Students

Al-Hubaishi Waleed iCompetence

Kaminsky Jens Informatik

Coaches

Kalunder Madlaina  
Agotai Doris

How to create a cool App <todo>

**Swiss Engineering  
Event App (SEEA)**

IP5 Documentation 2019-XX-XX

# Summary

The goal of this project is to create an app for the Swiss Engineering Association that will attract new members. To achieve this goal, a voice recognition/generation feature was suggested. This Document will summarize how such a feature could be implemented, what is currently possible and even what might be possible in the future by looking at cutting edge research that is currently being done in this section.

Contents

[Summary 1](#_Toc7371879)

[Introduction (problem, current state) 2](#_Toc7371880)

[Research (FACTS) 2](#_Toc7371881)

[state of the art 2](#_Toc7371882)

[Methodologies 2](#_Toc7371883)

[Topic in depth 2](#_Toc7371884)

[What voice recognition/generation is currently used on the market 2](#_Toc7371885)

[Technical research 3](#_Toc7371886)

[Google assistant 3](#_Toc7371887)

[Siri Assistant 4](#_Toc7371888)

[Alexa Voice Service (AVS) 5](#_Toc7371889)

[concept 6](#_Toc7371890)

[what did we come up with, what do we want to implement 6](#_Toc7371891)

[Potential ideas 6](#_Toc7371892)

[Prototypes, design, details 11](#_Toc7371893)

[implementation 11](#_Toc7371894)

[Analysis & future thoughts 11](#_Toc7371895)

[what's out of scope, what could be done in future 11](#_Toc7371896)

[attachments (sources, literature) 11](#_Toc7371897)

[everything that might be interesting for the expert 11](#_Toc7371898)

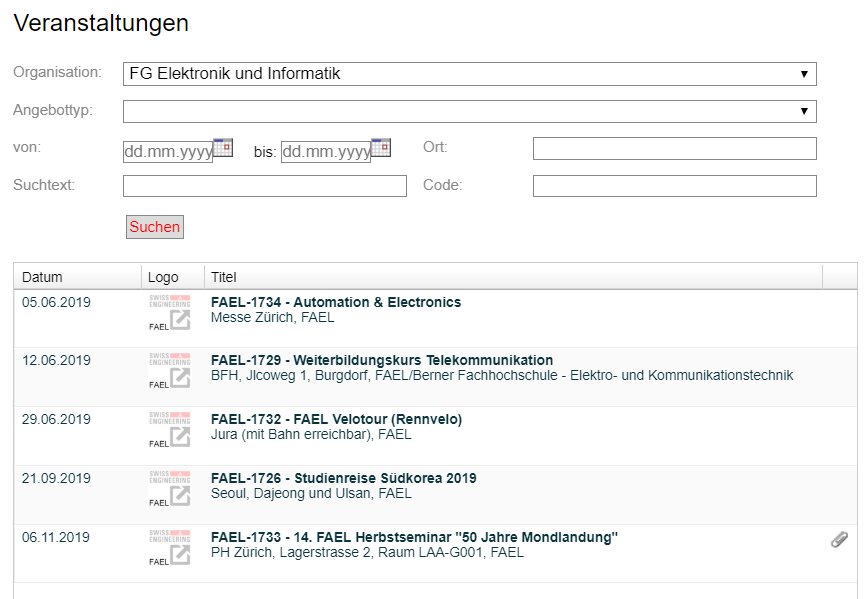
[literature references 11](#_Toc7371899)

[honesty policy 11](#_Toc7371900)

# Introduction (problem, current state)

Swiss Engineering is an association that spans over multiple professions and has a total of around 13’000 members in Switzerland. They are holding over 100 events annually over multiple of their Sections. These events are not only for current members also have the additional benefit of attracting new future members.

The Swiss Engineering association is currently in the process of updating their event booking website. To reach a greater audience they intend to also have a native mobile app to make booking on the go easier. This application would also be showcased on events with the idea to improve the image of the Swiss Engineering by showing that it has a modern app.



# Research (FACTS)

## state of the art

<https://ieeexplore.ieee.org/abstract/document/8462105> (automatic speech recognition ASR)  
<https://ieeexplore.ieee.org/abstract/document/8585066> (visual only recognition, promising for noisy environments)  
<https://arxiv.org/abs/1904.05862> (wav2vec, promising unsupervised learning that is character-based)

<https://ieeexplore.ieee.org/abstract/document/8461972> (multilingual support with a single model, might be very useful for countries like Switzerland)

<TODO, summarise findings>

## Methodologies

<TODO>

## Topic in depth

This section will go into depth about what voice recognition/generation applications are currently available on the market, what kind of devices are available and how the human-machine interface is used to enhance the user experience.

