

„Alexa, frage KIT-Bibliothek wo ich lernen kann“

Einsatz von digitalen Sprachassistenten im Umfeld von Bibliotheken und Erweiterung um eigene Dienste

Uwe Dierolf und Michael Skarupianski

Das Thema sprachgesteuerte persönliche Assistenten ist im wahrsten Sinne des Wortes in aller Munde. Die KIT-Bibliothek hat daher Ende 2017 entschieden, sich mit diesem Thema unter technischen Gesichtspunkten exemplarisch am Beispiel Amazon Alexa zu beschäftigen. Dazu wurde Anfang 2018 begonnen, den Alexa Skill „KIT-Bibliothek“ zu entwickeln, um Erfahrungen mit den von Amazon bereitgestellten Entwicklerwerkzeugen und dem Review-Prozess zu sammeln, der ähnlich wie bei der App-Entwicklung durchlaufen werden muss. Dieser Beitrag gibt Einblick in die gewonnenen Erfahrungen und Probleme und verweist auf ähnliche Entwicklungen an anderen Bibliotheken.

Voice assistants can be found in many households and more and more people are using such devices. At the end of 2017 the KIT Library decided to engage in this topic by developing an Alexa Skill called „KIT-Bibliothek“. The reasons for doing this step were primarily cause of gaining experience and was a technique-driven approach. The process of developing a skill is similar to the development of a smartphone app and includes a review process to publish the skill. This reading delivers an insight into the experiences we made but shows also the challenges and gives a short overview of developments made by other libraries.

Sprachassistentenz – mehr als nur ein Hype?

Ob Apples Siri, der Google Assistant oder Amazons Alexa – fast wie eine Flut kommen jeden Monat neue Geräte mit eingebauten Sprachassistenten auf dem Markt und buhlen um die Gunst der Kunden. Unbestrittener Vorreiter war dabei Apples Siri auf dem iPhone, jedoch unterscheidet man mittlerweile den reinen Software-Assistenten (z.B. Siri) und das Hardware-Gerät, auf dem dieser ausgeführt wird (z.B. iPhone).

Die Software-Hersteller gehen dabei unterschiedliche Wege. Während Apple seine Assistentin exklusiv auf die eigene Produktlinie beschränkt (iPhone, Apple TV, Homepod, etc.), vermarktet Google den Assistenten auf eigenen Lautsprechern, als App auf Android und iOS Geräten, sowie in Lizenz an diverse Hardware Hersteller. Auch Amazon geht diesen Schritt und baut nicht nur selber Hardware mit dem hauseigenen Assistenten „Alexa“, sondern ermöglicht auch Fremdherstellern diesen in (fast) beliebige Hardware einzubinden. Die Tatsache, dass Amazon der erste Hersteller mit einem reinen smarten Lautsprecher war, die aggressive Vermarktung und ein niedriger Einstiegs-

preis sind der Grund für einen großen Ansturm auf die smarten Lautsprecher von Amazon. Genaue Zahlen gibt es von den Herstellern nicht, allerdings hat Amazon alleine im dritten Quartal 2017 mindestens 20 Millionen Echo-Geräte verkauft. Laut einer repräsentativen Umfrage haben 16% der US-Amerikaner einen smarten Lautsprecher zu Hause stehen. Dabei hat Amazon einen Marktanteil von fast 70% [1]. Aus diesen Gründen fiel die Wahl bei der Auswahl eines möglichen Assistenten für die KIT-Bibliothek auf Amazons Alexa.

Wie bereits erwähnt ist Alexa der reine Software-Assistent. Um diesen zu benutzen wird eine Hardwarekomponente benötigt. Die erste verfügbare Hardware war der von Amazon selbst vertriebene Lautsprecher „Amazon Echo“. Die günstigste Variante nennt sich „Echo Dot“, ist kaum größer als ein Eishockey Puck und besteht aus einem Lautsprecher, vier Tasten sowie 7 Fernfeld-Mikrofonen (diese sorgen dafür, dass man auch trotz Nebengeräuschen oder aus dem nächsten Zimmer verstanden wird). Das obere Ende der Produktpalette der Smarten Lautsprecher bildet der „Echo Show“ mit eingebautem Bildschirm. Dazwischen gibt es noch den klassischen „Echo“ (etwa dreimal so groß wie der „Echo Dot“ und mit besserem Lautsprecher, den „Echo Plus“ mit eingebautem Smart-Home Hub und den „Echo Spot“, einen Radiowecker großen Lautsprecher mit kleinem Bildschirm. Der Normalpreis für den Echo Dot liegt bei 60 €, allerdings lockt Amazon alle paar Wochen mit Sonderangeboten und bietet diesen schon ab 35 € an, beim Oberklasse-Modell „Echo Show“ mit Bildschirm liegt der Normalpreis bei 220 €.

Neben den Geräten bietet Amazon seinen Assistenten auch in den Apps für Android (Alexa-App und Amazon Music) sowie in der Fernbedienung zur hauseigenen Set-Top-Box FireTV an.

