

## GRAFISCHE AUSGABE MIT ALEXA SKILL CARDS

# Bildschirmpräsenz

Alexa ist auf diversen Bildschirmen präsent – vom Smartphone bis zum Fernseher.

Echo-Endgeräte - abgesehen vom Echo Show und Echo Spot -, die über einen eigenen Touch-Screen eingerichtet werden können, erhalten ihre Verknüpfung mit einem Amazon-Konto und dem WLAN über die Alexa-App. Aber auch nach der Einrichtung bleibt die Alexa-App eine wichtige Komponente in der Nutzung.

## Cards

Eine sogenannte Card kann Informationen zur Antwort von Alexa anzeigen. Darüber hinaus enthält eine Card ein kleines Menü mit der Bezeichnung *mehr*, mit dem sich noch Details einblenden lassen ([Bild 1](#)). So wird aufgelistet, was Alexa verstanden hat und man kann sich die Anfrage auch noch einmal vorspielen lassen. Sollte beides voneinander abweichen, kann man ein Feedback geben und damit das Training von Alexa verbessern.

Welche Informationen ein Skill mit der Antwort zurückgeben kann, hängt vom Typ der Card ab. Diese Typen sind grob in zwei Kategorien unterteilt: Die normalen Cards, welche passive Informationen enthalten, und die interaktiven Cards, welche in ausgewählten Anwendungsfällen Informationen vom User erfragen können. Die Anwendungsfälle werden für letztere von Amazon stark reglementiert, da der Fokus weiterhin auf der Sprach-Interaktion liegen soll.

Ein Use-Case der rein per Sprache nicht ordentlich funktioniert ist zum Beispiel das Account-Linking (siehe Ausgabe 6/18, ab Seite 126). In der Response wird neben der normalen *outputSpeech* noch ein *Card*-Objekt zurück geliefert, welches nur den Typ *LinkAccount* hat ([Listing 1](#)).

Während der Echo den individuellen Text aus dem *outputSpeech* wiedergibt, enthält die erzeugte Card den Namen des Skills, einen Text der von Amazon vorgegeben wird sowie - als Besonderheit - einen Link, der auf die Skill-eigene Seite zum Account-Linking verweist ([Bild 2](#)).

Weitere interaktive Cards dienen dem Anfragen von Berechtigungen

([Bild 3](#)). Der Link führt dann innerhalb der App zu der Berechtigungsverwaltung ([Bild 4](#)), die als interaktives Element einen Schieberegler benutzt. Zwar gab es bereits Einwände, dass der Schieberegler direkt auf der Card angezeigt werden könnte, jedoch macht dieser Zwischenschritt durchaus Sinn, da auch die Übersichtsseite zum Aktivieren eines Skills nach der Aktivierung auf die Berechtigungsverwaltung verlinken und so alle Cards maximal einen Link als interaktives Element enthalten.

## Simple Card

Leider ist es bei normalen Cards nicht möglich, interaktive Links anzuseigen, über die auf einfache Weise auf weitere Informationen verlinkt werden kann. Da der Text auf einer Card

innerhalb der Alexa-App auch nicht selektiert und kopiert werden kann, müsste eine reine Text-URL vom Nutzer abgetippt werden. Diese wichtige Einschränkung zeigt aber auch gleich, was eine Card im einfachsten Fall enthält: reinen Text.

Im einfachsten Fall hat die Card auch den entsprechenden Typ *Simple* und kann einen Titel und/oder einen Text enthalten - muss sie aber nicht:

```
"card":  
{  
  "type": "Simple",  
  "title": "Dies ist der Titel",  
  "content": "Und hier ist einfach  
  eine Menge Text und der Link  
  http://test.me bleibt auch Text."  
}
```

Dabei ist zu beachten, dass eine leere Card und eine Response ohne Card zu unterschiedlichen Ergebnissen - insbesondere auf einem Echo Show beziehungsweise Spot - führen.

Eine Response ohne Card wird in der App nicht aufgeführt, aber auf einem Echo Show und Spot erfolgt eine Display-Anzeige, die den Namen des Skills (in diesem Beispiel *Card Builder*) enthält. Der Echo Show führt unter dem Namen auch noch ein Beispiel an, wie das Skill verlassen werden kann.