### What voice recognition/generation is currently used on the market

## Technical research

This section will go into depth about which existing implementations for natural language processing are available and how they could be integrated in the SEEA solution

## Google assistant

Google assistant is a platform independent natural language processor. It supports both Android and iOS devices.

It is recommended to use api.ai for the natural language processing in the background. Api.ai tries to extract intents and entities from the input.

#### Entity

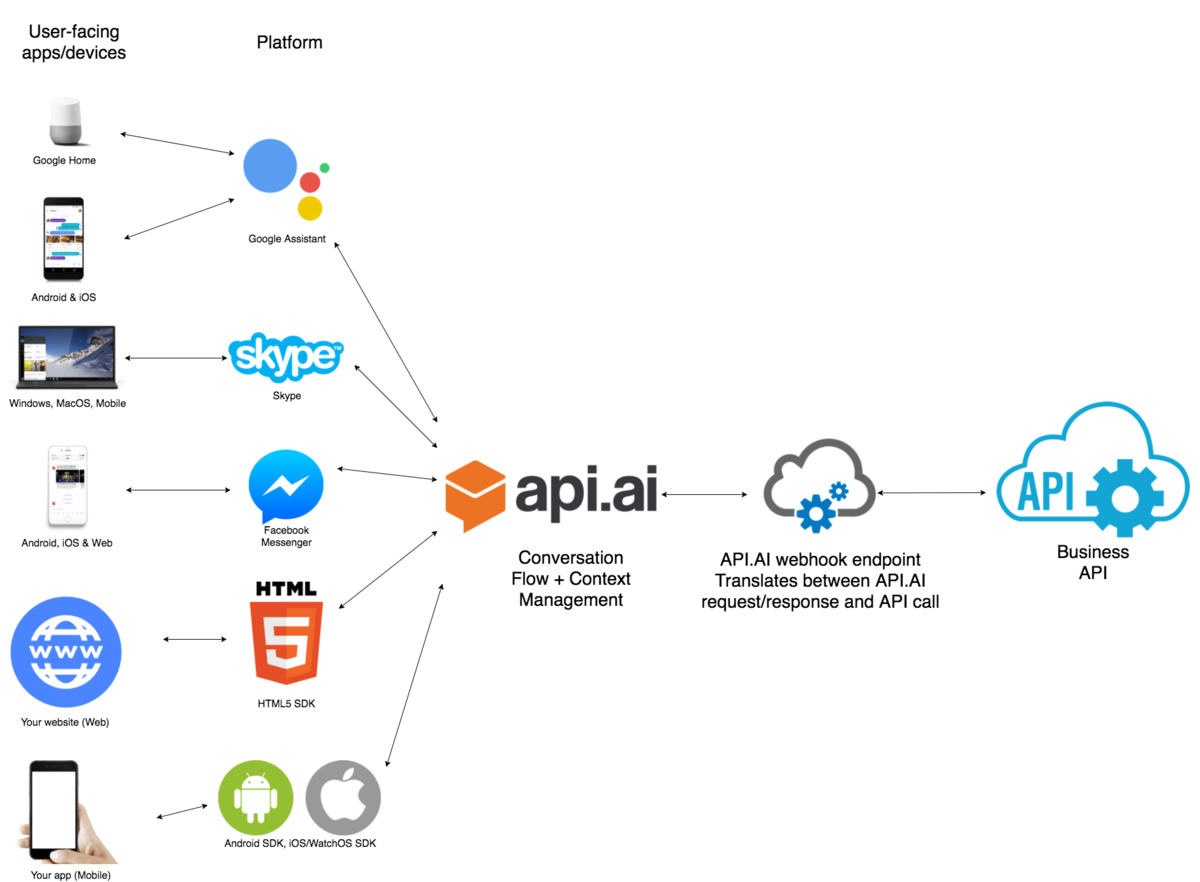
Entities could be for example Animal.

The Entity Animal contains multiple different animals, e.g. dog, cat, bird, etc.

After creating a list of things that are part of the “Animal” Entity you can also help the A.I. by providing synonyms. E.g. A “Puppy” is also a Dog, “Dogs” can also be interpreted as Dog, etc.

#### Intent

Intents are actions that can be derived from phrases the user tells the program. E.g. “Tell Me A Joke” would be an intent, the program will be prompted to execute the task (intent). Like for entities, the developer needs to provide the algorithm with example sentences that should trigger the intent. The more examples are given, the more non predefined sentences the A.I. can use to trigger the intent.



