



**Mittelstand 4.0**  
Kompetenzzentrum  
Handel



LEITFADEN

# Fakt oder Fiktion?

Einsatzmöglichkeiten Digitaler Sprachassistenten  
im Einzelhandel

Mittelstand-  
Digital 

Gefördert durch:



Bundesministerium  
für Wirtschaft  
und Energie

aufgrund eines Beschlusses  
des Deutschen Bundestages

# Inhalt

<b>1</b>	<b>Einleitung</b>	S. 4
<b>2</b>	<b>Digitale Sprachassistenten</b>	S. 5
<b>3</b>	<b>Aktuelle Sprachassistentensysteme</b>	S. 6
	3.1 Apple: Siri	S. 7
	3.2 Samsung: S Voice, jetzt Bixby	S. 7
	3.3 Microsoft: Cortana	S. 7
	3.4 Google: Assistant	S. 8
	3.5 Amazon: Alexa	S. 8
<b>4</b>	<b>Sprachassistenten in der Praxis</b>	S. 9
<b>5</b>	<b>Einsatzmöglichkeiten im Handel</b>	S. 10
	5.1 Front-Store: Kundenkontakt und Beratung	S. 10
	5.2 Back-Store: Interne Abläufe am Beispiel des Modehandels	S. 11

<b>6</b>	<b>Projekt „Rita“</b>	S. 14
	6.1 Zielsetzung	S. 14
	6.2 Auswahl der verwendeten Plattform	S. 15
	6.3 Entwicklungsumgebung und technische Voraussetzungen	S. 15
<b>7</b>	<b>Umsetzung und Aufbau von „Rita“</b>	S. 16
<b>8</b>	<b>Learnings: Technische und alltägliche Herausforderungen</b>	S. 18
<b>9</b>	<b>Fazit und Ausblick</b>	S. 20
	Impressum	S. 21
	Literaturverzeichnis	S. 22

# 1. Einleitung

**M**orgens ins vorgeheizte Auto steigen, sich vom Bordcomputer über die verkehrsgünstigste Strecke leiten lassen, während der Fahrt den Digitalen Sprachassistenten beauftragen, im Geschäft die Heizung hochzufahren und Kaffee aufzubrühen, dort angekommen gehorchen die Rollgitter aufs Wort und die Eingangstür lässt sich per Stimmerkennung entriegeln – was nach Zukunftsmusik klingt, könnte bald Arbeitsalltag sein.

Sprachsteuerung ist längst in aller Munde. Schon heute erleichtern viele kleine Helfer unser Leben und Verleihen der „Macht des Wortes“ eine ganz neue

Bedeutung. Siri, Cortana, Alexa, der Google Assistent und Bixby, um die bekanntesten Assistenten zu nennen, haben sich im Privat- und Geschäftsleben fest etabliert. Sie verwalten Termine, lesen E-Mails vor, lassen sich Textnachrichten diktieren, spielen Musik ab und versorgen uns mit nützlichen Informationen wie Nachrichten, Fahrplanauskünften, den aktuellen Wetteraussichten oder der Bestandsliste unseres Kühlschranks.

Wo sind Digitale Sprachassistenten heute schon im Einsatz?

Und welches Unterstützungspotenzial können sie dem Handel jetzt und in Zukunft bieten?

Die Antworten werden nachfolgend näher beleuchtet.



## 2. Digitale Sprachassistenten

**E**in Digitaler Sprachassistent ist eine spezielle Software, die es ermöglicht, ein elektronisches Gerät, wie ein Smartphone, Tablet-/Computer oder einen intelligenten Lautsprecher (sog. Smart Speaker), allein durch die Stimme zu bedienen. Die Sprachsteuerung ersetzt dabei die klassische Eingabe über die Tastatur oder per „Mausklick“.

Um mit der gewünschten Aktion reagieren zu können, muss das angesprochene Gerät die Worte „hören“, also über ein Mikrofon verfügen, die Informationen einordnen bzw. interpretieren und in einen für den Computer verständlichen Befehl umwandeln können.

Während das Gespräch mit mobilen Geräten in Science-Fiction-Filmen zum Alltag gehört, die Sprachassistenten dort sogar teilweise über eigene Persönlichkeiten verfügen, stellt die Sprachsteuerung für die derzeitigen Systeme eine Herausforderung dar. So müssen beispielsweise mögliche dop-

pelte Bedeutungen von Wörtern („Teekesselchen“) herausgefiltert und in den richtigen Zusammenhang gebracht werden und auch in regionale Dialekte „übersetzt“ werden. Gleichzeitig erwartet der Fragestellende (Nutzer:in) eine Antwort in angemessener Zeit, somit erfordert die Umwandlung in computergerechte Kommandos eine hohe Rechenleistung innerhalb eines möglichst kurzen Zeitraums.

Bereits vor den Smartphones<sup>1</sup> gab es Mobiltelefone (z. B. das Siemens i35S oder das Nokia 8210, die beide ab dem Jahr 2000 verkauft wurden), die über eine Diktierfunktion verfügten und in der Lage waren, Kontakte per Sprachbefehl anzurufen. Doch erst technische Fortschritte, wie die Entwicklung von Mikrochips mit hoher Speicherkapazität und die flächendeckende Verbreitung des (mobilen) Internets, ermöglichen die heutige Kommunikation mit den Digitalen Sprachassistenten.

### **„Wenn ich die See seh', brauche ich kein Meer mehr.“**

Ein einfacher Satz für einen Menschen und ein Quantensprung für die Maschine. Gesprochene Sprache zu verarbeiten und in für den Computer verständliche „Kommandos“ zu übersetzen, sodass (zeitnah) eine geeignete Reaktion bzw. sinnvolle Antwort erfolgen kann, erfordert enorme Rechenleistungen.

<sup>1</sup> Als erstes „Smartphone“, also Mobiltelefon mit Touchscreen-Display etc., wird das iPhone von Apple betrachtet, das 2007 auf den Markt kam.



### 3. Aktuelle Sprachassistenzsysteme

**S**eit Apple im Jahr 2011 der Weltöffentlichkeit „Siri“ als ersten funktionstüchtigen mobilen Sprachassistenten vorstellte, hat sich in diesem Bereich viel getan. Neben großen internationalen Konzernen wie Amazon (Alexa), Google (Assistant), Microsoft (Cortana) oder Samsung (S Voice/Bixby) bietet mittlerweile z. B. auch die Deutsche Telekom ihren Kunden ein System (Magenta) zur Sprachsteuerung an.

#### **Zauberhafte Assistentin?**

Siri und der Google Assistant bieten neben der weiblichen auch eine männliche Stimmauswahl an. Bei Siri ist die männliche Stimme je nach Land und Sprachversion, z.B. im Französischen, im Niederländischen und im britischen Englisch, bereits voreingestellt.