Grundsätzlich haben alle Geräte eines gemeinsam: Sie bestehen aus einem oder mehreren Mikrofonen um die Frage des Benutzers aufzunehmen und einem Lautsprecher um die Antwort auszugeben. Zum Betrieb benötigen Sie zudem eine Internetverbindung. Auf dieser Basis erlaubt es Amazon Fremdfirmen,

eigene Geräte zu entwickeln oder die Technik in bereits vorhandene Hardware zu integrieren. Da Amazon seine Geräte zuerst in den USA veröffentlicht, ist hier bereits ein großer Markt für Hardware mit Alexa-Integration entstanden. So gibt es z.B. Kopfhörer mit eingebauter Alexa oder kleine Sticks, die man im Auto in den Zigarettenanzünder steckt und die so das normale Auto in ein Smart Car verwandeln sollen. Für Mitte des Jahres ist sogar die erste Smartwatch angekündigt, die den Sprachassistenten eingebaut hat [4]. Auch die Autohersteller selber integrieren Alexa in ihre Infotainment-Systeme. Ford und Seat bieten entsprechend ausgestattete Autos bereits seit 2017 auf dem US-Markt an. BMW und Toyota haben angekündigt, dieses Jahr nachzuziehen [5].

Die Frage die sich nun unweigerlich stellt, ist: „Was kann ich damit tun?“.

Um den Lautsprecher zu aktivieren sagt man zuerst immer den Namen des Assistenten und stellt danach die gewünschte Frage. Für Faktenfragen sagt man z.B. „Alexa, was ist die Hauptstadt von Tuvalu?“. Der Lautsprecher antwortet mit: „Die Hauptstadt von Tuvalu ist Funafuti.“ Dabei wurde nach dem Aktivierungswort „Alexa“ automatisch die Aufnahme gestartet. Nach dem Ende des Satzes wurde diese Audioaufnahme in die Amazon Cloud geladen, dort analysiert und im Anschluss die Antwort (als Text) zurückgeschickt. Der lokale Lautsprecher hat diese Antwort dann vorgelesen.

Wenn Sie dieses Beispiel selbst testen möchten, genügt ein Amazon Account und ein Smartphone. Installieren Sie unter Android die Amazon Alexa-App [2] oder unter iOS die App „Amazon Music“ [3].

Eine der am häufigsten genutzten Funktionen ist natürlich auch das Abspielen von Musik [1]. Amazon macht hier natürlich Werbung für den hauseigenen Dienst „Amazon Music“, mit einem Echo-Lautsprecher können Sie aber auch kostenlos über das Radionetzwerk Tuneln auf Streams von Radiostationen zugreifen. Sagen Sie dazu einfach „Alexa, spiele Surf Radio“ und Alexa spielt den entsprechenden Stream ab.

Alexa kann natürlich eine schier unendliche Menge von Fragen beantworten und Aktionen ausführen. Sei es das Wetter („Alexa, wie wird das Wetter heute?“), die Nachrichten („Alexa, was sind die Nachrichten?“), ein Timer („Alexa, stelle Teetimer auf 5 Minuten“) oder die Einkaufsliste („Alexa, setze Nudeln auf meine Einkaufsliste“) – auf alles hat Alexa bereits die Antwort eingebaut. Dabei wird das System seitens Amazon ständig erweitert – der neueste Clou sind (Video-)gespräche zwischen Alexa-fähigen Geräten.

Und wenn dies alles nicht reicht – Amazons System

kann sowohl per Hardware als auch per Software erweitert werden. Dritthersteller können Ihre Produkte so erweitern, dass diese mittels Alexa gesteuert werden. Es gibt bereits jetzt schon steuerbare Glühbirnen, Thermostate, Steckdosen, Saugroboter und Fernseher. Sie müssen nur „Alexa, schalte das Licht in der Küche ein“ sagen und der Computer erledigt den Rest. Oft sind die Preise natürlich höher als bei gleichen nicht-smarten Geräten, aber durch die höhere Marktverbreitung sind auch diese gefallen und so gibt es z.B. per Alexa steuerbare Glühbirnen schon für unter 10 €.

Auf die Spitze treiben es Nissan und Ford, bei denen mittels Alexa der Benzinstand abgefragt, das Licht im Auto ein und ausgeschaltet, oder der Motor gestartet und gestoppt werden kann (wohlgemerkt auch aus der Ferne, nicht nur, wenn man selbst im Auto sitzt) [6][7]. Und wenn man glaubt, das ganze kann nicht überboten werden: Auf Kickstarter gibt es eine Kampagne für eine beheizbare Jacke, die mittels Alexa an- und ausgeschaltet werden kann [8].



Abbildung 1: Echo-Familie (Echo Dot, Echo, Echo Plus) im Größenvergleich zu einer Computermaus.

