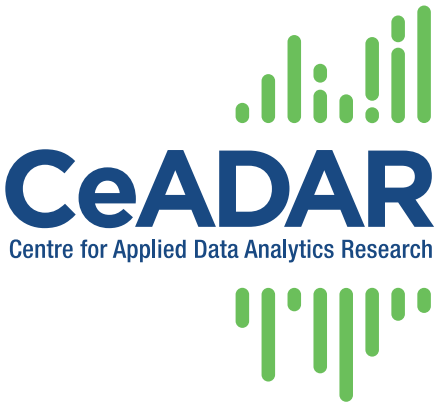

CeADAR – Centre for Applied Data Analytics Research

Enterprise Ireland Data Analytics Technology Centre

Intelligent Analytic Interfaces: Voice Controlled BI - Technical Specification

Document Type:	Technical Spec
Project Title:	CeADAR
DDN Theme:	1 – Intelligent Analytic Interfaces
Theme Leader:	Brian Mac Namee (DIT)
DDN Sub-Theme:	1.2 – Beyond the Desktop
Authors:	Eoghan O'Shea, Robert Ross & Brian Mac Namee
Document Version:	1.0
Date of Delivery to ISG:	26th May 2014
Number of pages:	5

ABSTRACT

Our Voice Controlled BI system (Speak+Seek) is one that allows a user, through the use of voice commands only, to extract useful information from Business Intelligence (or other) software systems. This document will provide a high-level technical specification of this system.

CeADAR is a research partnerhip comprising University College Dublin, University College Cork, and Dublin Institute of Technology.

http://www.ceadar.ie

Contents

1 Description of Industry Needs

There are many commercially available BI systems on the market today, including offerings from Nathean, SAS, Tableau and Qliktech. One of the most interesting recent developments in the design of these systems is that they are being targeted at non-technical users within organisations. Often referred to as *self-service BI* tools, these offerings typically focus on ad-hoc reporting and require interfaces that do not pre-suppose an understanding of technical topics such as statistics or query languages like SQL. Heavy use is made of data visualisations, *drag and drop* and *point and click* interfaces. Although these interfaces are useful in unlocking the power of BI to a wide range of users, they can become cumbersome and time consuming to use – particularly in a mobile setting. Voice recognition systems can be used to address this situation, allowing users to call up information using simple spoken commands.

Speak+Seek is a system that seeks to do just that, allowing a move away from drop-down lists and form filling towards a less complicated, voice controlled human-computer interaction. This document provides a high-level technical specification of this system.

2 System(s) Involved

Speak+Seek will be a voice-driven system that allows voice interaction with data held in executive dashboards (the dashboard shown in Fig. 1 is a typical example). This will allow users on mobile devices to query data normally displayed on a desktop or tablet dashboard by voice while on the move without trying to navigate a tricky mobile version of the dashboard. A typical user scenario would involve a user who is away from their desk but who requires a piece of information in a hurry, e.g., a piece of sales information before an important meeting. The user would activate Speak+Seek, speak their request, and Speak+Seek would retrieve the required information from the user's BI dashboard and present it back to the user. Therefore, the Speak+Seek system relies on access to a BI software systems in which a designer has already designed and built executive dashboards. The data accessed by Speak+Seek will be exclusively the data held in the dashboard and no access to raw data sources beyond this will be made.

The initial version of Speak+Seek will be modelled on dashboards found in Logix by Nathean. In this version the dashboards used will be developed by CeADAR and no actual integration with Logix will be performed. It will also be possible to integrate the system developed with any other similar BI system.

The other systems involved in building Speak+Seek will be a voice recognition engine, NLP toolkits and a voice generation or Text-To-Speech (TTS), engine. Likely candidates for these systems will be described in the next section.

3 Approach

The Speak+Seek system will be developed as an Android smartphone app initially, with the intention that it be used to access a user's BI dashboard when away from the desk. The app interface itself will be very simple, requiring just a control that allows the user to speak their query and see and hear a response. Fig. 2 shows a general overview of the proposed

Figure 1: A typical executive dashboard (from `www.elementwave.com`).