Die derzeit bekanntesten Sprachassistenten werden hier in der Reihenfolge ihres Markteintritts in Deutschland kurz vorgestellt:

### 3.1 Apple: Siri (2011)

Das Spracherkennungssystem „Siri“ ist seit dem „iPhone 4s“ auf allen Apple-Produkten vertreten. Konnte man den Sprachassistenten zunächst nur mit dem Home-Knopf aktivieren, reagiert er nun auf den Weckruf „Hey Siri“.

Neben der Verwaltung von Kalendereinträgen, Erinnerungen und dem Anrufen von Kontakten kann Siri Suchanfragen beantworten oder navigieren. Zudem wird Siri auch zur Bedienung der von Apple entwickelten HomeKit-Geräte eingesetzt. So kann beispielsweise auf Zuruf die Haustür geöffnet und das Licht angeschaltet werden oder per Sprachbefehl die Überwachungskamera eingeschaltet und die Bilder auf dem Smartphone angezeigt werden. Zusätzlich können andere Hersteller ihre „Smart Home“-<sup>2</sup>Produkte von Apple zertifizieren lassen und sind damit auch über Siri nutzbar, sodass die Palette der ansteuerbaren Geräte stetig wächst.

### 3.2 Samsung: S Voice (2012), jetzt Bixby (2019)

Nur ein Jahr nach Apple brachte Samsung den Sprachassistenten „S Voice“ mit einem ähnlichen Funktionsumfang auf den Markt. Aktivieren ließ sich das S Voice durch zweimaliges kurzes Drücken der Home-Taste, danach reagierte das Programm auf die Ansprache „Hallo Galaxy“.

Im Jahr 2019 wurde S Voice auch auf dem deutschen Markt durch seinen Nachfolger „Bixby“ ersetzt und wird nun mit „Hi Bixby“ geweckt. Der Sprachassistent kann an die Smart-Home-Geräte von Samsung gekoppelt werden und so per Sprachsteuerung beispielsweise mit dem Fernseher oder dem Kühlschrank kommunizieren. Eine Besonderheit von Bixby ist die Smartphone-Linse, die Produkte erkennt und diese oder ähnliche Produkte im Internet sucht.

### 3.3 Microsoft: Cortana (2014)

„Cortana“ ist das Sprachsteuerungsmodul aus dem Hause Microsoft. Der Digitale Sprachassistent wurde erstmalig mit dem „Windows Phone 8.1“ eingeführt und ist seit „Windows 10“ fester Bestandteil der Computer-Betriebssysteme.

Über Cortana können Fragen an die Suchmaschine „Bing“ gestellt werden. Darüber hinaus lassen sich E-Mails versenden, Spiele spielen, Apps öffnen und es können Listen, Termine und Erinnerungen

<sup>2</sup>Haustechnik, elektronische Haushaltsgeräte und Unterhaltungselektronik, die untereinander vernetzt sind und z. B. durch Apps ansteuerbar sind, fasst man unter den Begriffen „Smart Home“ oder „Connected House“ zusammen.

erstellt und verwaltet werden. Aktiviert wird der Sprachassistent durch den Befehl „Hey Cortana“ oder durch einen Klick auf das Mikrofon-Symbol. Eine Kommunikation und Steuerung weiterer Geräte ist mit Cortana bisher noch nicht möglich.

### 3.4 Google: Assistant (2016)

Der „Google Assistant“ ist seit 2016 im Einsatz und neben dem Android-Betriebssystem und Google Home auch für iOS-Nutzende verfügbar. Das System kann Termine vereinbaren, Mietwagen buchen, Fragen beantworten und wird mittlerweile auch in den Bordcomputern mehrerer Autohersteller verbaut, wo es sich um die Navigation sowie die Steuerung der Audiosysteme kümmert.

Ansonsten unterscheidet sich dieser Sprachassistent in den wesentlichen Funktionen kaum von den vorgenannten Geräten.

### 3.5 Amazon: Alexa (2016)

Der Sprachassistent von Amazon heißt „Alexa“ und ist seit 2016 in Deutschland mit den Echo-Lautsprechern erhältlich. Im ursprünglichen Funktionsumfang war Alexa vor allem auf das Abspielen von Musik oder Hörbüchern, die Terminverwaltung sowie die Erstellung von Einkaufslisten spezialisiert. Heutzutage umfasst die Produktpalette neben den „Smart Speakern“ auch Geräte mit Touchscreen-Display und/oder mit eingebauter Kamera.

Voraussetzung für die Nutzung ist ein Amazon-Kundenkonto, registrierte Prime-Mitglieder können dann auch direkt per Sprachbefehl über Alexa die gewünschten Produkte bestellen.<sup>3</sup> Damit bedient Amazon als einziges der hier vorgestellten Unternehmen den sog. Voice-Commerce.

Eine weitere Besonderheit ist, dass Alexa nicht nur auf Amazon-Geräte begrenzt ist, sondern durch sog. „Skills“ auch anderen Anbietern zur Verfügung steht. Die Spanne reicht dabei von der Mikrowelle über die Beleuchtung, Rolläden, smarte Rasenmäher bis hin zum Fahrzeugcomputer (Connected Car). Auch Telefonanrufe können mittlerweile über Alexa getätigt werden.

Aufgrund von stetig neu hinzukommenden Kooperationen wird die Anzahl der mit Alexa kompatiblen Produkte weiter steigen.

Alle Geräte haben gemeinsam, dass sie eine bestehende Internetverbindung benötigen, um zu funktionieren. Zudem setzen alle Sprachassistenten zur Nutzung die Anmeldung bei einem bestehenden Nutzerkonto voraus. Lediglich Cortana bietet eine abgespeckte Version ohne vorherige Registrierung an, der volle Funktionsumfang wird jedoch nur mit Benutzerkonto angeboten.

#### So aktivieren Sie Ihren Sprachassistenten

Während die meisten Sprachassistenten nur auf einen „Weckruf“ hören, reagiert Alexa derzeit auf vier verschiedene Stichwörter.

- Alexa: „Alexa“, „Amazon“, „Echo“ oder „Computer“
- Bixby: „Hi Bixby“
- Cortana: „Hey Cortana“
- Google Assistant: „Ok Google“
- Siri: „Hey Siri“

<sup>3</sup> Alexas Werkseinstellungen erlauben zunächst einmal jedem, der in ihrer Hörweite ist, über das mit ihr verbundene Amazon-Kundenkonto Bestellungen auszulösen. Um Missbräuche zu vermeiden, lassen sich Einkäufe in den Einstellungen komplett unterbinden oder durch die zusätzliche Abfrage einer PIN einschränken.

# 4. Sprachassistenten in der Praxis

**E**xperten zufolge könnten 2023 bereits bis zu 1,6 Milliarden Smart-Home-Geräte<sup>4</sup> weltweit im Einsatz sein. Bisher werden die Digitalen Sprachassistenten jedoch vor allem für den Privatgebrauch genutzt.