#### Pricing:

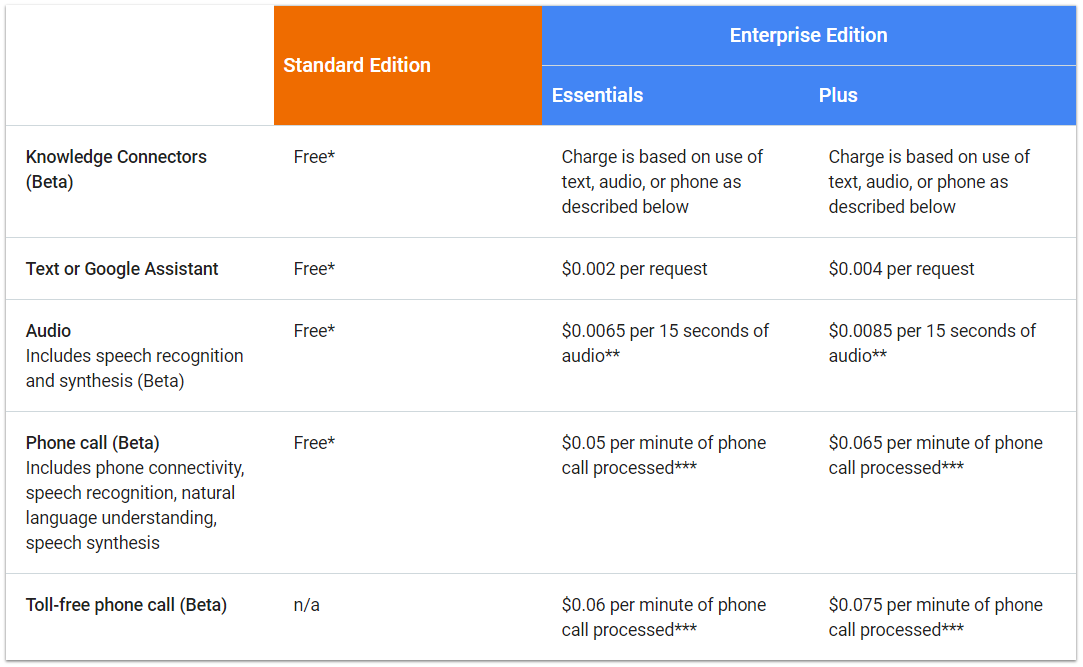
**Dialogflow Standard Edition** provides all of the core features of Dialogflow, but interactions are limited by usage quotas, and support is provided by the community and e-mail. It is ideal for small to medium businesses that want to build conversational interfaces or those who want to experiment with Dialogflow.

**Dialogflow Enterprise Edition** provides higher usage quotas and support from Google Cloud support. Dialogflow Enterprise Edition is a premium offering, available as a pay-as-you-go service. It is ideal for businesses that need an enterprise-grade service that can easily scale to support changes in user demand.

Available in two pricing plans:

**Essentials**: This plan contains all features offered by Dialogflow Standard Edition, plus enterprise-ready quotas for speech recognition, speech synthesis, and telephony gateway.