BILD: Amazon

Was hältst du von der Datenschutz Grundverordnung?

Der Schutz deiner Daten ist uns sehr wichtig. Bitte schaue dir hierzu die Amazon Datenschutzerklärung an, sowie die häufig gestellten Fragen zu Alexa und Alexa-Geräten. Ich habe dir einen Link in die Alexa App geschickt.

[AMAZON.DE-DATENSCHUTZERKLÄRUNG](#)

Sprachfeedback

▶ Alexa hat Folgendes gehört: „alexa was hältst du von datenschutz grundverordnung“

Hat Alexa das gemacht, was Sie wollten?

Ja      Nein

Karte entfernen      Mehr dazu

Weniger ▾

Card mit eingeblendeten Detail-Informationen (Bild 1)

SprachBox - Kontoeinrichtung  
SprachBox

Um diesen Skill optimal zu nutzen, verknüpfen Sie bitte Ihr Konto.

KONTO VERKNÜPFEN

Mehr ▾

AcountLinking-Card mit Link zur Anmeldeseite (Bild 2)

Berechtigung angefordert

Text Nachricht

Benachrichtigungen  
Skill erlauben, eine Kurznachricht an Ihre Warteschlange zu senden.

BERECHTIGUNGEN VERWALTEN

Mehr ▾

Berechtigungs-Card mit Link zur Rechteverwaltung des Skills (Bild 3)

Text Nachricht  
Emil Thies  
★★★★★

Skill-Berechtigungen

Text Nachricht erfordert Folgendes:

Alexa-Benachrichtigungen

Erlauben Sie diesem Skill, Ihnen Benachrichtigungen zu senden, die Sie abrufen können, indem Sie "Alexa, welche Benachrichtigungen habe ich erhalten?" fragen.

Einige Geräte unterstützen möglicherweise nicht die Benachrichtigungstöne oder optischen Benachrichtigungssignale.

EINSTELLUNGEN SPEICHERN

Skill aktivieren.

Rechteverwaltung eines Skills mit Notification (Bild 4)

Eine Response mit einer leeren Card führt auf allen Displays zu einer vollständig leeren beziehungsweise schwarzen Anzeige. In der App hingegen wird eine Card mit dem Skill-Namen angezeigt. Da die Cards in der App erhalten bleiben, ist der Skill-Name dort wichtig, um später noch nachvollziehen zu können, welches Skill diese Rückmeldung gegeben hat. Auf den Displays hingegen wird der Skill-Name nicht angezeigt, da die Card dort nach einigen Sekunden wieder ausgeblendet wird. Eine Ausnahme bildet hier das Fire TV, welches den Skill-Namen unter dem Titel anzeigt, wenn dieser in der Card enthalten ist.

Der gesamte Text aller Attribute einer Card ist prinzipiell auf 8000 Zeichen begrenzt. Dabei ist zu beachten, dass die verschiedenen Endgeräte diesen unterschiedlich verarbei-

ten. Selbst die Simulatoren, welche Amazon bereitstellt, bilden das Verhalten der echten Endgeräte nicht eins zu eins ab (Tabelle1).

Ein abgeschnittener Titel endet mit drei Punkten, mit denen weitere Zeichen angedeutet werden. Dies ist auch der Grund, weswegen die Zeichenangabe für den Titel nur als circa Angabe vorgenommen wurde, da der Titel für die ►