Doch die zunehmende Vernetzung der Systeme und die wachsende Anzahl an Geräten, die mit einem Sprachassistenten gekoppelt werden können, bergen auch für das Geschäftsleben große Unterstützungspotenziale. So sind erste Anwendungen bereits in verschiedenen Branchen im Einsatz.

Im **Dienstleistungssektor** wie der telefonischen **Kundenbetreuung** wird die Sprachwahl schon länger genutzt. Beispielsweise um das Anliegen eines Anrufers abzufragen und ihn so schneller an einen geeigneten Ansprechpartner weiterverbinden zu können. Auch Social-Media-Anfragen können heute schon durch intelligente **Chatbots** automatisch beantwortet werden.

Für den **Bankensektor** sind erste Anwendungen in Bereichen mit sensiblen Daten und hohen Sicherheitsanforderungen auf dem Markt. So können Kunden verschiedener Banken bereits kleinere Geldbeträge mittels Sprachassistenten **überweisen**.<sup>5</sup>

Immer mehr **Hotels** stellen Digitale Sprachassistenten in ihren Zimmern bereit. Sie übernehmen die Steuerung von Fernsehern, Musikanlagen, Beleuchtung und Temperatur per Sprachbefehl und bieten den Gästen nützliche Informationen wie Wettervorhersagen, Veranstaltungstipps, digitale Gästemappen oder ermöglichen die Bestellung beim Zimmerservice.

Doch schon den Weg in den Urlaub könnten die Sprachassistenten zukünftig erleichtern, so erlauben erste **Fluglinien** ihren Passagieren den **Check-in** per Sprachbefehl.<sup>6</sup>

Mit **Pick-by-Voice**<sup>7</sup> werden in der **Logistikbranche** Systeme bezeichnet, die die Lagermitarbeiter bei der **Kommissionierung von Waren** unterstützen. Informationen wie bestellte Artikel, gewünschte Anzahl und Lagerort werden auf das Headset des Kommissionierers übermittelt, der mit dem Sprachassistenten im Austausch steht. Je nach Ausstattung und Anbindung des Warenlagers können diese Systeme durch eine effiziente Wegführung erweitert werden.

<sup>4</sup> <https://www.idc.com/getdoc.jsp?containerId=prUS44971219>

<sup>5</sup> Bundesverband Digitale Wirtschaft (BVDW) e.V., 2019, S. 12

<sup>6</sup> Zum Beispiel United Airlines via Google Assistant.

<sup>7</sup> <https://www.prologistik.com/logistik-lexikon/sprachgestuetztes-kommissionieren/>; abgerufen am 21.01.2021

# 5. Einsatzmöglichkeiten im Handel

**A**uch im Einzelhandel wird derzeit in vereinzelt Pilotprojekten das mögliche Potenzial Digitaler Sprachassistenten ausgelotet.

## 5.1 Front-Store: Kundenkontakt und Beratung

Auf der Verkaufsfläche und im direkten Kundenkontakt dienen die Digitalen Sprachassistenten bisher als Marketing-Instrument und Image-Faktor.

### **Alexa-Skill von HIT Sütterlin<sup>8</sup>**

Nach eigenen Angaben als erster Händler in Deutschland setzt HIT Sütterlin Amazons Alexa in der Kundenkommunikation ein. Bereits seit 2017 stellt der Aachener Lebensmittelhändler Alexa-Nutzenden per Sprachbefehl Informationen zu aktuellen Angeboten, Wochenhighlights und den Öffnungszeiten zur Verfügung.

### **REWE Assistant<sup>9</sup>**

Ebenfalls im Jahr 2017 gestartet, bietet REWE über den Google Assistant auf Wunsch Kochrezepte und Zutatenlisten an. Zusätzlich ermöglicht es das Kölner Unternehmen, fehlende Zutaten auf Zuruf auf der virtuellen Einkaufsliste zu speichern.

### **H&M „Voice Interactive Mirror“<sup>10</sup>**

Während HIT Sütterlin und REWE bestehende Sprachassistenten um eigene Anwendungen erweitern, wurde das „Voice Interactive Mirror“-System eigens für H&M entwickelt. 2018 stellte H&M den nach eigenen Angaben weltweit ersten „sprachgesteuerten interaktiven Spiegel“ vor. Er befindet sich im Geschäft am New Yorker Times Square und verfügt über eine Stimm- und Gesichtserkennungs-Software.

Die Kundschaft kann dort per Sprachwahl zwei Möglichkeiten nutzen:

- die Selfie-Funktion, die ein Foto des Kunden aufnimmt und damit ein personalisiertes Titelbild des hauseigenen Magazins erstellt, welches durch das Scannen eines QR-Codes abgerufen werden kann, und
- die „Fashion-Inspirationen“, bei der vorgestellte Kleidungsstücke über das Abscannen eines QR-Codes direkt gekauft werden können. Um die Kunden zum sog. „Mirror Shopping“ zu motivieren, erhalten sie hierbei einen Extra-Rabatt.

Sprechende Obst- und Gemüsewaagen, die auf Anfrage weiterführende Informationen wie das Herkunftsland, Bio- oder Fair-Trade-Zertifikate oder passende Rezepte ausgeben können, zählen nicht zu den Digitalen Sprachassistenten, können jedoch ggf. von ihnen angesteuert werden.

<sup>8</sup> <https://www.youtube.com/watch?v=gXsMAkpLnBA>

<sup>9</sup> <https://www.youtube.com/watch?v=iUjKNkiTuyE>

<sup>10</sup> <https://www.youtube.com/watch?v=A1Pif06jdVA>

## 5.2 Back-Store: Interne Abläufe am Beispiel des Modehandels

Während die Sprachassistenten bisher branchenübergreifend überwiegend in der Kundenkommunikation verwendet werden, ist vergleichsweise wenig über das Unterstützungspotenzial interner Abläufe bekannt. Nachfolgend werden die Verwen-

dungsmöglichkeiten eines Digitalen Sprachassistenten anhand eines beispielhaften Tagesablaufs eines inhabergeführten Modegeschäfts mit einer weiteren festen Arbeitskraft und zwei Aushilfskräften näher aufgezeigt:

### Vor Mitarbeiter-Phase

- Ankommen
- Alarmanlage ausschalten  
(Die Bedienung ist hier zwar per Sprachbefehl möglich, wird aus Sicherheits- und Versicherungsgründen jedoch nicht empfohlen)
- Beleuchtung im Büro und den Betriebsräumen einschalten
- Zeiterfassung: Einstampeln
- Heizung/Klimaanlage im Büro, den Betriebsräumen, im Lager und auf der Verkaufsfläche einschalten