Neben Hardware kann das System auch mit Programmen (den sogenannten **Skills**) erweitert werden. Diese Skills können mittels Sprache oder über die Alexa-App aktiviert werden. Möchte man den Skill „Fernsehprogramm“ installieren, genügt ein „Alexa, aktiviere den Skill Fernsehprogramm“. Im Anschluss ist dieser sofort verfügbar und man kann ihm Fragen stellen wie „Alexa, frage Fernsehprogramm, was heute Abend läuft“. Im Moment sind laut Amazon über 30000 solcher Skills verfügbar, egal ob man seine Gitarre stimmen möchte („Alexa, aktiviere Gitarrenstimmer“) oder über die neuesten Weinfeste informiert sein will („Alexa, aktiviere Weinfestkalender“). In der Amazon-Suche [9] erhält man einen Überblick über verfügbare Skills. Da die Skills nicht tatsächlich auf dem Gerät installiert werden, sondern lediglich in der Cloud aktiviert sind, ist die Anzahl der aktiven Skills theoretisch nicht limitiert. Ebenso ist

ein aktivierter Skill auf allen Alexa-Geräten desselben Accounts verfügbar.

Warum beschäftigt sich eine Bibliothek mit dem Thema Sprachassistenten?

Das Thema wurde in erster Linie unter technischen Gesichtspunkten an der KIT-Bibliothek Ende 2017 in Angriff genommen. Aber natürlich sollen auch mögliche Einsatzszenarien evaluiert werden. Dazu im Folgenden ein paar Beispiele.

Die KIT-Bibliothek ist seit April 2006 eine 24/7-Bibliothek. Täglich besuchen knapp 3000 Personen die KIT-Bibliothek. Bei so vielen Menschen wird natürlich die Auskunft rege genutzt, um auf allerlei Fragen Antworten zu erhalten. Die Auskunft ist werktags bis 19 Uhr und am Samstag bis 12.30 Uhr besetzt. Zu allen anderen Zeiten ist lediglich Wachpersonal der Campus-sicherheit anwesend. Dieses auch häufig wechselnde Personal kann nicht in allen Fällen adäquat Auskunft auf Fragen geben.

Daher ist ein naheliegendes Szenario für den Einsatz von Sprachassistenten die Beantwortung von Faktenfragen. Der Einsatz einer Alexa-Box scheidet aus Datenschutzgründen natürlich aus (aber dazu später mehr). Alexa kann jedoch auch direkt vom Smartphone aus genutzt werden, um Fragen nach Räumlichkeiten „wo befindet sich das Medienzentrum“, „wo finde ich Bücher zum Thema ...“ zu beantworten. Weitere ggfs. für manche Nutzer relevante Fakten sind Öffnungszeiten oder die Nennung von Telefonnummern, Räume von Fachreferenten oder Direktion etc.

Auch die sprachgesteuerte Nutzung dynamischer Services wie dem Seatfinder zur Ermittlung freier Arbeitsplätze ist sicherlich eine nützliche Angelegenheit für manche Menschen. Auf die Frage „wo kann ich lernen“ kann der Skill die maschinenlesbaren Schnittstellen des Seatfinders verwenden, um Auskunft über freie Lernräume zu geben. Dies wurde bereits für den Skill „KIT-Bibliothek“ implementiert.

Als ein mögliches Szenario, bei dem Alexa sogar sinnvoll zu Hause via Alexa-Box genutzt werden könnte, wäre die Beantwortung von Fragen nach der Rückgabe von Büchern mit ggfs. durchzuführender Einzel- oder Pauschalverlängerung. Bis zur Klärung der Frage nach dem Datenschutz ist die Kontoverknüpfung bisher allerdings nur angedacht, jedoch noch nicht implementiert worden.

Die UB Bamberg hat für ihren Skill „Unibib“ diese Verbindung von Amazon-Konto und Bibliothekskonto sogar schon implementiert. Allerdings ist dieser Skill noch nicht für die Öffentlichkeit sichtbar, da auch hier noch nicht sicher ist, was rechtlich erlaubt ist. Sie bieten auch den umfangreichsten Skill an, der auch die

Neuigkeiten zur UB umfasst.

In Österreich hat sich die FH Burgenland mit dem Thema „Alexa Skill“ beschäftigt und bietet den **Skill „Bücherregal“** an – allerdings auch nur im Dev-Modus und somit nicht sichtbar für alle Nutzer.

Wie im Fall der UB Bamberg wird hier nach dem Starten des Skills nicht nur erklärt, was der Skill kann, sondern es wird dem Benutzer die Frage gestellt, für welches Thema er sich interessiert. Das ist insofern sehr praktisch, da man nur den Namen des Themas nennen muss, um eine Antwort zu erhalten. Der Nutzer kann den Skill also auch nutzen, ohne die Art und Weise zu kennen, wie man mit dem Skill sprechen muss, um verstanden zu werden. Dieses Thema „was muss ich tun, damit mich Alexa versteht?“ wird im kommenden Kapitel genauer betrachtet.

Wie funktioniert Spracheingabe bei Alexa?

Jeder Skill besitzt neben dem eigentlichen Namen noch einen **Aufrufnamen („invocation name“)**. Der Aufrufname muss aus zwei Wörtern bestehen (Ausnahmen sind nur für Markennamen und zusammengesetzte Namen erlaubt) und wird benutzt um den Skill zu aktivieren oder zu starten. Ein Skill kann z.B. den Namen „Virtuelle Buchsuche mit Buchguru“ haben, wird aber mit „buch guru“ aufgerufen.