▶ Tabelle 1: Texte auf verschiedenen Endgeräten

Gerät	Titel	Content
Alexa App (iPhone X)	bis zu 3 Zeilen; circa 70 Zeichen	Die Card wird entsprechend groß abgebildet, um allen Text aufzunehmen; 42 Zeichen pro Zeile
rechteckiger Simulator	1 Zeile; circa 44 Zeichen	6 Zeilen sichtbar; nicht scrollbar; 22 Zeichen pro Zeile
Runder Simulator	bis zu 2 Zeilen; circa 38 Zeichen	initial circa 5 Zeile; scrollbar; circa 7 Zeilen sichtbar; 13 Zeichen pro Zeile
Echo Show (1. Gen)	1 Zeile; circa 48 Zeichen	Initial 6 Zeilen; scrollbar; 22 Zeichen pro Zeile
Echo Spot	bis zu 2 Zeilen; circa 26 Zeichen	initial circa 5 Zeilen; scrollbar circa 7; 13 Zeichen pro Zeile
Fire TV (bis Sep. 2018)	bis zu 2 Zeilen; circa 78 Zeichen	5 Zeilen; scrollbar, aber Titel und Skill Name bleiben stehen 29 Zeichen pro Zeile
Fire TV (ab Okt. 2018)	1 Zeile; circa 55 Zeichen	8 Zeilen; scrollbar, aber Titel bleibt stehen; 29 Zeichen pro Zeile

### ▶ Listing 1: Skill Response

```
{
  "version": "1.0",
  "response": {
    "outputSpeech": {
      "type": "PlainText",
      "text": "Um die Sprachbox nutzen zu können, muss das Skill mit deinem SprachBox Daten verbunden werden."
    },
    "card": {
      "type": "LinkAccount"
    },
    "shouldEndSession": true
  },
  "sessionAttributes": {}
}
```

drei Punkte gekürzt wird. So wird auf einem Echo Show ein Titel mit 49 Zeichen vollständig angezeigt, während ein Titel mit 50 Zeichen nur mit seinen ersten 48 Zeichen wiedergegeben wird. Zeilenumbrüche werden unabhängig vom Wort oder Trennungsregeln durchgeführt. Vorher wird allerdings versucht diese Situation zu umgehen, indem das Leerzeichen vor einem betreffenden Wort in einen Zeilenumbruch umgewandelt wird. Diese Regel wird auch angewendet, wenn der Wortumbruch nicht verhindert werden kann und führt so unter Umständen zu Platzverschwendungen (**Bild 5**). Um durch die unterschiedliche Zeichenlänge etwas Kontrolle über die Zeilenumbrüche behalten zu können, können diese mit der Escape-Sequenz \n im String explizit angegeben werden.

In den Simulatoren gibt es noch die Besonderheit, dass ein einzelnes sehr langes Wort das Scrollen deaktiviert und im rechteckigen Simulator auch nur eine Zeile genutzt wird.

Dieses Bundesanwaltsschaftshauptgebäude ist ein Bundesanwaltsschaftshauptgebäude wie jedes andere Bundesanwaltsschaftshauptgebäude

**Zeilenumbruch** führt zu Platzverschwendungen (**Bild 5**)

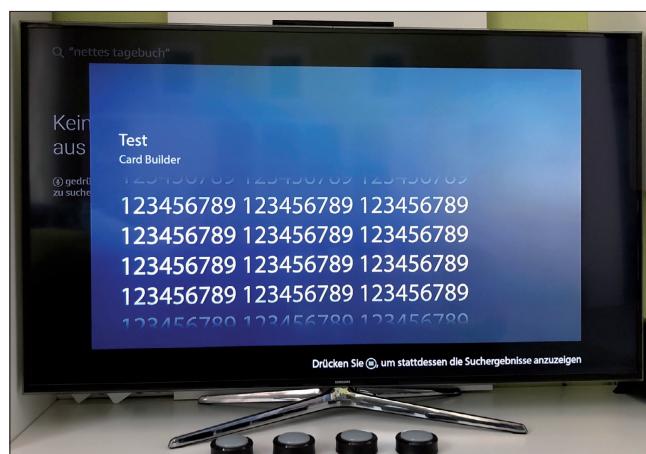
Auf dem Fire TV trat ein ähnlicher Effekt auf, (\*) wenn kein Titel genutzt wurde (das heißt, wenn das Titel-Attribut fehlt beziehungsweise wenn es nur einen leeren String enthält), dann wurde ein bisher scrollbarer Text (**Bild 6**) in der sechsten Zeile mit drei Punkten abgeschnitten und war nicht mehr scrollbar (**Bild 7**). Am 30. September lieferte Amazon für das Fire TV ein Update aus, wodurch eine Card nun nicht mehr als blaues Overlay 10 Prozent kleiner als der Bildschirm eingeblendet wird, sondern stattdessen ein dunkles Muster (wie auf dem Echo Show), welches den ganzen Bildschirm füllt.