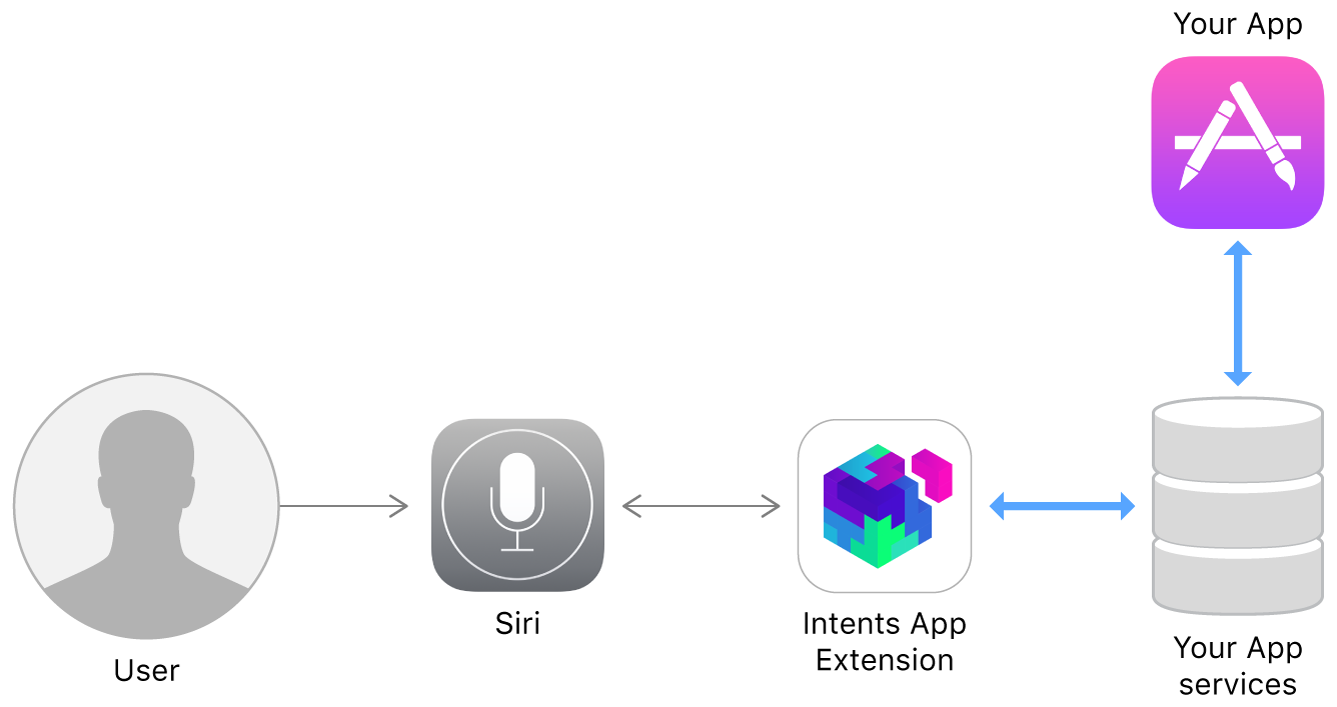
**Plus**: This plan contains all features offered by Essentials, plus enterprise-ready quotas for knowledge connectors. Each request from an Enterprise Plus agent performs the regular intent recognition and entity extraction, as well as a knowledge connector search.



### Siri Assistant

The development environment of Siri is called SiriKit and it offers the possibility to translate calls to the assistant into calls for the app you want to use. For this to work an internet connection is mandatory as the voice data is sent to apple servers for the natural language processing.

Apple has a strict List of so-called Domains and Intents which the application call has to be assigned to for it to work.



Domains:

* Lists: Simple list creation and editing
* Visual Codes: Display QR codes
* Ride Booking: Requesting a ride, this is intended for services like Uber
* Messaging: Simple text messaging
* Photo Search: Making a search request for photos and videos and display them.
* Payments: Creating a payment request or sending money
* VoIP Calling: Make a call (also supports video)
* Workouts: Starting, pausing and finishing a routine, explicitly specified as workout
* Climate and radio:

Pros:

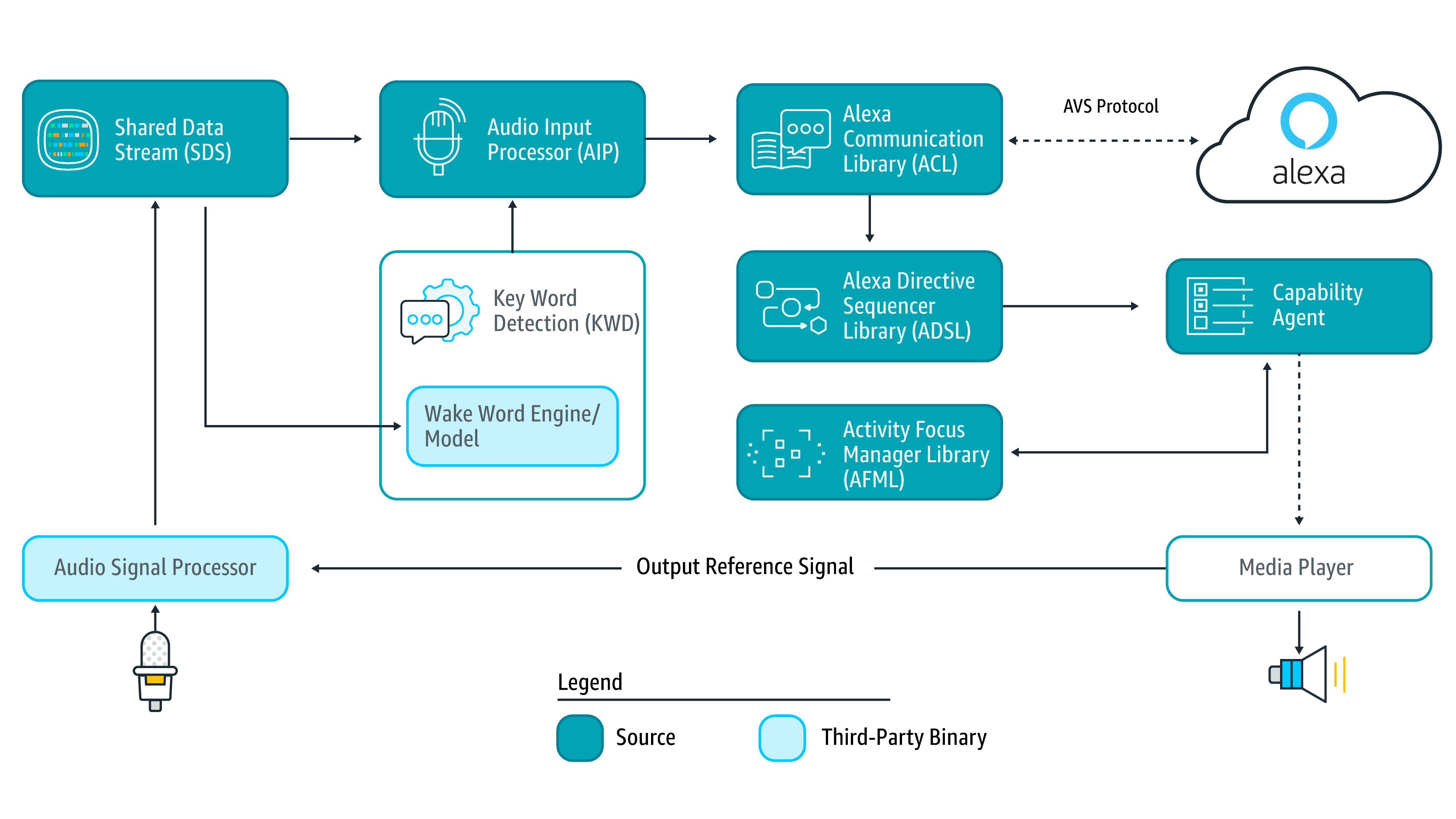
+ Apple would take over the NLP part

Cons:

* Only works with an internet connection
* Domains limit the usability for the app
* Always requires specific keywords for it to work

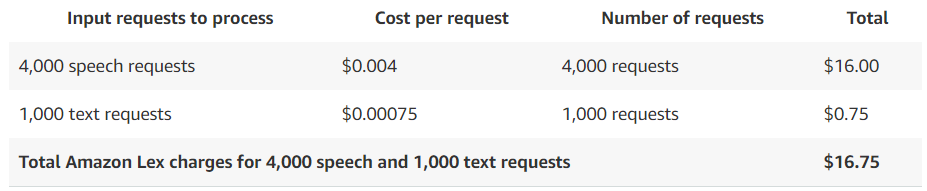
### Alexa Voice Service (AVS)

AVS is a platform independent service that is used over an HTTP/2 api.

AVS is built like Siri in the way that it is task oriented. The tasks of AVS are called directives and events, the following are supported:  


* Alerts: timer and stopwatch features
* AudioPlayer: control music playback, etc.
* Bluetooth: manage Bluetooth connections
* DoNotDisturb: enable DoNotDisturb mode on device
* EqualizerController: control equalizer settings and equalizer modes
* InteractionModel: Allow the client to support complex interactions and Alexa routines
* Notifications: API for notifications
* PlaybackController: navigate playback queue via GUI or buttons
* Speaker: volume control, mute and unmute
* SpeechRecognizer: API for speech capture
* SpeechSynthesizer: Text to Speech API
* System: System state
* TemplateRuntime: visualize metadata of requests

Pricing:



# concept

## what did we come up with, what do we want to implement

### Potential ideas

## Prototypes, design, details

# implementation

# Analysis & future thoughts

## what's out of scope, what could be done in future

# attachments (sources, literature)

## everything that might be interesting for the expert

# literature references

[https://github.com/alexa/avs-device-sdk 2019-04-26](https://github.com/alexa/avs-device-sdk%202019-04-26)

<https://aws.amazon.com/lex/pricing/> 2019-04-26

<https://developer.apple.com/sirikit/> 2019-04-26  
[https://cloud.google.com/dialogflow-enterprise/pricing 2019-04-27](https://cloud.google.com/dialogflow-enterprise/pricing%202019-04-27)  
<https://developers.google.com/actions/distribute/> 2019-04-28

<https://www.swissengineering.ch/index.php?option=com_content&view=article&id=4&Itemid=6&lang=de> 2019-04-28

<https://xamarinhelp.com/interacting-google-assist-xamarin/> 2019-04-28

# honesty policy