Die Funktionen, die ein Skill zur Verfügung stellt werden **„Intents“ (Absichten)** genannt. Für jede Funktion, die ein Skill zur Verfügung stellen soll, muss ein Intent definiert werden (z.B. ein Intent für die Suche, einer für die Auskunft, etc.). Schließlich müssen zu jedem Intent passende Beispieläußerungen (im Alexa-Jargon heißen die Äußerungen **„Utterances“**) angegeben werden.

Der Buchguru hat z.B. einen Intent Buchsuche. Nun gibt man einen Beispielsatz an, was der Benutzer sagen würde, wenn er diesen Intent ausführen will. Soll der Skill z.B. auf „Alexa frage buch guru nach Neuerscheinungen“ reagieren, würde der Beispielsatz „nach Neuerscheinungen“ lauten.

Damit Alexa weiß, welcher Intent gemeint ist, sollte man möglichst viele Variationen angeben. Immerhin kann der Skill nicht nur mit dem Schlüsselwort „frage“, sondern auch mit „sage“, „lade“, „starte“, „öffne“ und noch weiteren Möglichkeiten eingebunden werden. So wären für die Neuerscheinungen auch folgende Beispiele möglich: „Alexa, starte buch guru und suche nach Neuerscheinungen“, „Alexa rufe bei buch guru die Neuerscheinungen ab“ sowie „Alexa, was sind die Neuerscheinungen von buch guru“.

Dazu würde man passende Utterances erstellen: „suche nach Neuerscheinungen“ bzw. nur „Neuerscheinungen“.

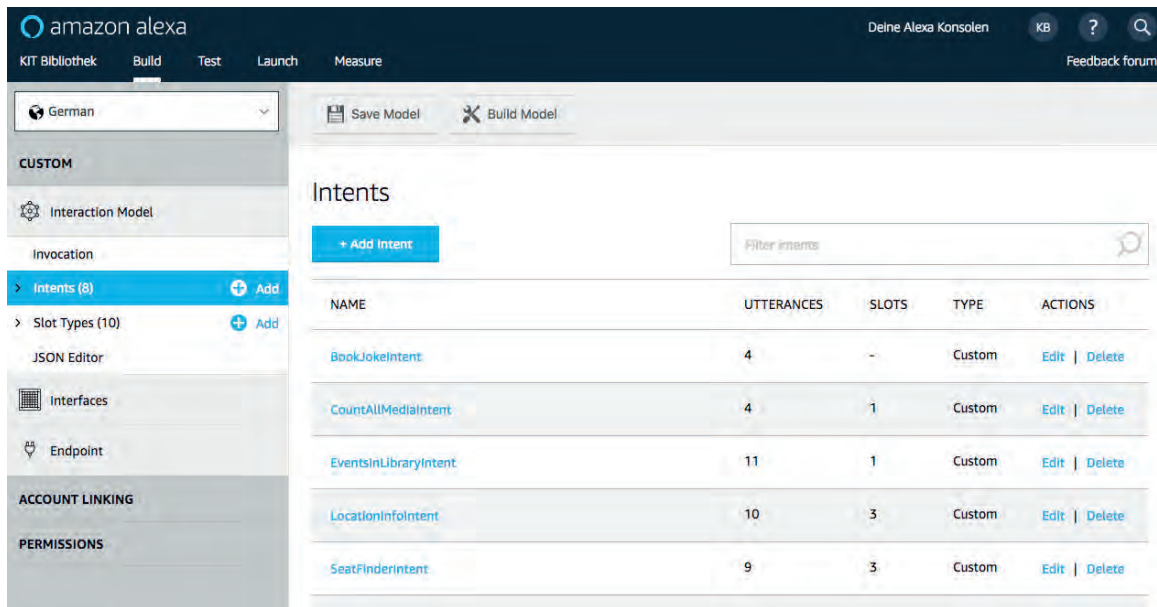


Abbildung 2:
Alexa Developer
Konsole mit bereits
erstellten Intents
(Absichten)