### Standard Card

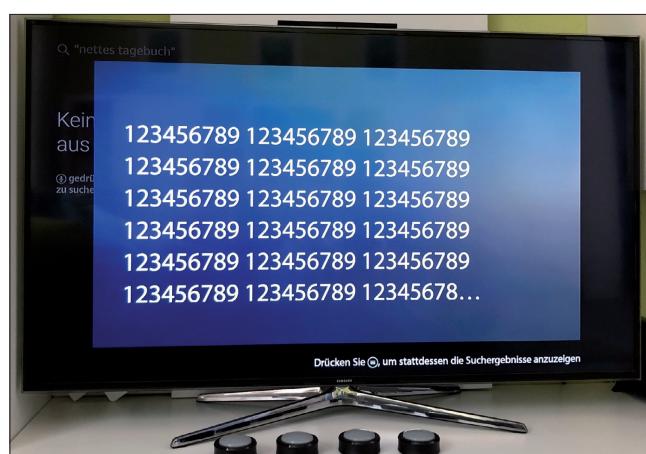
Neben reinem Text kann eine Card auch ein Bild enthalten. Hierzu muss das Attribut *type* der Card auf *Standard* geändert werden. Zwar kann das Attribut *title* weiter für den Titel genutzt werden, der zusätzliche Text ist nun jedoch im Attribut *text* zurück zu gegeben, anstatt im Attribut *content*. Die Nutzung eines Bildes ist optional. Auch wenn keines verwendet wird, ändert sich die Darstellung der Card, da die Schriftgröße angepasst wird und der Platz für das Bild reserviert bleibt. Positiver Gesamteffekt ist, dass in Summe mehr Text auf einmal angezeigt werden kann. Negativ hingegen ist die nicht dokumentierte Einschränkung, dass ein zu langer Text auf dem Fire TV und Echo Show bei diesem Card-Type nicht mehr scrollbar ist (**Tabelle 2**).

### Bilder – Vorbereitende Maßnahmen

Bevor ein Bild zu einer Card hinzugefügt werden kann, muss der Server meist noch vorbereitet werden. Zum einen muss das Bild über eine HTTPS-URL erreichbar sein, deren SSL-Zertifikat von einer durch Amazon anerkannten Certificate Authority signiert ist. Wenn das Bild auf dem gleichen Server liegt wie das PHP-Skript, ist diese Voraussetzung bereits erfüllt, da die URL des Scripts als Skill-Endpunkt akzeptiert



Card mit Titel auf Fire TV, Text ist scrollbar (**Bild 6**)



Card ohne Titel auf Fire TV: Text wird abgeschnitten und ist nicht mehr scrollbar (**Bild 7**)

Tabelle 2: Darstellungsvariante Standard Card

Gerät	Titel	Content
Alexa App (iPhone X)	bis zu 3 Zeilen; circa 70 Zeichen	Die Card wird entsprechend groß abgebildet, um allen Text aufzunehmen; 42 Zeichen pro Zeile (keine Platzreservierung)
rechteckiger Simulator	1 Zeile (ohne ...); 46 Zeichen	10 Zeilen; nicht scrollbar; 24 Zeichen pro Zeile
runder Simulator	bis zu 2 Zeilen circa 38 Zeichen	circa 4 Zeilen; nicht scrollbar; circa 13 Zeichen pro Zeile; (keine Platzreservierung)
Echo Show (1. Gen)	1 Zeile; circa 48 Zeichen	11 Zeilen; nicht scrollbar 26 Zeichen pro Zeile
Echo Spot	bis zu 2 Zeilen; circa 26 Zeichen	initial circa 5 Zeilen; scrollbar circa 7; 13 Zeichen pro Zeile (keine Platzreservierung)
Fire TV (ab Okt. 2018)	1 Zeile; circa 55 Zeichen	18 Zeilen; nicht scrollbar; 51 Zeichen pro Zeile

