ISSN: 0976-3031

Available Online at http://www.recentscientific.com

CODEN: IJRSFP (USA)

International Journal of Recent Scientific Research
Vol. 9, Issue, 1(G), pp. 23349-23352, January, 2018

International Journal of Recent Scientific Research

DOI: 10.24327/IJRSR

Research Article

A STUDY ABOUT VOICE ENABLED APPLICATIONS FOR DOCUMENT SEARCHING SERVICE

Jeevitha V and Kavitha E.S

Department of Library & Information Science, Periyar University, Salem-11

DOI: http://dx.doi.org/10.24327/ijrsr.2018.0901.1451

ARTICLE INFO

Article History:

Received 17th October, 2017
Received in revised form 21st November, 2017
Accepted 05th December, 2017
Published online 28th January, 2018

Key Words:

Voice Application, Voice Search Software, SAP Netweaver

ABSTRACT

This paper studies and evaluates the voice enabled applications for in build the current system. The further develops new library system with enhanced features which are the new searching subsystem by using SAP Net weaver a high-performance, full-featured voice search software for library written entirely in Java or XML. So it is introduce to improve the searching function of current system.

INTRODUCTION

Why speech?

There are hundreds of apps that let users search, write emails, take notes and set appointments with their systems. But, for some people, the small size of a keyboard or touch screen can be limiting and difficult to use.

Speech is ideally suited to mobile computing, as users tend to be usually on the move and speech can not only take commands but also read out messages. Not to say that voice based apps will improve accessibility for a person with a disability. Progress in technologies needed to help machines understand human speech, including machine learning and statistical data-mining techniques, is fuelling the adoption of speech in smart devices.

Successful speech implementations

There have been several successful case studies of successful implementation or introduction of voice to apps. Ask.com, a question answering-focused web search engine, has integrated Nuance's speech recognition technology into its iOS and Android apps. This linkage lets users use their voice to ask and answer questions as well as comment on answers.

Amazon has also updated its Kindle iOS app with support for Apple's "VoiceOver" reading and navigation feature for blind and visually impaired users of the iPhone and iPad. Amazon mentioned that more than 1.8 million e-books will support the technology, which automatically reads aloud the words on the page. The decision to use Apple's VoiceOver is notable in part because Amazon had previously acquired a company, IVONA Software that provides Kindle "text to speech" and other features.

A recent survey by Forrestor also indicates the growing popularity of voice control apps among mobile workers. A majority of those who use voice recognition use it to send text messages, 46 percent use voice for searches, 40 percent use it for navigation, and 38 percent use voice recognition to take notes.

Integrating speech

The speech technology is in fact built in two parts. The first is a speech synthesizer (often referred to as Text to Speech or TTS in short) that the device or app can use to communicate with user –for example, read text on demand or keep users informed about the status of a process.

The second is a speech-recognition technology that allows users to talk to the app to send commands to it or a message/email dictation what we usually do with a keypad. An ideal voice application should include both, however for starters; it is a good idea to start with one to get a knack of it.

Corresponding author:* **Jeevitha V
Department of Library & Information Science, Periyar University, Salem-11

At first glance, implementing a speech synthesizer/recognizer library may seem like a relatively simple task of determining the phonetic sounds of the input and outputting a corresponding sequence of audible signals. However, it is in fact quite difficult to produce intelligible and natural results in the general case.

This is because of linguistic and vocalization subtleties at various levels that human speakers take for granted when interpreting written text. In fact, the task requires a considerable grasp of both Natural Language Processing and Digital Signal Processing techniques.

To develop a worthwhile speech algorithm from scratch would take tens of thousands of programming hours, hence it makes sense to use one of the several existing tools for it. There are a number of technologies in the marketplace to speech-enable apps. One thing to figure out before choosing your speech SDK, is the deployment model. There are primarily two deployment models:

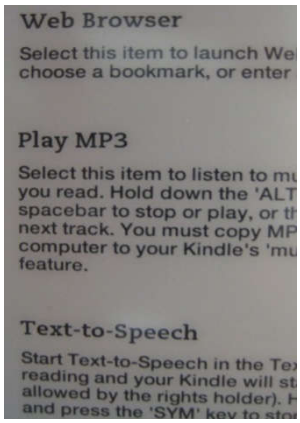
Cloud: In this case, the Automatic Speech Recognition (ASR) or Text to Speech (TTS) happens in the cloud. This offers significant advantage in terms of speed and accuracy, and is the commonly used mode. This also means that the app will require Internet connection at all times. On the positive side, the app will be lightweight.