■ = Unterstützungspotenzial durch Sprachassistenten

### Vor Kunden-Phase

- Kaffeemaschine anschalten
- Spülmaschine ausräumen
- Beleuchtung im Lager einschalten
- Steuerung der Überwachungskamera an Lieferanten-/Mitarbeiter- und weiteren Eingängen
- Warenannahme
- Wareneingangskontrolle
- Warenverbuchung im Lager
- Bestandskontrolle auf der Fläche
- Auflisten der dort fehlenden Artikel
- Ware im Lager kommissionieren
- Ware vorbereiten, z. B. Aufbügeln etc.
- Reduzierte Beleuchtung auf der Verkaufsfläche einschalten
- Regale auffüllen
- Schaufenster-/Auslagengestaltung
- Volle Beleuchtung auf der Verkaufsfläche anschalten
- Store-Radio/-TV/-Entertainment-System einschalten
- Rollgitter herauffahren
- Verkaufsräume aufschließen

## Während der Öffnungszeiten

- Die Kundschaft beraten
- Wege zum gewünschten Produkt weisen
- Ware verkaufen, Kassivorgang
- Rabatte oder Währungen errechnen
- Ware aus dem Lager nachfüllen
- Sonst Nachbestellung direkt aus dem System heraus bzw. nach Abgleich mit dem Warenbestand auf der Fläche
- Kabinenkontrolle
- Diebstahlprävention auf der Fläche
- Ordnung halten auf der Verkaufsfläche
- Umgestaltung der Auslagen/Warenpräsentation
- Geld abholen lassen/zur Bank bringen
- Lieferavise abrufen und ins System buchen
- Bei Pausenzeiten: Zeiterfassung: Ein- und Ausbuchung
- Bei Schichtwechsel: Zwischenabrechnung, Aktualisierung des Kassenbuchs
- Beleuchtung auf der Verkaufsfläche je nach Tageszeit und Witterungslage anpassen
- Heizung/Klimaanlage auf der Verkaufsfläche je nach Tageszeit und Witterungslage anpassen

## Bei Onlineshop zusätzlich

- Verfügbarkeit der Artikel aktualisieren
- Neue Artikel einstellen
- Online-Bestellungen kommissionieren
- Lieferscheine erstellen
- Pakete packen
- Pakete abholen lassen

## Nebenbei

- Personaleinsatzplanung
- Anrufe und E-Mails beantworten
- Stellenanzeigen aufgeben
- Termine/Bewerbungsgespräche mit Aushilfskräften, Lieferunternehmen etc. vereinbaren und wahrnehmen
- Wettervorhersage abfragen
- Neue Stücke für den Verkauf aussuchen z. B. ähnliche Stücke durch die Cortana-Bildsuche finden

## Vor Geschäftsschluss

- Vorziehen, Ordnung auf der Fläche
- Bestandsaufnahme
- Nachbestellungen bei den Lieferanten
- Heizung/Klimaanlage auf der Verkaufsfläche reduzieren bzw. ausschalten

■ = Unterstützungspotenzial durch Sprachassistenten

Es gibt Handelsunternehmen, die Roboter auf der Verkaufsfläche einsetzen, um ihrer Kundschaft zusätzliche Kontaktpunkte anzubieten. Obwohl die Roboter per Sprache mit den Kund:innen kommunizieren und ihnen mit Rat und Tat zur Seite stehen, zählen sie nicht zu den Digitalen Sprachassistenten, da das Sprachsteuerungsmodul hier nur einen Baustein des gesamten Systems ausmacht.

## Nach Geschäftsschluss

- Rollgitter herunterlassen
- Kundenzugang zu den Verkaufsräumen abschließen
- Store-Radio/-TV/-Entertainment-System ausschalten
- Beleuchtung in den Verkaufsräumen reduzieren
- Saug-/Wischroboter auf die Fläche schicken
- Kassenendabrechnung, Aktualisierung des Kassenbuchs
- Kassenlade im Tresor deponieren
- Kaffeemaschine reinigen bzw. Reinigungsprogramm anstoßen
- Spülmaschine einräumen und anstellen
- Kaffeemaschine ausschalten
- Heizung/Klimaanlage im Büro und in den Betriebsräumen ausschalten bzw. reduzieren
- Beleuchtung im Büro, den Betriebsräumen und im Lager ausschalten
- Zeiterfassung: Ausstempeln
- Abschließen



# 6. Projekt „Rita“

Das Mittelstand 4.0-Kompetenzzentrum Handel hat die Einsatzmöglichkeiten Digitaler Sprachassistenten im Handel genauer geprüft und für die **„Warenverbuchung im Lager“** und die **„Bestandskontrolle auf der Fläche“** einen sog. Demonstrator, den Prototypen **„Rita“** entwickelt.

## 6.1 Zielsetzung

Die oben beschriebene Boutique ist ein gut florierendes Geschäft mit Onlineshop, das stets am Puls der Zeit bleibt und die Kundenbedürfnisse im Blick hat. Das Lager wird täglich aufgefüllt, neue Kollektionen und saisonale Artikel werden bestellt und abverkaufte Größen nachgeordert, bei besonderen Wünschen der Kundschaft werden auch Einzelbestellungen bei anderen Handelsunternehmen ausgelöst, so dass mehrfach täglich neue Ware verschiedener Marken und Hersteller angeliefert wird. Dabei verfolgen die Lieferfirmen eigene Lieferkonzepte, vom Express- über den Speditionsversand bis hin zur Anlieferung per Lastenfahrrad. Die Inhaberin des Geschäfts prüft die eingegangene Ware sehr gewissenhaft und hat auch ihre Mitarbeiterin dazu angehalten, neben dem Abgleich mit der Bestellliste die Qualität der Ware im Blick zu halten.

Bisher werden die Lieferscheine mit den einzelnen Artikeln, Liefermengen und Preisen manuell ins Warenwirtschaftssystem gebucht, eigene Artikelnummern vergeben und aus dem System heraus Preisetiketten gedruckt. Der Onlineshop wird dabei wie eine zweite Filiale behandelt. Zukünftig soll die „Warenverbuchung im Lager“ per Spracheingabe erfolgen.

Die gute Beratung, die Qualität der Ware und die liebevolle Auswahl der Stücke hat sich nicht nur im eigenen Viertel herumgesprochen, auch der Onlineshop erfreut sich wachsender Beliebtheit und eines immer höheren Bestellaufkommens. Noch lohnt es sich nicht, eine neue Arbeitskraft dafür einzustellen, doch die vorhandene Personaldecke stößt langsam an ihre Grenzen, und auch wenn die beiden Aushilfskräfte hier viel abfedern können, wäre es eine Erleichterung, wenn man die „Bestandskontrolle auf der Fläche“ quasi nebenbei erledigen könnte. Bisher werden hier die einzelnen Artikelnummern mit Mengen, Größen und ggf. Farben noch handschriftlich erfasst und mit den im Warenwirtschaftssystem angegebenen Werten verglichen. Ein Vorgang, der zur Kundenberatung, zum Kassieren oder um Ordnung auf der Fläche zu schaffen, unterbrochen werden muss. Daher soll der Digitale Sprachassistent „Rita“ auch hier in Zukunft Unterstützung leisten, sodass der Gang zum Computer überflüssig wird.

