



Technische Universität Berlin



DOCUMENT BUILD DATE: February 14, 2018

DOCUMENT STATUS: Beta

Development and Evaluation of a Service Bot in the e-Government Sector

Bachelor Thesis

am Fachgebiet Agententechnologien in betrieblichen Anwendungen und der
Telekommunikation (AOT)

Prof. Dr.-Ing. habil. Şahin Albayrak
Fakultät IV Elektrotechnik und Informatik
Technische Universität Berlin

vorgelegt von
Mohamed Megahed

Betreuer: **Dr. Andreas Lommatzsch**
Gutachter: Prof. Dr.-Ing. habil. Şahin Albayrak
Prof. Dr. Odej Kao

Matrikelnummer: 342655

Abstract

Though not a recent phenomenon, chatbots and voice assistants are increasingly gaining unprecedented attention as a successor for mobile and web apps. While still emerging with no defined standards or set protocols, with a hype on the rise, tensions between industry giants with products like Amazon's Alexa, Apple's Siri, the Google Assistant or IBM's Watson unveil new examples in favour of providing an enriched user experience on consumer and business level. The surrounding ecosystem also plays a major role in widening the platforms available while exploring new horizons with alternative approaches and business models. Today voice assistance are already present around indoor spaces, in the car or on the go but are still a new terrain to discover and great potential to unleash.

One such use cases involves the public sector. In this work, we are going to explore Amazon's Alexa and respective platforms to develop a voice assistant for the local city council extending the current chatbot's functionality available on <http://service.berlin.de>. We will touch on the technical challenges and possibilities in implementing a system for eGovernment inquiries and touch on its usability as well as effectiveness in replacing a traditional lookup service. We will then examine the goals we define for our use case to what we were able to achieve with the available APIs and SDKs. With respect to those, we will also report on the limitations developers could face in the process.

Finally, we aim at analysing the current state of voice assistants and service bots in the market and the future of this trend from a technical and a social point of view.

Zusammenfassung

👉 **NOTE:** translate to German to English or vice-versa.

[possibility to make inline notes] **[CITATION]**when a citation is missing

✍️ **To-DO:** todos

[TO CITE]

Contents

List of Figures	vi
List of Tables	vii
1 Introduction	1
1.1 Motivation	3
1.2 Approach and Goals	5
1.3 Structure of the Thesis	5
2 Background	7
2.1 Related Work	7
2.1.1 Information bots	8
2.1.2 more on AI in bots (optional)	8
2.2 State of the Art	8
2.2.1 Node.js	8
2.2.2 Alexa Skills, Amazon Voice Services	9
2.2.3 comparison to Amazon Lex / Amazon Polly	9
2.3 D115	9
2.4 Frameworks and Data Structures (change title)	9
2.4.1 Intents and Slots	9
2.5 currently deployed bot	11
2.6 Implementation Possibilities	11
3 Implementation as Alexa Skill	12
3.1 All about Alexa	12
3.2 Difference Between Lex and Alexa Skills	12
3.3 APIs and SDKs	13
3.4 challenges	13
4 Implementation as Facebook Messenger Bot / Google Action	14

5	Evaluation	15
5.1	Results	16
5.2	Discussions	16
6	Conclusion and Future Work	17
6.1	Summary	17
6.2	Conclusion	17
6.3	Future Work	17
	Bibliography	18
	Appendices	19
	Appendix A: Abbreviations	20
	Appendix B: Glossary	21

List of Figures

2.1	Dienstleistungen.json - Primary Nodes	10
2.2	Dienstleistungen.json - secondary Nodes	10

List of Tables

Chapter 1

Introduction

With over a third of the world's population projected to own a smartphone in 2018 [10] and a substantial fraction thereof using smarthome devices on a daily basis, AI's role has become more interesting than ever for productivity and entertainment. Many technologies we take for granted today, such as dictation and word prediction, recommender systems or other number-based analytics depend on Machine Learning and Natural Language Processing techniques that were only made possible thanks to the high computational power shipped in most devices gradually overtaking the consumer market. This transition also facilitated the introduction of a new form of interaction through conversation with the hardware, paving the way to an aspiration we long strove for [7]. Conversational bots were already prevalent since the 80s in the form of Question/Answer systems based on query programming languages like PROLOG and SQL. ELIZA, considered as the world's first chatbot and though quite superficial as an NLP-based programme for psychoanalysis, already at its early stages demonstrated how humans can become emotionally attached to machines, transcending over the anomaly of making conversation not with a human [11]. Today, combining ML with the retrieval-based approach allows a more advanced interaction with the system and yields smarter and more personalized chatbots. Consequently, it is no longer a surprise that chatbots acquire social skills to make Xiaoice, the empathetic bot from China, possibly a new kind of friend made of silicon.