Listing 2: .htaccess limitiert CORS

```
1: SetEnvIf Origin
   "^.+https?:\/\/ask-ifr-download.s3.amazonaws.com$"
   origin_is=$0
2: Header set Access-Control-Allow-Origin
   %{origin_is}e env=origin_is
```

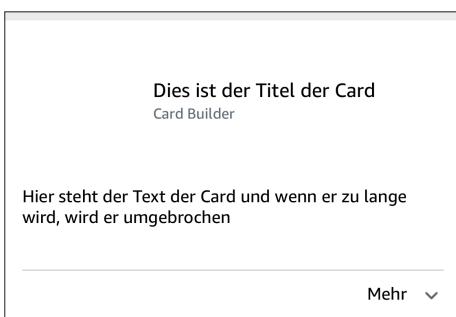
### Listing 3: PHP Script limitiert CORS

```
<?
$http_origin = $_SERVER[HTTP_ORIGIN];
if ($http_origin == "http://ask-ifr-download.s3.
amazonaws.com" || $http_origin ==
"https://ask-ifr-download.s3.amazonaws.com")
{
header("Access-Control-Allow-Origin: $http_origin");
}
?>
```

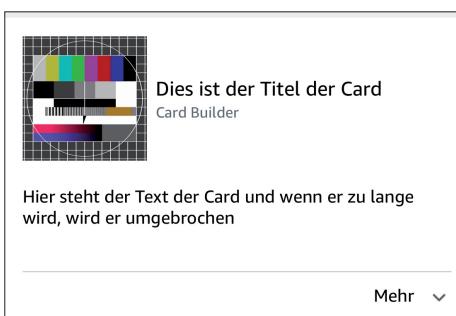
wird. Zusätzlich muss der Server im http-Header des Bildes aber auch Cross-Origin Resource Sharing (CORS) erlauben. Für eine CORS Freigabe reicht es, auf einem Apache-Server eine `.htaccess`-Datei mit diesem Inhalt zu hinterlegen:

```
Header set Access-Control-Allow-Origin "*"
```

Hier passieren aber gleich zwei Dinge: Erstens wird CORS für alle abfragenden URLs akzeptiert, wobei wir hier eigentlich



**Card mit Bild-Platzhalter** in der App, wenn die Bild-URL fehlerhaft ist  
**(Bild 8)**



**Card mit Bild** in  
der App bei  
maximaler  
Ausnutzung  
des Bildberei-  
ches (Bild 9)



**Card mit Bild auf Echo Show:** Text ist scrollbar und das Bild rechts davon fixiert (Bild 10)



**Card mit Bild auf Fire TV:** Text ist nicht scrollbar und Bild rechts davon (Bild 11)

nur zwei erlauben müssten ([Listing 2](#)), zum anderen gilt die Freigabe für alle Dateien. Wenn die Bilder nicht bereits als Datei vorliegen, sondern ebenfalls von einem Skript erstellt werden, so kann der notwendige Header vom Skript auch erzeugt werden ([Listing 3](#)).

## Aussehen von Cards mit Bild

Wenn die App das Bild nicht laden kann, entweder weil es nicht per CORS freigegeben ist, oder unter der URL nicht existiert, dann wird die Card zwar erzeugt, die für das Bild vorgesehen Fläche links neben Titel und Skill-name bleibt jedoch weiß (**Bild 8**). Der Text nutzt dann unterhalb des Bildes wieder die gesamte Breite. Die für ein Bild vorgesehene Fläche ist quadratisch und immer gleich groß (**Bild 9**). Sollte das Bild nicht quadratisch sein, so wird die kleinere Kante mittig ausgerichtet und die nicht genutzte Fläche bleibt unverändert weiß.

Auf dem Echo Show sieht die Sache so aus: Das Bild befindet sich rechts vom Text. Die Breite des Bildbereiches wird vom Textbereich abgezogen. Der Text bleibt innerhalb seines Bereiches scrollbar, während das Bild in seinem Bereich fixiert ist ([Bild 10](#)).