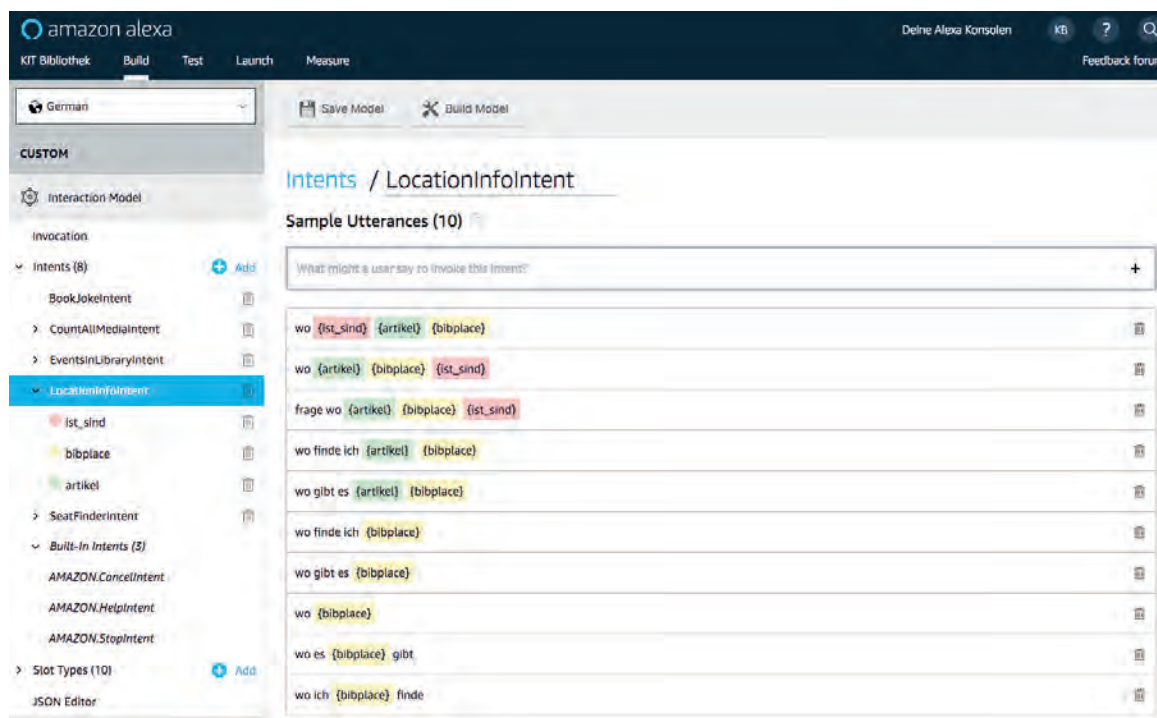


Abbildung 3:
Mögliche Bei-
spielsätze für
die Lokalisierung
eines Gerätes
oder Ortes
innerhalb der
Bibliothek

Je komplizierter der Intent ist, desto komplizierter können auch die Beispielsätze werden. Möchte man den Neuerscheinungen-Intent um eine Datumsfunktion erweitern, muss man nicht alle Möglichkeiten eintippen (z.B. Neuerscheinungen diese Woche, Neuerscheinungen nächste Woche, Neuerscheinungen gestern), sondern kann hier **Variablen** einsetzen (diese werden bei Alexa „Slots“ genannt). So genügt die Angabe eines Slots in Klammern z.B. „Neuerscheinungen für {ERSCHEINUNGS_DATUM}“ dem Slot ERSCHEINUNGS_DATUM kann man dann entweder einen vordefinierten Typ (Datum) zuweisen oder einen eigenen Typ definieren. Beim Datum würde man den vorhandenen Typ AMAZON.DATE auswählen und Amazon erkennt automatisch das Datum und liefert es an den Entwickler

zurück. Wenn kein vorgegebener Typ (Nummer, Daten, Regionen, Vornamen, etc..) passt, kann man auch eigene definieren (z.B. für Buchgenres). Dabei sind die Slots nur als mögliche Beispiele zu sehen, d.h. sie sollten z.B. zwar möglichst viele Buchgenres angeben, es müssen aber nicht alle angegeben werden.

Ein leichtgewichtiger Ansatz

Die KIT-Bibliothek hat mit einem leichtgewichtigen Ansatz einen Prototypen eines Skills entwickelt. Für diesen war es nötig einen kostenlosen Developer-Account bei Amazon anzulegen. Unterstützt wird man von einer umfangreichen Dokumentation, vielen Videos, Beispielen und kostenlosen Webinaren. Dabei muss jedoch beachtet werden, dass fast alle Medien

auf Englisch sind und vorrangig die Amazon-Dienste benutzen. Die Entwicklung findet in der sog. „Alexa Skill Kit Developerkonsole“ statt. Dies ist eine Sammlung von Online-Tools, die beim Erstellen eines Skills unterstützen.

Bereits beim Einrichten des Namens gab es das erste Problem, da Amazon das Wort KIT nicht als Abkürzung einzelner Buchstaben (kah-iih-tee) sondern als ein Wort (KIT wie in Fugenkit) ansieht. Hier wurde also der Skillname „KIT-Bibliothek“ und der Rufname „k. i. t. bibliothek“ benutzt.

Als Funktionen wurde die Abfrage des Seatfinders (eines Systems zur Lokalisierung von freien Arbeitsplätzen in Bibliotheken), die Lokalisierung von Orten („Wo finde ich“), das Zählen aller Medien und die Abfrage der aktuellen Veranstaltungen gewählt.

miersprache auf dem eigenen Server). Da die KIT-Bibliothek Webservices erfolgreich mit PHP einsetzt, wurde dieser Weg gewählt.

Amazon verschickt seine Anfragen in einem bestimmten JSON-Format und erwartet auch eine entsprechende JSON-Antwort. Aus diesem Grund wurde ein kleines externes PHP-Framework eingesetzt, um die Anfragen zu parsen und Antworten zu verschicken [10]. Mit diesen Voraussetzungen können nun die gestellten Fragen beantwortet werden. Wichtig dabei ist, dass Amazon nur den Intent und die Slots an den Entwickler sendet, aber nicht den genauen Satz, den der Benutzer gesagt hat. Man erhält also keine Audio-Datei sondern wirklich nur JSON-Text.