So far, Voice assistants represent an additional layer of abstraction from software beyond the graphical user interface (GUI) and are hence the closest we have come to human communication. They take away another barrier between the user and the application as they generally do not require profound computer literacy and are modelled to rely on our inherent ways of expression. As they simulate the human aspect and immitate its behaviour for instance with small-talk abilities [9], voice assistants are regarded as a convenience for daily tasks and are on their way to becoming a de-facto replacement to sophisticated actions we perform on the screen. In fact, since voice is a

primal stimulus in the human brain [6], using it becomes more instinctive, making us in turn process our ideas starting with an "inner voice" that translates easiest to words when we speak it before "funneling" it into actions [2]. Arguably, on a market scale, this gives voice assistants today a competitive innovation advantage (CIA) for they are hence more accessible to further demographic groups with a growing wider acceptance.

Further, the speech-to-text/text-to-speech domain has become more powerful with steadily increasing processing power, an effect of Moore's law we only get to understand lately in addition to the gradual lessening of the dominance of Chomskyan theories of linguistics (e.g. transformational grammar) [3]. With an integration in most modern operating systems, speaking to a device has become no longer a absurd novelty. Inadvertently, Google Now, the voice assistant built into the OS currently with the highest usage shares among computer and smartphones [4], supports understanding of multiple languages in the same sentence for multilingual interaction. As we constantly challenge our expectations towards technology, a not too far-fetched imagination makes us question the ability of a machine to react to everything we say. Although we are still far from this step, at least for the consumer level, we dedicate a lot of effort to make it happen with examples like IBM's Watson. And as long as we still do not exactly understand human intelligence yet, it is hard to fathom AI as a holistic field. As such, it is therefore more realistic to consider the current works in the field as *intelligence amplification* [2] empowering human take better decisions beyond their normal brain processing power.

To-do: Why Bots

- mention how these are centralized at alexa somewhere **SKILLS**
- **Related work:** what are classic use cases for their use with prominent examples? Booking tickets (KLM bot)
- fun bots and more
- unfortunately forums and FAQ pages are not as effective as talking to a human.
- then again, as a customer, if I want assistance, I want the customer to tell me related information that he/she might not know, e.g. model number etc.

biased pros/cons:

- it would speak as an advantage for bots if they can determine these things automatically z.B.
- besides, I could be a bit more sure in customer support scenario that a bot won't trick me
- as a novice I am usually not sure if the help article / Kbase I am reading is the right one
- and forums have mostly Schrott anyway.
- what bots already achieved is at least not to give wrong answers.
- they could sometimes say idk, which is annoying, but at least it doesn't confuse the user.
- answer suggestions functionality - next step is to get around the user's frustration by making the bot at least more human.

To-do: funnel towards Motivation

- ...modularization and Einteilung of the paper

1.1 Motivation

Berlin.de is an online one-stop-shop for appx. 3,7 million residents [1] with [**hundreds/thousands**] of visitors daily for information lookup, appointment bookings and even access to local news. As part of a federal modernization procedure with the help of the ministry of interior, D115 was launched in 2009 [5] as a phone service to help residents find relevant information about a public service or municipality, something that can be tricky if a person has no overview of the local government structure and still not always easy even with the help of search engines nowadays. [**leading sentence**]. Meanwhile, statistics [8] show that time spent on messaging apps already surpassed average uptime on social media, which indicates how the former is more desirable as a communication format on mobile platforms. It could therefore be worth exploring, how to offer D115 services in a fashion that takes advantage of conversational abilities beyond its human personnel.

For now, although local authorities rely heavily on their websites to communicate information to the public, the challenge is mainly finding the right service. In a metropolis with a high influx of expatriates, it is also very likely that certain services are frequently pursued, meaning that helping find the right public service or authority is a repetitive task. In this context, thinking of a chatbot as a public service could have several advantages, like offloading some traffic from the phone service, getting over the language barrier in the case of non-german speakers or expatriates or simply helping customers formulate the right wording for a query in a more intuitive way than using a search box.