Ein ähnliches Erscheinungsbild nimmt die Standard Card auch auf dem FireTV an. Allerdings ist hier ein längerer Text auf dem TV-Bildschirm nicht mehr scrollbar (**Bild 11**). Auf einem Echo Spot ist zwar der Text scrollbar, aber das Bild wird nicht verwendet (**Bild 12**). ▶

Die URL zum Bild wird nicht direkt in der Card hinzugefügt, sondern innerhalb eines *image*-Objektes. Dies hat den Vorteil, dass für ein Bild mehrere URLs für unterschiedliche Auflösungen bereitgestellt werden können. Zurzeit unterscheidet Amazon nur zwischen kleiner (720x480) und großer (1200x800) Auflösung:

```
"card": {
  "type": "Standard",
  "title": "Dies ist der Titel",
  "text": "Und hier ist einfach eine Menge Text und
  http://test.me bleibt
  auch Text.",
  "image": {
    "smallImageUrl":
      "https://woachimmer.
      de/klein.png",
    "largeImageUrl":
      "https://oauchimmer.
      de/gross.png"
  }
}
```

An den Empfehlungen ist zu erkennen, dass Amazon



**Card mit Bild auf Echo Spot:**  
Text ist scrollbar und Bild wird nicht verwendet ([Bild 12](#))

### ➤ Update zu SSML (Artikel in 10/18)

Amazon erlaubt es nun auch zwischen den 27 verschiedenen Polly-Stimmen direkt in SSML zu wechseln: <https://developer.amazon.com/de/docs/custom-skills/speech-synthesis-markup-language-ssml-reference.html#voice>

Des Weiteren kann auch die Sprache festgelegt werden, um zum Beispiel englische Begriffe in deutschem Text richtig auszusprechen: <https://developer.amazon.com/de/docs/custom-skills/speech-synthesis-markup-language-ssml-reference.html#lang>

eigentlich von rechteckigen Bildern ausgeht. Amazon selbst schreibt nur, dass jedes Endgerät eine der beiden URLs priorisiert. Sollte diejenige URL nicht vorhanden sein, wird einfach die andere verwendet. Dabei werden zu große Bilder skaliert, was dann zu unnötig langen Ladezeiten aufgrund größerer Dateien führt. Zu kleine Bilder werden ebenfalls skaliert, bilden dabei jedoch in den meisten Fällen sichtbare Qualität ein.

Wenig überraschend ist, dass die Alexa-App das kleine und der Fire TV das große Bild verwendet. Beim Echo Show (1.Gen) wären beide Bilder potentielle Kandidaten, aber beim Ausprobieren stellt sich heraus, dass dieser das große Bild verwendet.

### Fazit

Mit einfachen Mitteln kann Alexa eine Textausgabe - begleitet bei Bedarf auch von einem Bild - auf den ihr zugänglichen Bildschirmen ausgeben. In der Alexa-App bleiben diese sogar persistent erhalten. Als Entwickler muss man sich allerdings bewusst sein, dass die zahlreichen unterschiedlichen Bildschirmgrößen auch ein unterschiedliches Verhalten nach sich ziehen.

Unterschiedlich viele Zeichen pro Zeile führen zu automatischen Zeilenumbrüchen an unterschiedlichen Stellen. Die unterschiedliche Anzahl an Zeilen führt zu unterschiedlich viel Text, der ohne Scrollen (wenn überhaupt möglich) sofort sichtbar ist.

Mit der Einführung des Echo Show mit integriertem Touch-Screen, hat Amazon die Möglichkeiten für Skills um die Ausgabe zusätzlicher Informationen und sogar um die Möglichkeit zu Eingabe, erweitert. Hierfür wurde das sogenannte Display-Interface eingeführt, welches ich in meinem nächsten Artikel vorstellen werde. ■



**Emil Thies**

ist in der IT Abteilung der Deutschen Telekom angestellt und probiert auch nebenberuflich neue Technologien aus, indem er eigene Apps, Skills und Actions programmiert und veröffentlicht.