwareentwicklung für die öffentliche Verwaltung. Im Center of Competence „Data Science“ beschäftigt er sich intensiv mit dem Thema Machine Learning.



Nedislav Nedyalkov ist bei der msg systems ag als Lead IT Consultant im Public Sector tätig. Er verfügt über langjährige Erfahrung in der Softwareentwicklung und der Architektur für Verfahren der öffentlichen Verwaltung. Aktuell liegen seine Schwerpunkte bei der Softwarearchitektur und der technischen Leitung des Projekts DAISY (Deutsche Auslands-Schularbeit-Informationen-System) für das Auswärtige Amt. Des Weiteren ist er Mitglied im Center of Competence „Data Science“ im Geschäftsbereich Public Sector.



Alexander Salvanos verfügt über langjährige Praxiserfahrung als IT-Experte für Java-/Jakarta-EE-Technologien und ist Mitglied der Java EE 8 Expert Group (JSON-B). Als IT-Senior-Consultant bei der msg systems ag berät er Behörden und Unternehmen beim professionellen Einsatz moderner Java-Technologien. Darüber hinaus ist er Autor von Bestsellern, wie „Professionell entwickeln mit Java EE 8“ und „UML 2.5“. Außerdem ist er Referent in E-Learning-Kursen auf LinkedIn.



Martin Sernow ist Wirtschaftsinformatiker und bei der msg systems ag als IT-Consultant in der öffentlichen Verwaltung tätig. Seine Expertise liegt in der Entwicklung von Software mit innovativen Komponenten wie Kotlin, MongoDB, AsciiDoc sowie in der Entwicklung von Fullstack Java-Systemen mit Webtechnologien, Spring, REST und Oracle-Datenbanken.



Hans-Ulrich Striffler ist Lead Business Consultant und bei der msg systems ag als Softwarearchitekt, Prozess- und Anforderungsmanager tätig. Der Dipl.-Informatiker mit Schwerpunkt E-Health und künstliche Intelligenz verfügt über viel Erfahrung in der Modellierung von Geschäfts- und Entscheidungsprozess-Systemen in BPMN und in DMN. Er referiert auf internationalen Konferenzen über „Entscheidungsmanagement in Prozessen“.



Dr. Andreas Zamperoni ist promovierter Diplom-Informatiker (Nebenfach kognitive Psychologie) mit langjähriger Erfahrung als Projekt- und Programm-Manager für unternehmenskritische Softwaresysteme. Er berät und coacht zu agilen und klassischen Projektmanagement- und Softwareengineering-Themen, unterstützt Transitionsprozesse von Software- und IT-Organisationen sowie die Entwicklung von IT- und Fachstrategien. Zudem ist er PMI-PMP®, Certified Scrum Master, Chefredakteur der .public.



msg Security Advisors

FÜR EINE RESILIENTE DIGITALE VERWALTUNG

msg hat ihr IT-Sicherheits-Know-how gebündelt und setzt mit mehr als hundert Expertinnen und Experten Maßstäbe. Die msg Security Advisors unterstützen Ihre Organisation bei Cybersicherheit und Datenschutz: mit Strategie, Konzeption oder in der methodischen und operativen Umsetzung sowie im IT-Risiko- und Compliance-Management.

Wir beraten Sie gerne!

<https://www.msg.group/services/security-advisors>



IHNEN GEFÄLLT
DIE AUSGABE?
DANN ABONNIEREN
SIE .public UND
EMPFEHLEN SIE UNS
WEITER.

