Software Project (Verto Analytics)

Libby, the library assistant for Amazon Alexa

# Introduction

* Verto Analytics is an audience measurement company that analyses what people do with their personal devices. Personal assistants are a new emerging class of devices. Verto Analytics wants to start investigating them.
* Conversational UI and personal assistants (Amazon, Google, Microsoft, etc.) provide

modern interaction models for new and traditional applications alike. Finland is being

serviced to some extent by these devices, which are exceedingly popular in supported

countries, through use of the English language. Upcoming versions of such devices may

support Finnish natively. The project being proposed will build experience in design and

implementation for personal assistants, skills for them, and their adaptation to the Finnish

market.

* The project is designed as a pilot project that would initially target Aalto students and

employees as end-users. To be designed is a personal assistant that answers questions

about Aalto services, with a focus on libraries.

The product will mostly be used in the Library. It has to be “cool” enough (i.e. fast, reliable, enough features, etc.) for people to actually want to use it. It should not be placed in noisy environments, such as close to a woodworking room, but should obviously not be placed in quite quarters. This means that suitable places for Libby would be in, e.g., the Library lobby. There was also discussion about placing them in guild rooms and possibly implementing extra features for these environments, but this is secondary to the Library. Verto will not give up the idea of having Libby placed in the Library, but if have ideas we should absolutely propose them.

# Project Goals

## Goal A

Goal A is to implement a “skill” to search for information on Aalto services in spoken

communication. This goal is designed to result in a usable product based on existing

technologies. The “Libby” skill to learn about Aalto libraries will be implemented by students as a

cloud service for integration with Amazon Alexa, and communicate with end users in a language

supported by Alexa (i.e. English). For example:

* “Alexa, ask Libby where the learning center is”.
* “Find Theory and methods of programming”. Etc.)

Logic is implemented by querying a third-party service, in particular, aalto.finna.fi. The skill will

answer in English, e.g. “I found Theory and methods of programming in the Collection of Computer

Science in Learning Hub Greenhouse”)

## Goal B

Goal B is to assemble and configure a hardware device that could be used with this skill. After the embedded hardware is chosen, the hardware should be configured with an Alexa Voice Service (AVS) client. A small instruction leaflet should be designed to teach users how to query the Libby service using this device. An engine will be configured or trained to activate the hardware using a custom wake word. The hardware can be installed e.g. in student facilities and used by anyone who is passing by.

## Goal C

Goal C is a stretch goal that builds on the prototype of Goal B by customizing the software so that Alexa is no longer required. The task is to replace AVS calls with API calls to Amazon Lex. The improved product will respond to queries of the form: “Libby. What are the opening hours of the learning center?”. Users of this improved product can interact with it directly, without knowing how to address Alexa skills.

## Goal D (ultimate goal)

Goal D is the ultimate goal, which will adapt the product to an international audience and open development up to experimental features. Queries spoken by the end-user will be processed using speech recognition (e.g. Bing REST API) and pre-processed on the device using simple natural language processing (NLP) techniques. The aim of this ultimate goal is to move as many NLP-related steps as possible out of AWS Lex and into the project’s own control. From the perspective of end users, this will permit additional languages (Swedish, Russian, etc.) to be understood and synthesized.

# Future vision and extra comments

## Authentication

An idea for authentication was presented, which would include personalized features such as a calendar. This would provide personalization, but would be complicated to actually implement. If this idea would be developed, a companion app for Libby could be a good way for authentication purposes.

## Private Version

Another idea that emerged during the meeting would be a private version for people who have an Echo dot at home (if the finalized product runs through Amazon services). This version would also be able to access normal Libby info. Implementing this idea would probably also need a companion app.

## User Data

No user data is stored as. Verto is not interested in obtaining user data. However, having some data on which features are used the most is very good to have and should be provided. This data should be such that they can be represented on e.g. pie-charts.

## Extra Hardware

The possibility to add more hardware to the system can be developed into an idea. However, adding hardware, such as a screen, has to work seamlessly and actually be a part of the system. It should not take away from core functionality (i.e. speech recognition). This means that adding a screen could potentially be a huge task and should be discussed properly. However, an idea for a screen would be that Libby presents a map with relevant locations when asked for, e.g.: “Where is [something] book?”.

## Languages of interest

Languages that Libby should support are English, Swedish, and Finnish. Verto stated that the product vision would stay unfinished if no Finnish support is implemented. We should thus, as a team, research tech for Finnish voice recognition, Microsoft has something for voice recognition that we could look into.

*When Verto was asked for the product vision the answer was: “All of this”. This refers to everything above. As explained earlier, all features do not have to be actually implemented, but if asked for a product vision with these features in mind the finalized vision would be a product with everything.*