## 6.2 Auswahl der verwendeten Plattform

Bei der Auswahl des zu verwendenden Sprachsteuerungsmoduls stehen folgende Kriterien im Vordergrund:

- das Sprachverständnis<sup>11</sup>
- der technische Support und die Aktualität der verwendeten Software
- die vorhandenen Schnittstellen des Systems und weitere Vernetzungsmöglichkeiten sowie
- die Anschaffungs- und die Betriebskosten

Aufgrund der bereits genutzten Computersysteme, der persönlichen Vorliebe der Inhaberin und der Erfahrungswerte der Programmiererin fällt die Entscheidung, dass „Rita“ auf dem Google Assistant aufbauen soll. Damit die Spracheingabe im Zweifelsfall direkt am Sprachassistenten visuell kontrolliert und wenn notwendig korrigiert werden kann, entscheidet man sich für ein Display mit Google-Assistent-Funktion.

## 6.3 Entwicklungsumgebung und technische Voraussetzungen

Dabei wird „Rita“ auf der cloudbasierten Plattform „Dialogflow“ entwickelt. Diese Plattform ist auf die Spracheingabe über dialogorientierte Kommunikation und das sog. Natural Language Processing spezialisiert.

Was bedeutet, dass der Computer das Gesagte analysiert und in für ihn verwertbare Befehle übersetzt, ohne dass die sprechende Person festgelegte „Kommandos“ nutzen muss, die in ihrer Wortwahl nicht abweichen dürfen. Anderenfalls würde beispielsweise ein Zeilenumbruch in einem Text nur erfolgen, wenn das Wort „Enter“ ausgesprochen wird. „Neue Zeile“ oder „Zeilenumbruch“ würde der Computer nicht erkennen und somit auch nicht umsetzen können.

Dialogflow ist hauptsächlich auf die Erstellung sog. „Sprachdienste“ für den Google Assistant spezialisiert, unterstützt gleichzeitig jedoch weitere Programmiersprachen für Plattformen wie Skype oder den Facebook Messenger. Weitere Vorteile von Dialogflow sind die dort integrierte Künstliche Intelligenz, das maschinelle Lernen und die Cloud-Dienste von Google. Der einfache Aufbau ermöglicht es auch Laien, schnell mit diesen Technologien arbeiten zu können. Neben der eigentlichen Implementierung mit Dialogflow wird eine Schnittstelle zum Warenwirtschaftssystem benötigt. Im Rahmen dieses Projekts wird zunächst eine Test-Datenbank zur Verwaltung der Bestände und Bestellungen in Google Firebase angelegt.

Für die Verwendung der Plattform wird ein Google-Konto vorausgesetzt.<sup>12</sup>

### So hört der Digitale Sprachassistent aufs Wort

Um Ihren Kommandos zu folgen, benötigt der Digitale Sprachassistent zunächst ein paar Stimmproben. Um sich Ihre Aussprache und die Klangfarbe Ihrer Stimme einzuprägen, werden vorgegebene Texte von Ihnen abgefragt.

<sup>11</sup> Siehe hierzu: Munster & Thompson: Annual Smart Speaker IQ Test, 2018.

<sup>12</sup> Die Anbindung von Sprachassistenten ist derzeit nur in der US-Version von „Dialogflow“ verfügbar.

# 7. Umsetzung und Aufbau von „Rita“

**D**iese Datenbank konzentriert sich im Wesentlichen auf zwei Menüpunkte: Den Lagerbestand und die offenen Bestellungen.

Im Lagerbestand sind hinter jedem verfügbaren Artikel die Artikelnummer und der aktuelle Bestand hinterlegt. In der zweiten Komponente sind die offenen Bestellungen mit Bestellnummern, den dazugehörigen Artikelnummern und den bestellten Mengen enthalten.

Die Struktur von „Rita“ lässt sich in vier Menüpunkte unterteilen:

- Start (1)
- Bestandsabfrage (2)
- Bestandsbuchung (3)
- Aufgabenliste (4).

Zu Beginn kann die nutzende Person den Sprachassistenten mit **„Ok Google, ich möchte mit Rita sprechen“** aktivieren. Nach diesem „Aufwecken“ des digitalen Sprachassistenten wird der Nutzende begrüßt und nach der gewünschten Aktion gefragt. Sollte es Verständnisschwierigkeiten geben, was beispielsweise beim Sprechen unter einer FFP2-Maske oder bei lauten Hintergrundgeräuschen der Fall sein kann, fordert „Rita“ die sprechende Person auf, die Eingabe zu wiederholen, und gibt ggf. Hinweise, welchen Wortlaut sie als Eingabe verstehen kann. Darüber hinaus hat man die Möglichkeit, den Sprachassistenten mit Stichworten wie „Ende“ oder „Stopp“ zu beenden.



Da der Sprachassistent selbstlernend ist, wird „Ritas“ Wortschatz, die Bestätigung ihres „Lernerfolgs“ durch den Nutzenden vorausgesetzt, stetig erweitert, sodass ihr Verständnis und die richtige Zuordnung der Eingaben laufend verbessert werden.

Im Menüpunkt Bestandsabfrage (2) fragt „Rita“ nach einer Artikelnummer oder einer GTIN. Sollte der Sprachbefehl nicht verstanden werden, erfolgt auch hier eine Rückfrage.

Genauso erfolgt die Buchung von Wareneingängen (3). Zur Bestimmung des zu verbuchenden Eingangs muss die nutzende Person eine in der Datenbank hinterlegte Bestellnummer und die eingegangene Menge nennen. Anhand der Eingabe wird die offene Bestellmenge in der Warenbestellungs-Datenbank, um die genannte Anzahl reduziert und der Warenbestand entsprechend erhöht. Wenn die bestellte Menge vollständig geliefert wurde, wird die Bestellung in der Datenbank ausgeblendet.

Auch die Aufgabenfunktion (4) wird durch entsprechende Schlagwörter aufgerufen.

Hier informiert „Rita“ den Benutzenden über die neuen Wareneingänge. Als zusätzliche Information weist sie darauf hin, dass schlechtes Wetter vorausgesagt ist, und verbindet dies mit der Empfehlung, den Bestand der Regenjacken auf der Fläche und im Lager zu prüfen und eine bestimmte Bestandsgröße vorzuhalten. Von hier aus kann man mit den bekannten Schlagwörtern in die Bereiche Bestandsprüfung (2) und Warenbuchung (3) springen.