Sagt der Benutzer z.B. „Alexa, frage k. i. t. Bibliothek wo das Medienzentrums ist“, schickt der

Abbildung 4:
Mögliche Werte
für den Typ
BIB_PLACE
(inklusive
Synonyme)

Slot Type	ID	Enter ID	Add synonym	+	Slot Value	Remove
BIB_PLACE	lsw	Enter ID	Add synonym	+	lsw, l s w	Remove
BIB_PLACE	lbs	Enter ID	Add synonym	+	l b s, l b s	Remove
BIB_PLACE	lehrbuchsammlung	Enter ID	Add synonym	+		Remove
BIB_PLACE	lst	Enter ID	Add synonym	+	l s t, l s t	Remove
BIB_PLACE	lsn	Enter ID	Add synonym	+	l s n, l s n	Remove
BIB_PLACE	scanner	Enter ID	Add synonym	+		Remove
BIB_PLACE	kopierer	Enter ID	Add synonym	+		Remove
BIB_PLACE	gruppenarbeitsräume	Enter ID	Add synonym	+		Remove

Das Anlegen der Intents und die Konfiguration erfolgt bequem über die Webtools.

Vor allem bei den Beispielsätzen muss man beachten, auf wie viele mögliche Arten eine Frage gestellt werden kann. „Wo ist der Scanner“, „Wo sind die Scanner“, „Wo finde ich Scanner“, „Wo gibt es scanner“ sind nur einige Beispiele.

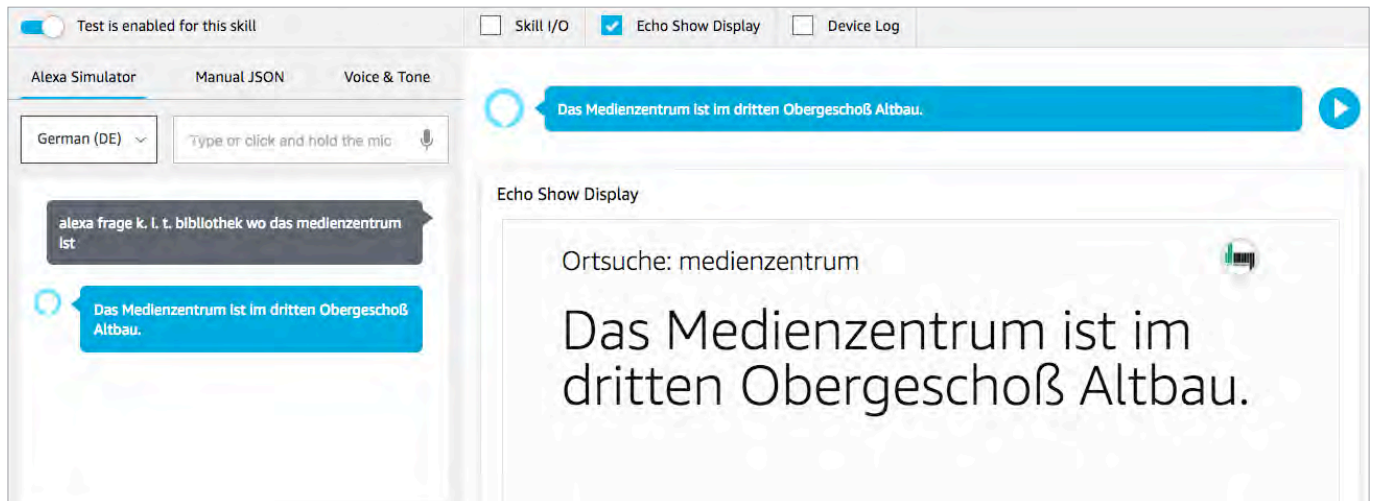
Die Variablen (sog. „Slots“) bestehen immer aus einem Typen, dem Werte zugewiesen werden. Abbildung 3 zeigt mögliche Werte für den Typ BIB_PLACE, der dem Slot {bibplace} zugewiesen ist.

Sind alle Intents erstellt und mit Beispielsätzen versehen, muss noch angegeben werden, wo sich der Code befindet, der auf den Skill antworten soll.

Dazu kann man entweder die von Amazon betriebene Cloudlösung AWS Lambda benutzen (mit Java-, Node.js-, C#- und Python Unterstützung) oder einen eigenen Webservice betreiben (in beliebiger Program-

Amazonserver an den eigenen Server eine Anfrage mit dem Inhalt {Intent: LocationInfoIntent, Slot{bibplace:medienzentrum}}. Der Inhalt wäre gleich, wenn der Nutzer: „Alexa, wo ist das Medienzentrums von k. i. t. bibliothek“ gesagt hätte.

Auf diese Frage erstellt der Server dann eine Text-Antwort „Das Medienzentrums ist im dritten Obergeschoß Altbau.“. Diese Antwort landet bei Amazon, wird dann zum Benutzer weitergeschickt und das Echo-Gerät liest dieselbe vor. Hat der Benutzer ein Gerät mit Display, kann man mit der Antwort auch einen Text mitschicken (dieser muss nicht unbedingt dem gesprochenen Text entsprechen). Dabei soll die Antwort nach dem voice first Prinzip sein: Das wichtigste ist die gesprochene Antwort, der Text ist nur eine Begleitung derselben. Aus diesem Grund darf der angezeigte Text auch keine Links zu anderen Webseiten erhalten, oder gar andere Apps öffnen.