Embedded*:* In an embedded mobile speech recognition or TTS, the entire process happens locally on the mobile device. When the voice feature is completely embedded, the app can work offline but becomes very heavy.

For instance, TTS engines use a database of prerecorded voice audio where there is a clip for every possible syllable. Using it offline means including all these clips inside your app. In case of IVONA voices, developers can download voice data for US English (Kendra, female) or UK English (Amy, Female) – each has data size of approximately 150 MB. Another advantage of such systems is that they are not affected by the latency that's involved with transmitting and receiving information from a server.

Popular speech libraries

Nuance is perhaps the most popular provider of speech libraries for mobile apps.



Nuance's app Dragon Dictation, a simple speech-to-text app that is free on iOS, is probably the best known speech app. The app sends a digital sample of your speech over the Internet to do the speech recognition, so it requires a wireless connection. But this process is speedy, and the app soon displays your transcribed text in the main window. Apple, Google and Microsoft all now offer direct speech-to-text recognition in their mobile operating systems, and you can use those features wherever there is an option to type in text. One new API available to app developers in iOS 7 is Text to Speech.

In the past, developers would have to integrate their own text to speech solution, which adds time and cost to development and makes apps larger. With iOS 7, Apple is integrating an API that will allow developers to make their apps speak with just three lines of code. Not only will it allow iOS app developers to implement speech in their apps with ease, it will also work in Safari for developers building web applications.

OpenEars is an offline text-to-speech and speech-to-text Opensource library, and supports Spanish as well as English. Like other offline libraries, it can drastically enlarge the app size (since the OpenEars is more than 200MB). However, developers can reduce their final app size by doing away with voices or any features of the framework that they aren't making use of. Depending on which voices they're using and which feature, they might see an app size increase of between 6 MB and 20MB unless they're using a large number of the available voices.

Other popular SDKs include Ivona, iSpeech, Vocalkit and Acapela. All the three are online and paid.

Based on their requirements, developers can select any of these libraries in their apps. The key metric to look at is efficacy of the solution versus total cost of ownership. As these tools become more adept in handling problems such as cocktail party problem, context awareness, various accents and stutters handling etc., the user experience of using speech in apps will also get better.

Developers, however, need to take care of some things when embedding speech. They shouldn't use speech everywhere as of now, and can start at places where changes of error are low or perhaps where the impact will be high (e.g. filling in a form).

Trying to speech-enable the entire app at once can be too big a problem to solve. Rather, it makes sense to voice-enable a couple of the killer use cases first. Find out what are the primary methods or functions that people use in your app and voice-enable these first. It's an iterative process, and the rest of the app can be speech-enabled or speech controlled in a phase II or beyond development cycle.

Similarly, if the app does not contain a lot of text that needs to be read out, using simple audio clips may make more sense rather than using a TTS engine. If you want to track just few keywords, you should not look for speech recognition API or service. This task is called Keyword Spotting and it uses different algorithms than speech recognition. Speech recognition tries to find all the words that has been said and because of that it consumes way more resources than keyword spotting. Keyword spotter only tries to find few selected keywords or key phrases. It's simple and less resource-consuming.

What Solutions Can Be Implemented?

We use SAP NetWeaver Voice to develop applications that allow customers and employees to interact with SAP or non-SAP solutions. These solutions can be internal to or external from your organization.

Practical examples of voice-enabled applications include:

Customer Relationship Management (CRM)

- Travelers call to book a flight or to check departure times.
- Customers dial in to create an order or to check its status.
- Product users call for support.
- Utility clients use the telephone to report their meter readings.

Remote User (RU) Services

- Field technicians dial in to report a malfunction.
- Truck drivers call to update their location or travel status.

Employee Self-Service (ESS)

- Employees call in at work to report their absence.
- System users reset a password by telephone.

What Are the Advantages?

Some major benefits of a voice-enabled application include:

- Providing system access via a telephone to anyone anywhere at anytime
- Making business transactions easy to use
- Eliminating paper-based transactions
- Limiting dependency on mobile devices, computers, and Internet access
- Saving costs by automating telephone processes
- Expanding your business by reaching users on a larger scale
- Leveraging the easy-to-use Visual Composer development environment

What Are the Key Capabilities?

Some key Net Weaver Voice capabilities include:

- Full integration with SAP NetWeaver
- Graphical application modeling using Visual Composer (Voice Kit)
- Voice-enabled access to business transactions using Enterprise Services, Remote Function Call (RFC), or BAPI
- Complete software lifecycle (administering, deploying, logging, and monitoring) for voice applications
- Detailed reporting through Business Information Warehouse (BI) (optional)
- Investment protection in ERP and SAP Business Suite

What Systems Do You Need?