To-DO: 2 ¶

- **Chatbot vs. human:** Analyze differences between bot and human response
 - disadvantage: a bot wants a sentence broken down in small pieces to avoid errors in lengthy interpretation
- **Why can't robots understand us:** language ambiguities - the need to understand context
 - Syntactical: Homonymy
 - Semantic: Metaphors, sarcasm, and puns
 - dialects: enunciations
 - underlying grammar
 - underlying sentiment
- NLP Progress:** How does it help in enriching the bot experience
 - neural networks: help understanding language patterns and get better over time
 - thought vectors: helps connect different words with related meanings
- **wrap-up:** can bots replace services offered by humans? – mention transition from facets (Altavista) to metasearches to all-in-one (Google).
 - chatbots as enablers in customer service industry
 - conclusion: Although not impossible, it is a bit too far-fetched at this stage.

To-DO: Topology of Bots

- use cases and purpose categories (leisure, productivity) - quick survey of respective 'AppStores'
- platforms
- physical locations (home, office, car, phone, in a business)

To-DO: Aufgabenstellung:

- 1- es sollen die Stärken und Schwächen eines solchen System zu analysieren.
- 2- Es sollte zunächst eine Dienstleistung aus dem Berliner Service-Katalog mit dem Chatbot beauftragt werden können.
- 3- Nuancen beachten (e.g. 10243 / FHain)
- 4- Smalltalk Fähigkeiten

1.2 Approach and Goals

- making the bot become something beyond a Q&A:
 - Alexa Documentation
 - retaining sessions (explain requests/responses - GET/POST)
 - fulfilling intents
 - nested handlers
-
- for facebook: implementing the three-answer suggestions
-
- internationalization / customization based on Locale - why is it important?
 - many international users prefer a chatbot than a phone since the bot will communicate more accurately, will not have language probs if it understands the foreign lang etc.
 - what are other approaches to localization? refer to IRS lecture notes
 - use of translators, Stammsprache, etc., detecting the language and say it does not support it.
-
- Alexa Skill will work in germany in english and german -> add english after german
- AL: Anschließend soll das Ziel der Arbeit formuliert werden: Entwicklung und Evaluation eines Prototypen für den Anwendungsfall.

1.3 Structure of the Thesis

This thesis endeavours to shed light on the following: In Chapter 2, we discuss related work as currently available. We first show a few use cases for chat bots and voice assistants and their implementation. For that we consider [the City of Vienna chat bot service "WienBot" for German and Singapore/ LA/... for English] and compare them to the current application available on service.berlin.de. We then introduce the Alexa platform and the idea behind Skills in chapter 3 and discuss how it can overcome a few struggles with respect to its counterparts [like redundant boilerplate code, smalltalk, but also limitations]. We also present the frameworks we use in our own implementation, the prerequisites and the artefacts provided initially. Chapter 3 and 4 represent a detailed analysis of our implementation. In particular, [SOME MORE SENTENCES. In Chapter 4, our solution is presented. This solution covers ... (SOME MORE SENTENCES).] Chapter 5 evaluates our implementation with respect to our target in the defined [4-5] use cases. [SOME MORE SENTENCES] In Chapter 6, we conclude with how the implementation of this e-Government solution

is attainable through Alexa and where future works can be directed.

Chapter 2

Background

2.1 Related Work

To-DO:

by platform:

- API.ai
- Facebook Messenger Chatbots
- wit.ai
- motion.ai

by purpose

Information bots - mention available service types (information system as a "webpage/database")

- vs an interactive bot that gives you customized information on demand hier soll der D115 Anwendungsfall "Beauskunftung" kurz erläutert werden

social bots

- with advantages / disadvantages
- fake news / online reviews

more on AI in bots (optional) - use of ML

Handyversicherungsbeispiel

- from business perspective, the bot is aiming to sell more policies,
- the bot tries to determine if there is a nuance in the user's answer (machine acting as a judge!) - e.g. "how did the phone fall off" - MKTG - Aufwand

by use case: - Wienbot

- Singapore / LA /...

2.2 State of the Art

 **TO-DO:** - difference alexa + AVS / Lex & Polly

2.2.1 Node.js

 **TO-DO:** why node js: event-driven, non-blocking I/O model

2.2.2 Alexa Skills, Amazon Voice Services

2.2.3 comparison to Amazon Lex / Amazon Polly


2.3 D115

- summarize infobroschuere_ BMI08324_screen_barrierefrei.pdf
- Use case im Detail
- Welche Daten gibt es?
- Was sind die Erwartungen?
- wie kann man die Güte des Systems beurteilen?
- Meist sollte man in diesem Kapitel die Lösung schon im Auge haben, um die Erwartungen so zu formulieren, dass die Lösung auch geeignet ist?