Testen und Veröffentlichen eines Skills

Während der Entwicklung der Antworten für die Intents stellt sich natürlich die Frage wie diese getestet werden können.

Zum einen bietet Amazon selber in der **Developer-Konsole** ein Tool an, in das man entweder per angeschlossenen Mikrofon oder Texteingabe den Skill testen kann. Man schreibt z.B. „alexa frage k. i. t. bibliothek wo das medienzentrum ist“ und erhält die Antwort gesprochen mit passender Anzeige (falls eine mitgeschickt wird).

Hier lauert schon eine der ersten Fallen der Spracheingabe: Was eingetippt wird, ist nicht immer das, was man sagt. Hätte man im oberen Satz hinter das Wort Alexa ein Komma gesetzt, lautet die Antwort „Das weiß ich leider nicht“.

Um einen Test auf „richtigen“ Alexa Geräten durchzuführen, ist es möglich, bis zu 500 Benutzer (bzw. deren Amazon Account Email Adressen) anzugeben, die den Skill noch während der Entwicklung auf den eigenen Geräten testen können.

Vor allem während dieser Tests auf realen Geräten kamen einige Unzulänglichkeiten zu Tage. Eines der größten Probleme stellte nicht die Erkennung der Intents sondern das des eigentlichen Skill-Namens dar. Während der Satz „Alexa, wo kann ich lernen von k. i. t. Bibliothek“ relativ gut verstanden wurde, gab es mit „Alexa, frage k. i. t. bibliothek wo ich lernen kann“ in fast 50% der Fälle Probleme. Dies führte dann beim Testen zu so absurden (für Alexa aber syntaktisch korrekten) Sätzen wie „Alexa, sag k. i. t. bibliothek wo das medienzentrum ist“. Ähnliche Probleme mit der Einbindung sind uns auch von der Universitätsbibliothek Bamberg bekannt.

Leider erhält man hier oft keine genaue Rückmeldung von Amazon (Alexa antwortet nur mit einem „Das weiß ich leider nicht“) und man kann dann nur mutmaßen, was genau schief gegangen ist.

Ist die App genug getestet, kann man diese für den Review Prozess freigeben. Dazu muss man drei Beispielsätze angeben, auf die Alexa antwortet und (im Falle eines eigenen Webservers) mehrere Sicherheitsfeatures wie z.B. eine Signaturprüfung implementieren. Dafür gibt es aber bereits fertige Skripte [11]. Amazon überprüft den Skill innerhalb von fünf Werktagen und gibt detaillierte Rückmeldung, wenn das Review nicht geklappt hat. War das Review erfolg-

Abbildung 5: Beispielfrage nach dem Medienzentrum in der Developer Konsole

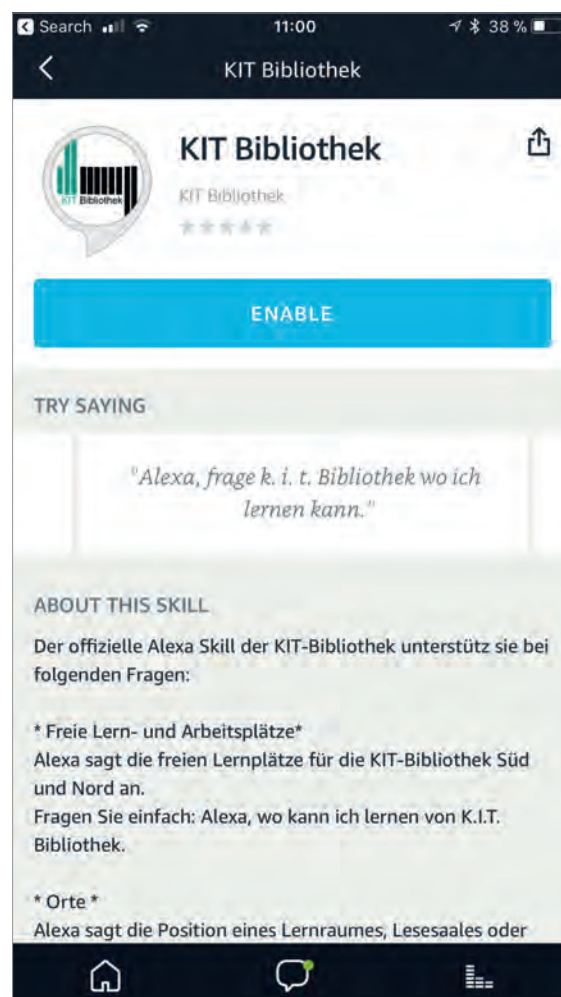


Abbildung 6: Der fertige Skill im Alexa-Store

reich, erfolgt die sofortige Verfügbarkeit für alle Alexa fähigen Geräte.

„Wer ist Alice?“ – Erfahrungen mit der Spracheingabe

Noch liegen wenig Erfahrungen mit der Spracheingabe außer den selbst gewonnenen vor, da der Skill noch nicht beworben wurde.