### Was wird benötigt?

- Ein Google-Benutzerkonto
- Ein Lautsprecher oder ein Display mit Google-Assistant-Funktion
- Ein Smartphone/Tablet, das folgende Anforderungen erfüllt:
  - Betriebssystem: Android 6.0 oder höher
  - Mit einem 2,4-GHz- und 5-GHz-WLAN kompatibel (Ein WPA2-Enterprise-Netzwerk ist nicht ausreichend)
- Die aktuelle Version sowohl der Google-App, der Google-Voice-App als auch der Google Home App
- Eine aktivierte Bluetooth-Schnittstelle
- Eine Internetverbindung und ein sicheres WLAN
- Beide Geräte (Smartphone/Tablet und das Google-Assistant-Gerät) müssen sich im selben WLAN befinden

Zudem muss die Anmeldung auf den Geräten und in den Apps mit demselben Google-Konto erfolgen!

# 8. Learnings: Technische und alltägliche Herausforderungen

Im Laufe des Projektes kristallisierten sich einige Herausforderungen bei der Entwicklung eines Sprachassistenten heraus.

## ■ **WLAN-Verfügbarkeit**

Voraussetzung zur Nutzung des Sprachassistenten ist eine bestehende Internetverbindung, dazu sollten die örtlichen Gegebenheiten wie beispielsweise die Bausubstanz des Geschäfts im Vorfeld berücksichtigt werden. Um miteinander kommunizieren zu können, müssen sich zudem die verschiedenen Geräte, Tablet/Display und das ggf. zur Steuerung verwendete Smartphone, im selben WLAN befinden oder über denselben Hot Spot eingewählt sein.

## ■ **Funktionsänderungen nach Updates**

Eine weitere technische Herausforderung waren eingespielte Updates mit Erweiterungen oder

Änderungen der genutzten Funktionen. Dies kam im Projektverlauf zwar nur einmal vor und war im Grunde genommen eine vorteilhafte Funktionsänderung, man sollte jedoch ein Auge darauf haben.

## ■ **Anzahl der benötigten Google-Benutzerkonten**

Damit verschiedene Personen auf den Sprachassistenten zugreifen können, benötigt jede:r ein eigenes Google-Benutzerkonto, was dann mit dem Sprachassistenten verknüpft werden muss.

## ■ **Die mögliche Sensibilität des Sprachassistenten**

Damit die Spracheingabe der Nutzenden erfolgen kann, müssen diese auch als berechtigte Personen erkannt und verstanden werden. Da im Projektverlauf die FFP2-Maskenpflicht eingeführt wurde, wurden vorhandene Stimm-

## **Auf ein Wort: Der Datenschutz**

Damit der Digitale Sprachassistent seinen „Weckruf“ vernehmen und darauf reagieren kann, muss das Gerät aktiviert und auf die Umgebung sensibilisiert sein. Welche Informationen in der Zwischenzeit gefiltert werden, ist von Anbieter zu Anbieter unterschiedlich und wird bisher nicht vollkommen transparent kommuniziert.

Die meisten Sprachassistenten bieten jedoch eine Möglichkeit an, um die Spracherkennung auf Wunsch ab- und das Gerät somit „gehörlos“ zu schalten.

Hinzu kommt, dass die Spracheingaben auf den Servern bzw. in der Cloud des jeweiligen Anbieters verarbeitet werden. Diese Server befinden sich in der Regel in Rechenzentren in den USA. Zwar unterliegen in Deutschland eingesetzte Systeme auch der hier gültigen Datenschutz-Grundverordnung (DSGVO), dennoch lohnt sich im Zweifel immer ein Blick in die Datenschutzrichtlinien des Anbieters.

profile teilweise nicht mehr auf Anhieb erkannt und mussten neu eingesprochen werden.

■ **Sprachsteuerung nur über einen bestimmten Wortlaut**

Gerade in der Eingewöhnungsphase des Systems benötigt der Sprachassistent bestimmte Wortfolgen, um die gewünschte Aktion aufzurufen. Befolgt man dabei nicht den genauen Wortlaut, birgt dies anfänglich ein hohes Frustrationspotenzial.

■ **Die Eingabe langer Artikelnummern**

Die Spracheingabe der 14-stelligen GTIN birgt Fehlerpotenzial. Eine mögliche Abhilfe wäre hier – sofern vorhanden – die Verwendung eigener, kürzerer Artikelnummern.



# 9. Fazit und Ausblick

**D**ie Fortschritte der Spracherkennung durch innovative Algorithmen der Künstlichen Intelligenz lassen erwarten, dass Digitale Sprachassistenten eine breite Nutzung in Privat- und Geschäftsumgebungen erreichen werden. Allerdings fristen die Digitalen Sprachassistenten derzeit noch ein Nischendasein im Handel.

Einzelne „Pilotprojekte“ auch in anderen Branchen zielen überwiegend auf die Kundenkommunikation ab und dienen dazu, sich als innovatives Unternehmen zu positionieren. Allerdings bleibt abzuwarten, ob die Kundschaft tatsächlich mit Maschinen sprechen möchten oder nicht doch lieber die persönliche Kommunikation zwischen Menschen vorziehen.

Kundenferne Anwendungen sind teilweise über das Erprobungsstadium hinaus und binden die Digitalen Sprachassistenten, beispielsweise mit

Pick-by-Voice im Kommissionierungslager, erfolgreich in bestehende Arbeitsprozesse mit ein. Noch ist vergleichsweise wenig über das Unterstützungspotenzial interner Abläufe bekannt.

Dieser Leitfaden bietet – auf Basis des eigens entwickelten Sprachassistenten „Rita“ – bereits einen Ausblick auf die Einsatzmöglichkeiten Digitaler Sprachassistenten im Tagesgeschäft des Einzelhandels. Dabei zeigen sich der grundsätzliche Nutzen und auch der weite Weg, der noch zurückzulegen ist, bis eine flüssige Sprachsteuerung und -assistenz im Geschäft flächendeckend ankommt. Es ist jedoch anzunehmen, dass durch die zunehmende Vernetzung der Systeme und die steigende Anzahl der Geräte, die mit einem Sprachassistenten interagieren können, die Einsatzfelder zukünftig weiter zunehmen werden. Eine Auseinandersetzung mit dieser sich schnell entwickelnden Technologie bleibt daher unerlässlich.

# Impressum

Herausgeber:  
Mittelstand 4.0-Kompetenzzentrum Handel  
c/o EHI Retail Institute GmbH  
Spichernstraße 55, 50672 Köln

Geschäftsführung:  
Michael Gerling  
Registergericht Amtsgericht Köln  
Registernummer HRB 1941

[www.kompetenzzentrumhandel.de](http://www.kompetenzzentrumhandel.de)  
T +49 / (0)221 / 57993 –992  
F +49 / (0)221 / 57993 – 45  
[info@kompetenzzentrumhandel.de](mailto:info@kompetenzzentrumhandel.de)

Im Text wird aus Gründen der besseren Lesbarkeit häufig nur die männliche Form verwendet. Selbstverständlich sind alle Personen gemeint.