2.4 Frameworks and Data Structures (change title)

- AL: Ich würde erst etwas die Algorithmen und Datenstrukturen (Textanalyse, JSON, ggf. Graphen beschreiben. -AL: Anschließend die Frameworks vorstellen
 - AL: Wichtig ist: Aus den Beschreibungen eine Schlussfolgerung ableiten, welche Art von Lösung entwickelt werden soll.
- for current bot:
- Lucene **as the golden standard**: spell check, unscharfe suche, Tika / detect language / ...
 - Solr - explain what's an intent, whats a slot <https://service.berlin.de/virtueller-assistent/virtueller-assistent-606279.php>
<https://www.itdz-berlin.de/>

2.4.1 Intents and Slots

 **TO-DO:** explain json

provided in JSON for value lookup, there are

Figure 2.1: Dienstleistungen.json - Primary Nodes

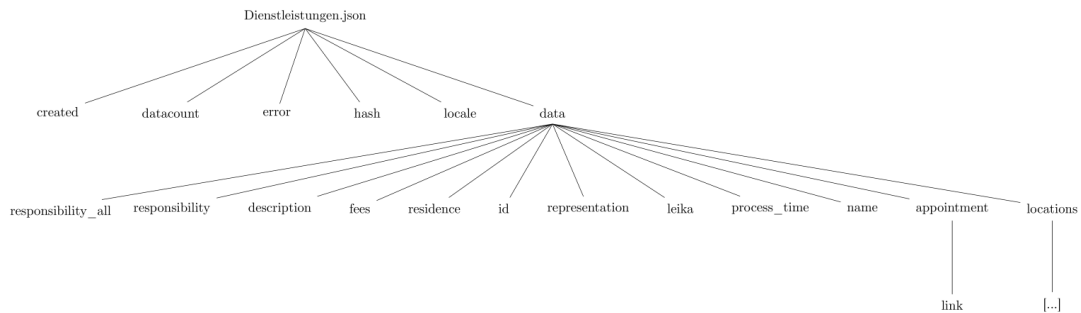
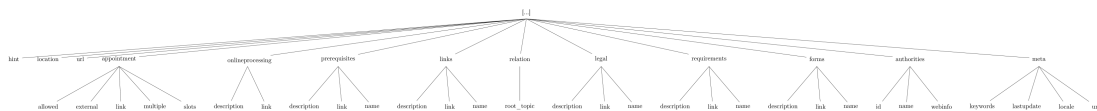



Figure 2.2: Dienstleistungen.json - secondary Nodes



- 616 Intents as `data`, each containing

 **TO-DO:** missing variables e.g. are required papers, flag: persönliche Vorsprache ja nein, ...

- `<string> responsibility` denoting in which city halls a service is available
- `<boolean> responsibility_all` a flag set to true in case the service is available in all local authority offices / service points
- `<HTML list string> description` not unified and includes text
- `<string> not unified` and might need to have an `\lstinlineint`— added to it and set to 0 in case service is free
- `<int>residence`
- `<int>id`
- `representation`
- `<long>leika`
- `<string> process_time` need to derive minimum, average and maximum service times instead of a string, as well as conditions

- `<string> name` the name of the service that would make sense to a human
- `<node> appointment` with
 - * `link` (Key value with URL to /terminvereinbarung page) - check if orphan or if it is for each behörde and in that case how it gets the right one
- `<node> locations`
 - * `hint`
 - * `<int> location` one of the 12 authorities
 - * `url` of that service at that authority
 - * `<node> appointment` (a second one)
 - .
- `<node> onlineprocessing`
- `<node> prerequisites`
- `<node> links`
- `<node> relation`
- `<node> legal`
- `<node> requirements`
- `<node> forms`
- `<node> authorities`
- `<node> meta`

2.5 currently deployed bot

- dienstleistungen.json structure (finding the info through hierarchical nodes)
- interpreting the nodes as intents - traversing the nodes (one level up then to next node)
- no session/no persistence

2.6 Implementation Possibilities

- structure of Hitlist on berlin.de is provided by ITDZ - as opposed to Versicherungs-firma z.B (ML tries to detect irregular patterns in case customer is lying). - unfortunately forums vs. FAQs did not work. if i want assistance, i want the customer to tell me the model number - and forums have mostly Schrott!
- what the bot curenly achieved is at least not give wrong answers, sometimes says idk but it doesnt confuse u. same attitude like in german shops (nur unpassende antworten

sind frustrierend!