Wenn man mit dem Thema beginnt, hat man in der Regel eher zu große Erwartungen. Man denkt, dass die KI im Hintergrund Fragen unabhängig davon versteht, wie sie formuliert werden. Dem ist bei weitem (noch) nicht so. Einzelne Wörter werden ohne Anlernphase i.d.R. sehr gut erkannt. Nicht so gängige Begriffe und somit meistens auch alle Abkürzungen müssen wie oben beschrieben der KI beigebracht werden. Ein witziges Beispiel ist der Begriff des Lesesaals, der in Bibliotheken oft als „LS“ abgekürzt wird. Obwohl dem Skill „KIT-Bibliothek“ die Abkürzungen wie LSM (Medienzentrum), LST (Technik), LSW (Wirtschaft) „beigebracht“ wurden, funktionierte die Erkennung beim Testen oft nicht. Ein Blick in die ausführlichen Anfragedaten, die man intern in so einem Fall im JSON-Format erhält, führte dann zumindest zu einer Erklärung. LS klingt schnell ausgesprochen für einen Menschen aus dem anglo-amerikanischen Sprachraum wie der Vorname „Alice“ und auf die Frage wo sich „Alice“ befindet hatte unser Skill keine Antwort parat. Hier merkt man der KI von Alexa dann doch sehr ihre Herkunft an.

Fazit

Derzeit in Klärung befindet sich die Frage an die am KIT bestellte datenschutzbeauftragte Person, ob es sich bei Sprachsamples um personenbezogene Daten handelt. Das würde den Einsatz einer im Bereich einer Bibliothek aufgestellten Alexa-Box definitiv verbieten. Selbst dann, wenn man – und das ist mit einem Alexa-Business-Account für wenige Euro pro Monat technisch möglich – eine solche Box so konfiguriert, dass sie nur den eigenen Skill versteht.

Allerdings ist das Thema Sprachassistenten generell ein heißes Eisen und öffentliche Einrichtungen wie Bibliotheken tun vermutlich gut daran, hier kein Öl ins Feuer zu gießen. Denn zu Sprachassistenten gibt es meist nur zwei Positionen – die einen mögen es und haben keine Probleme damit, sich in jedes Zimmer ein solches Gerät zu stellen und die anderen fühlen sich schon bei dem Gedanken unwohl, dass sie beim Besuch von Freunden inzwischen nicht mehr wissen, wer sonst noch so mithört, unabhängig davon, ob die Gastgeber das Gerät deaktiviert haben oder nicht.

Bei der technischen Umsetzung wird man von Ama-

zon zwar mit vielen Tools und guter Dokumentation unterstützt, wieso Alexa plötzlich etwas anderes versteht, bleibt aber oft ihr Geheimnis. Für die Entwicklung ist dies natürlich hinderlich, da man nie weiß, ob der Fehler bei der KI oder beim Entwickler selbst liegt. Dabei ist das „Internet of Voice“ noch ein relativ neuer Trend und es fehlen entsprechende Literatur und Erfahrungen mit Sprachinterfaces. Ebenso müssen wir noch unsere Sprache umständlich an die Syntax der Systeme anpassen, vielleicht wird das in Zukunft schon anders herum sein. Wir sind schon auf die nächste Generation der Geräte gespannt. Was die weitere Entwicklung des eigenen Skills betrifft, so werden an der KIT-Bibliothek weiterhin häufig gestellte Fragen von Nutzern gesammelt. Fragen, die einfach zu verstehen und hoffentlich auch einfach zu beantworten sind, sollen dann vom Skill verarbeitet werden können. Die Suche nach der „Killer“-Anwendung für einen Skill im Bibliotheksumfeld dürfte weiterhin spannend sein. ■

Links & Literatur

- [1] The smart audio report https://bit.ly/kit_skill_1
- [2] Amazon Alexa Android Apps bei Google Play https://bit.ly/kit_skill_2
- [3] iTunes – Amazon Music App https://bit.ly/kit_skill_3
- [4] mVoice G2: World's First „Watch“ with Voice https://bit.ly/kit_skill_4
- [5] Alexa in the car: Which vehicles will get Alexa this year? https://bit.ly/kit_skill_5
- [6] NissanConnect | Amazon Alexa https://bit.ly/kit_skill_6
- [7] Alexa in the Car: Ford, Amazon to Provide Access https://bit.ly/kit_skill_7
- [8] The First Intelligent Heated Jacket https://bit.ly/kit_skill_8
- [9] Amazon Alexa Skill Suche https://bit.ly/kit_skill_9
- [10] GitHub – Amazon Alexa PHP Library https://bit.ly/kit_skill_10
- [11] GitHub – Validator for incoming Amazon Alexa requests https://bit.ly/kit_skill_11

Dipl.-Inform. Uwe Dierolf

Leiter der IT-Abteilung

M. Sc. Michael Skarupianski

Mitarbeiter der IT-Abteilung

Anfragen bitte an

alexa@bibliothek.kit.edu

Karlsruhe Institut für Technologie (KIT)

KIT-Bibliothek

Straße am Forum 2

76131 Karlsruhe