Speak+Seek system. In step 1 a user activates the Speak+Seek system and speaks a defined command into their smartphone, requesting, for example, a piece of sales information from a graph sitting within a dashboard in their BI software. In step 2 the smartphone speech recognition system, activates and transforms the spoken words into text. In step 3 this text, the user command, is converted from natural language into a machine-understandable query. In step 4 this machine-understandable query is used to extract the requested information from the user's BI dashboard. Finally, in step 5 this information, in the form of natural language, is converted into speech using a TTS conversion and spoken back to the user (it will also be written to the screen).

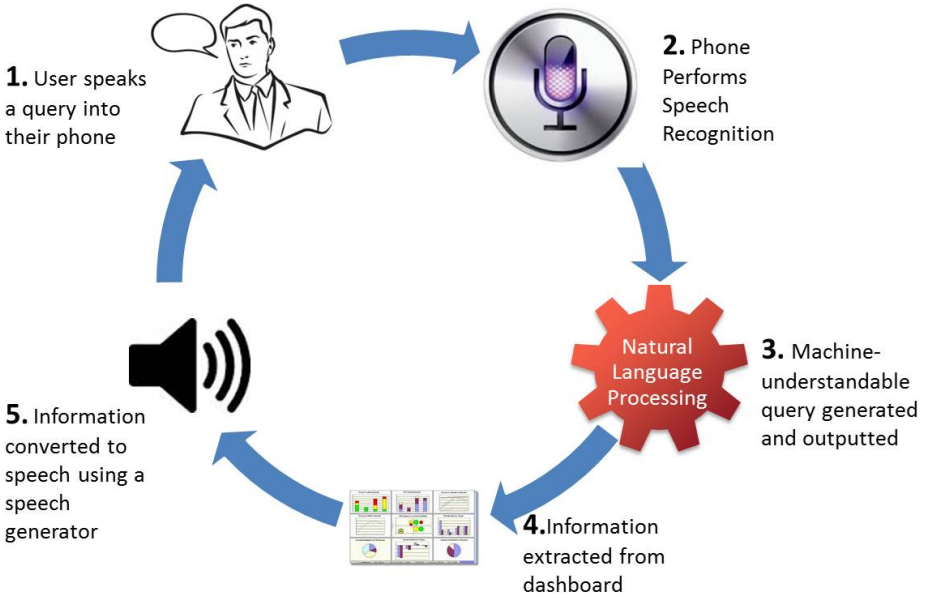

Figure 2: The key components of the Speak+Seek system.

There are four major elements involved in the Speak+Seek system:

1. Speech recognition: Speech Recognition has improved significantly in recent years. In smartphones there are a number of powerful built-in speech recognition systems that we can leverage - for example the Android voice recognition system. All of these speech recognition systems are known to perform particularly well on the type of short utterances and commands that would be used in a voice controlled user interface system such as Speak+Seek.
2. Data model: In order to perform queries on the data behind any chart in a dashboard a defined data model is required. For most dashboard charts a simple data model containing two columns is sufficient to model any data series displayed. The first column will contain the category labels while the second will contain the values associated with each category. For example, the data underlying the pie chart in Fig. 1 can be represented using this model as follows:

Category	iPhone	iPad	A. Tablet	Android
Value	65	15	10	10

   In order to voice enable individual Dashboards additional meta data is also required. At its simplest this meta data can just be a name for each data series but we will explore more sophisticated solutions such as the provision of a Wordnet (`wordnet.princeton.edu`) synset for each series. Access to any other annotations associated with a dashboard chart (such as title, axis labels etc) will also offer opportunities for richer interactions.
3. Translation from natural language into dashboard data queries: Translation from natural language to formal queries (e.g., SQL) or other commands is a long-standing research area. Both open source NLP toolkits such as NLTK (`www.nltk.org`) and licenceable products such as TextRazor (`www.textrazor.com`) provide many of the underlying tools which will be required for this parsing and trasnlation task.
   The following are examples of the types of queries that the initial system will be built to handle (assuming a dashboard similar to that in Fig. 1):
   - *What is the most common installed device?*
   - *What month had the highest app usage?*
   - *How many pushes were sent in October?*
4. Speech Production: The results of a user's query or request, in the form of some specific piece of information will be realised into a natural language answer and converted back into speech using a Text To Speech (TTS) engine (for example the Android TTS engine). The fields of natural language generation and TTS technologies are both well-developed research areas in which DIT have in-house expertise. The language generation will be based on template techniques which will be facilitated by the focus on reasonably simple interactions around a dashboard.