Soweit keine redaktionelle Kennzeichnung für den Inhalt, Verantwortliche im Sinne des Presserechts und des Rundfunkstaatsvertrags:

Michael Gerling  
Spichernstraße 55, 50672 Köln

Text:  
Simone Sauerwein, EHI  
Nicole Bongartz, GS1

Gestaltung:  
Huyen Nguyen-Le, S3 Advertising GmbH & Co. KG

Bildquellen:  
Titel: Erikona/istockphoto.com  
Seite 4: Prostock-Studio/istockphoto.com  
Seite 6-7: Vertigo3d/istockphoto.com  
Seite 13: Ralf Geithe/stock.adobe.com  
Seite 16: Natalya Burova/istockphoto.com  
Seite 19: LuckyStep48/istockphoto.com

Druckerei:  
Cede Druck

Auflage:  
500

Stand:  
Juni 2021

# Literaturverzeichnis

**Amazon (o.A.):** Alexa Skills.

[https://www.amazon.de/b/?ie=UTF8&node=10068460031&ref\\_=topnav\\_storetab\\_a2s](https://www.amazon.de/b/?ie=UTF8&node=10068460031&ref_=topnav_storetab_a2s) ;  
abgerufen am 14.01.2021

**Apple (o.A.):** Siri macht jetzt noch mehr. Schon bevor du fragst.

<https://www.apple.com/de/siri/> ; abgerufen am 14.01.2021

**Apple (o.A.\_b):** Zubehör: HomeKit.

<https://www.apple.com/de/shop/accessories/all/homekit> ; abgerufen am 14.01.2021

**BVDW/Bundesverband Digitale Wirtschaft e.V. (2019):** Sprachassistenten im Smart Home.

[https://www.bvdw.org/fileadmin/bvdw/upload/publikationen/smart\\_home/Sprachassistenten\\_im\\_SmartHome.pdf](https://www.bvdw.org/fileadmin/bvdw/upload/publikationen/smart_home/Sprachassistenten_im_SmartHome.pdf)

**DIVSI (04.11.2014):** .

[www.divsi.de/publikationen/studien/wissenswertes-ueber-den-umgang-mit-smartphones/5-basisdienste-und-datenermittlung/5-3-sprachsteuerung-aktivieren/index.html](http://www.divsi.de/publikationen/studien/wissenswertes-ueber-den-umgang-mit-smartphones/5-basisdienste-und-datenermittlung/5-3-sprachsteuerung-aktivieren/index.html) ; abgerufen am 20.01.2021

**EHI (2019):** EHI-Whitepaper „Smart Store: Teil 2“.

<https://www.ehi-shop.de/en/studies/whitepaper/whitepaper-smart-store-2> ; abgerufen am 20.01.2021

**Google (o.A.):** Hey Google: Wo er dir hilft: Jetzt auch in deinem Auto.

[https://assistant.google.com/intl/de\\_de/platforms/cars/](https://assistant.google.com/intl/de_de/platforms/cars/) ; abgerufen am 14.01.2021

**IDC (2021):** IDC Forecasts Double-Digit Growth for Smart Home Devices as Consumers Embrace Home

Automation and Ambient Computing; <https://www.idc.com/getdoc.jsp?containerId=prUS47567221> ;  
abgerufen am 08.04.2021

**Mittelstand 4.0-Kompetenzzentrum Rostock (2020):** Leitfaden Smart-Hotel.

<https://www.kompetenzzentrum-rostock.digital/wp-content/uploads/2020/07/Leitfaden-Smart-Hotel-14022020-digital.pdf> ; abgerufen am 21.01.2021

**Munster, Gene & Thompson, Will (2018):** Annual Smart Speaker IQ Test. <https://loupventures.com/annual-smart-speaker-iq-test/> ; abgerufen am 08.01.2021

**proLogistik (o.A.):** Logistik-Lexikon: Sprachgestütztes Kommissionieren.  
<https://www.prologistik.com/logistik-lexikon/sprachgestuetztes-kommissionieren/> ;  
abgerufen am 21.01.2021

**reichelt elektronik Magazin (2017):** Themen-Specials: Das Smart Home – mit Sprachsteuerung wird es richtig smart!. <https://www.reichelt.de/magazin/themen-projekte/das-smart-home-mit-sprachsteuerung-wird-es-richtig-smart/> ; abgerufen am 14.01.2021

**Samsung (o.A.):** Smart TV: Smart Home mit Smart Home Geräten.  
<https://www.samsung.com/de/tvs/smart-tv/smart-home-with-iot-devices/> ; abgerufen am 14.01.2021

**Schaber, Krieger-Lamin & Peissl (2019):** Digitale Assistenten.  
(Institut für Technikfolgen-Abschätzung der Österreichischen Akademie der Wissenschaften, Hrsg.).  
<https://doi.org/10.1553/ITA-pb-2019-01>

**SPLENDID RESEARCH GmbH (2019):** Digitale Sprachassistenten.  
<https://www.splendid-research.com/de/studie-digitale-sprachassistenten.html> ;  
abgerufen am 06.12.2020

**Wendel, Mariella (2020):** Home&Smart: Alexa, Siri, Google, Cortana & Bixby Sprachassistenten-Test: Die 5 beliebtesten Sprachassistenten im Überblick;  
<https://www.homeandsmart.de/smart-home-sprachassistenten> ; abgerufen am 14.01.2021



## Über das Mittelstand 4.0-Kompetenzzentrum Handel

Das Mittelstand 4.0-Kompetenzzentrum Handel gehört zu Mittelstand-Digital. Mit Mittelstand-Digital unterstützt das Bundesministerium für Wirtschaft und Energie die Digitalisierung in kleinen und mittleren Unternehmen und dem Handwerk.

Weitere Informationen unter  
[www.kompetenzzentrumhandel.de](http://www.kompetenzzentrumhandel.de)

Mittelstand-  
Digital

## Was ist Mittelstand-Digital?

Mittelstand-Digital informiert kleine und mittlere Unternehmen über die Chancen und Herausforderungen der Digitalisierung. Die geförderten Kompetenzzentren helfen mit Expertenwissen, Demonstrationszentren, Best-Practice-Beispielen sowie Netzwerken, die dem Erfahrungsaustausch dienen. Das Bundesministerium für Wirtschaft und Energie (BMWi) ermöglicht die kostenfreie Nutzung aller Angebote von Mittelstand-Digital.

Der DLR Projektträger begleitet im Auftrag des BMWi die Kompetenzzentren fachlich und sorgt für eine bedarfs- und mittelstandsgerechte Umsetzung der Angebote. Das Wissenschaftliche Institut für Infrastruktur und Kommunikationsdienste (WIK) unterstützt mit wissenschaftlicher Begleitung, Vernetzung und Öffentlichkeitsarbeit.