-Vorgehensweise: XML -> index über Lucene - > solr knoten...based on sth like when i say am 10. augustit gets me masalan events..aha august ist ein monat, monat relates to calendar, calendar relates to events

Chapter 3

Implementation as Alexa Skill

- as an example for voice
- System Specifications
- System Structure
- UML Diagrams
- Design Choices
- scopes and granularity

3.1 All about Alexa

https://en.wikipedia.org/wiki/Amazon_Alexa
<https://medium.com/@robinjewsbury/how-to-create-bots-and-skills-for-facebook-messenger-and-amazon-echo-4>
- Alexa Appstore had over 5,000 functions ("skills") available for users to download,[18] up from 1,000 functions in June 2016. McLaughlin, Kevin (16 November 2016). "Bezos Ordered Alexa App Push"Paid subscription required. The Information. Retrieved 20 November 2016.

Perez, Sarah (3 June 2016). "Amazon Alexa now has over 1,000 Functions, up from 135 in January". TechCrunch. Retrieved 5 August 2016.

3.2 Difference Between Lex and Alexa Skills

<https://stackoverflow.com/questions/42982159/differences-between-using-lex-and-alexa#URL>
<https://aws.amazon.com/lex/faqs/>
<https://aws.amazon.com/about-aws/whats-new/2017/09/export-your-amazon-lex-chatbot-to-the-alexa-skills-kit/>

Amazon Lex is a service for building conversational interfaces using voice and text. Powered by the same conversational engine as Alexa, Amazon Lex provides high quality speech recognition and language understanding capabilities, enabling addition of sophisticated, natural language chatbots to new and existing applications. Amazon Lex reduces multi-platform development effort, allowing you to easily publish your speech or text chatbots to mobile devices and multiple chat services, like Facebook Messenger, Slack, Kik, or Twilio SMS. Native interoperability with AWS Lambda, AWS MobileHub and Amazon CloudWatch and easy integration with many other services on the AWS platform including Amazon Cognito, and Amazon DynamoDB makes bot development effortless.

3.3 APIs and SDKs

- swagger for handling JSON requests?
- <https://github.com/alexa/alexa-skills-kit-sdk-for-nodejs>

3.4 challenges

- und Lösungen dafür
- eine Überführung in Alexa, not writing everything new in alexa. such that when you want to do it in another system what do u want to integrate?
- use external web service maybe? in case that helps instead of alexa doing everything..
- konten hosting to be on alexa
- wo hilft mir alexa, was mach ich lieber woanders?
- Ähnlichkeitsmaße -levenstein-distanz, IFTTT

Chapter 4

Implementation as Facebook Messenger Bot / Google Action

- as an example for text
 - implementing the answer suggestions as buttons
 - passing data to the Bürgeramt terminseite
- <https://console.dialogflow.com/api-client/>
<https://console.actions.google.com>

Chapter 5

Evaluation

- benchmarks
- strengths and weaknesses
- challenges
- performance
- usability
- feasibility of using the studied agents
- node.js?
- amazon's system testing options (incl. Betas)
- system usability scales (ISO, DIN)
- Con: Alexa skills are listed in the amazon shop page. Sehr unübersichtlich just like prime
- impression: Amazon collects data and makes something "intuitive out of it for you". e.g. fire stick setup already had account linked before connecting to the internet! scary/funny/ but then it could be counterintuitive at some point if u want to do ur own customizations.
- removing bias in recruitment of participants (diversify based on what categories?)
- EVAL: AUC/ROC, true positives, false...no of utterances to text
- compare with Wiener Stadportal as a benchmark for a bot
- <https://www.wien.gv.at/bot/>
- <http://www.vienna.at/wienbot-chatbot-der-stadt-wien-informiert-als-virtueller-beamter/5590853>
- <https://digitalcity.wien/wienbot-auszeichnung-fuer-chatbot-der-stadt-wien/> singaporebot

5.1 Results

usability metrics: - heuristic eval - guidelines (**jakob nielsen, ralf molich whitepaper**)

- biggest usability flaw
- cognitive walkthrough
- step-by-step approach
- questions..will the user tr and achive
- pluralistic walkthrough
- panel method
- hallway testing
- A/B Test
- speed and Bottlnecks