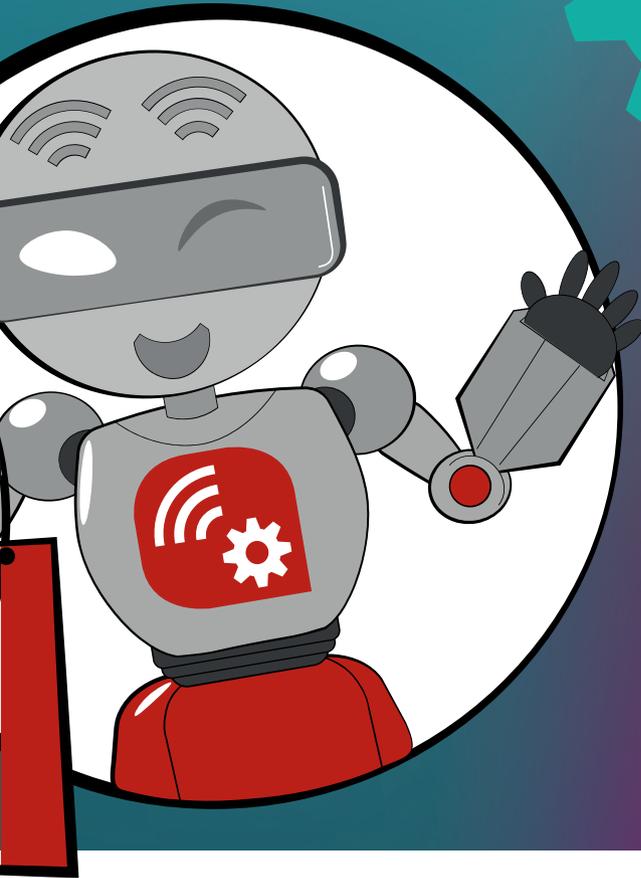
Weitere Informationen finden Sie unter  
[www.mittelstand-digital.de](http://www.mittelstand-digital.de)







**Mittelstand 4.0**  
Kompetenzzentrum  
Handel



Online  
Sichtbarkeit

Innovationen  
auf der  
Fläche

Prozesse  
digitalisieren

Wir unterstützen Händler  
kostenfrei bei der Digitalisierung!

[kompetenzzentrumhandel.de](https://www.kompetenzzentrumhandel.de)

-  Mittelstand 4.0-Kompetenzzentrum Handel
-  @Mittelstand 4.0-Kompetenzzentrum Handel
-  @handelkompetent

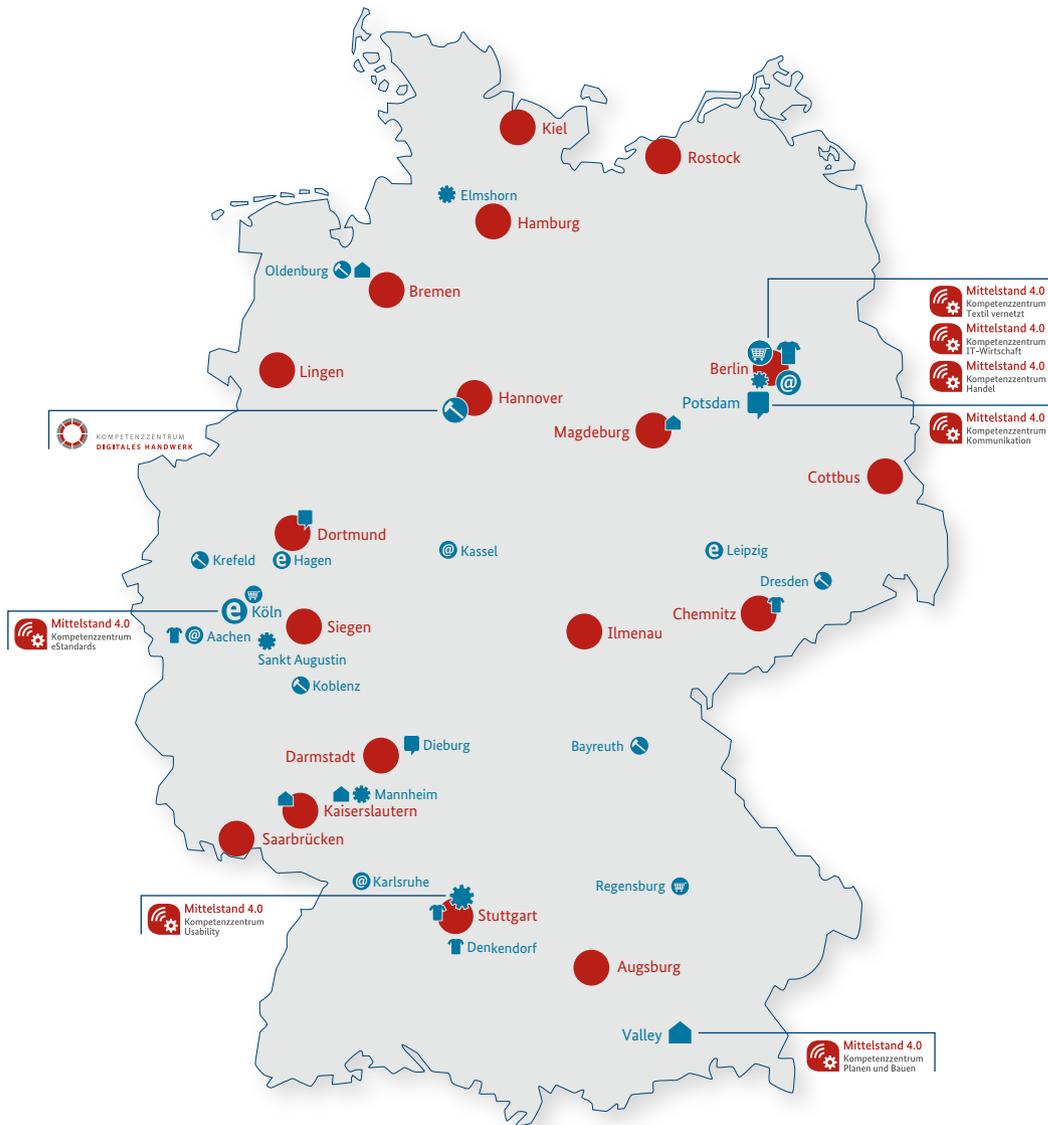
Mittelstand-  
Digital 

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages

# Die regionalen Mittelstand 4.0-Kompetenzzentren und Themenzentren mit ihren Stützpunkten



Weitere Informationen finden Sie unter  
[www.kompetenzzentrumhandel.de](http://www.kompetenzzentrumhandel.de)