- clientele: census / SOEP, who can use the bot
- make a small prediction (Bus Analytics)
- this Hassloch thing from MKTG

5.2 Discussions

- Evaluate the system:
 - is it trivial to build such a bot or not / what is the aufwand
 - how does it react with longer sentences? some service names are long
 - what does levenstein distanz cause
 - wie leicht kann ich eine antwort finden auf das was ich suche?
 - how am i going to classify my tests?
-
- are chatbots being pushed on the market or is there a demand? (kleine Umfrage basteln?)
 - how easy or difficult it is to make a bot: planing poker - varianz anschauen zw. leicht und schwer und iterativ darüber sprechen
 - wo kann der Kunde (Sawa2 kan el end user or the senat in our case) help optimize the bot masalan bürgeramt beyektebo, welche Rechtsgrundlage keine auffällige Probleme masalan zay Perso, PA, personalausweis, how to introduce expert mode so that if u add it with a special character it knows what u want, just like alexa knows when u rename the lamp - refer again to use cases and exper vs personal field

Chapter 6

Conclusion and Future Work

6.1 Summary

6.2 Conclusion

6.3 Future Work

- use machine learning to rank higher demands for more popular services.
- matkhoshesh fel 7etta di awi - for now hitlist already given.
- future of bots. deren Einsatz. roles (As judges, catereres in hotels (that hotel botler)

Bibliography

- [1] Amt für Statistik Berlin-Brandenburg. Statistischer Bericht - A I 5 hj 1. https://www.statistik-berlin-brandenburg.de/publikationen/stat_berichte/2017/SB_A01-05-00_2017h01_BE.pdf, August 2017. Accessed on 06.02.2018.
- [2] Brian Roemmele, Dave Isbitski. Episode 018 - A Voice First Future with Brian Roemmele. *Alexa Dev Chat Podcast*, 11 2017. Accessed on 13/02/2018.
- [3] Wikipedia contributors. History of natural language processing — wikipedia, the free encyclopedia, 2017. [Accessed 14 February 2018].
- [4] Wikipedia contributors. Usage share of operating systems — wikipedia, the free encyclopedia, 2018. [Accessed 14 February 2018].
- [5] Geschäfts- und Koordinierungsstelle 115 im Bundesministerium des Innern. {https://www.115.de/SharedDocs/Publikationen/Service_Publikationen/Infomaterialien/infobroschuere_%20BMI08324_screen_barrierefrei.pdf?__blob=publicationFile&v=8}, November 2013.
- [6] Tobias Grossmann, Regine Oberecker, Stefan Koch, and Angela D Friederici. The developmental origins of voice processing in the human brain. 65:852–8, 03 2010.
- [7] Judy Goldsmith and Nicholas Mattei. Science fiction as an introduction to ai research, 2011.
- [8] M. Ballve. Messaging apps are now bigger than social networks, 2016.
- [9] Shankar Vedantam. Radio Replay: I, Robot. <https://www.npr.org/podcasts/510308/hidden-brain>, 2018. Accessed on 25.01.2018.
- [10] Statista. Smartphone user penetration as percentage of total global population from 2014 to 2020, November 2016. Accessed on 06.02.2018.

- [11] Joseph Weizenbaum. *Computer Power and Human Reason: From Judgment to Calculation*. W. H. Freeman & Co., New York, 1976.

Appendices

Appendix A: Abbreviations

AWS	Amazon Web Services
ASK	Alexa Skills Kit
AVS	Alexa Voice Service
ARN	Amazon Resource Name
MVP	Minimum Viable Product
AI	Artificial Intelligence
NLP	Natural Language Processing
ML	Machine Learning
GUI	Graphical User Interface
appx	approximately

Appendix B: Glossary

Intent	erklärung
Slot	erklärung
Utterance	erklärung
Alexa	erklärung
Alexa Skill	erklärung
Lambda Function	erklärung
Alexa Skills Kit	erklärung
Amazon Developer Console	erklärung
AWS Lambda	erklärung
Amazon Lex	erklärung
Amazon Polly	erklärung
ElasticSearch	erklärung
node.js	Framework built on top of JavaScript
Interaction Model	erklärung
Service	bot, AWS, Berlin.de
https://docs.aws.amazon.com/general/latest/gr/glos-chap.html	
Application ID	erklärung
Skill ID	erklärung
Bot	Unless otherwise mentioned, yeb2a Cha
Hitlist	erklärung