The development system is the Visual Composer Voice Kit. The Voice Kit extends the capabilities of the Visual Composer for voice-enabled application modeling. It provides a complete set of modeling tools for back-end access, data manipulations, dialog elements, and dialog flow. The Voice Kit provides a complete and codeless implementation of a fully interactive voice application.

To use NetWeaver Voice, you need the following systems with activated usage types:

SAP NetWeaver Composition Environment with an activated voice modeling add-on.

This is the development environment you need to develop voice applications.

SAP NetWeaver 7.0 BI (optional)

The Visual Composer Voice Kit is integrated with BI to provide basic reporting. More sophisticated reporting functionality is optional.

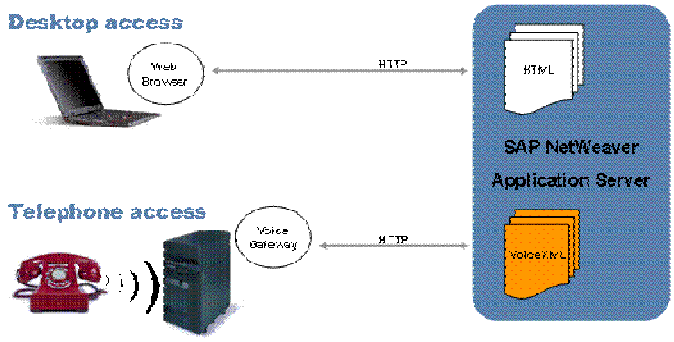
Third-party VoiceXML-compatible gateway (external system)
You render the output of the NetWeaver Voice application on a third-party VoiceXML compliant gateway.

What Technologies Are Involved?

SAP NetWeaver Voice is similar to a Web application, but based on these technologies:

- Telephone applications, also called Interactive Voice Response (IVR)
- Speech recognition and telephone keypads (touch-tones) for input
- Recorded speech and computer-generated speech, called text to speech (TTS) for output
- Voice-enabling access to business transactions using Enterprise Services, RFC, or BAPI

In common with Web applications is the Hyper Text Transfer Protocol (HTTP) used as the application communications protocol:



While the Hypertext Markup Language (HTML) is used to write Web pages, the Voice Extensible Markup Language (VoiceXML) is used to specify the interactive dialog between the caller and the system in a voice-enabled application.

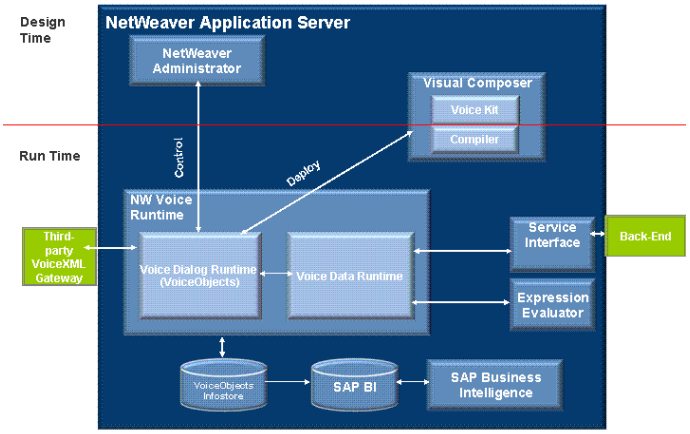
How Does NetWeaver Voice Fit into the Composition Environment Strategy?

NetWeaver Voice addresses the component and process modeling stages and consumes services and business objects. You can use NetWeaver Voice to create individual voice UI components, such as a dialog to collect a credit card number, and you can reuse these components in new business-process models.

How to Test Your Voice Applications?

- You can either use a VoiceXML gateway or SAP Text Application Gateway (STAG) to test your voice applications.

- Using VoiceXML gateway for testing allows you to test the voice application more close to a real-time scenario.
- Using STAG to test your voice applications is not considered as an efficient way of testing because it cannot recognize natural voice. It allows you to test your voice application by submitting inputs only in the form of text and not voice. However, STAG provides a quicker way of testing your voice applications as it is a runtime component of SAP NetWeaver Voice and does not require any configuring unlike VoiceXML gateway.



SAP Net Weaver Voice Architecture

The following table provides descriptions of the design-time components:

Design-time component	Description
NetWeaver Administrator	Allows you to perform various administrative tasks such as creating users, stopping and restarting voice applications and viewing the logs
Visual Composer Voice Kit	Allows you to model voice applications using the voice and non-voice entities. It also allows you to deploy your voice applications

The following table provides descriptions of the runtime components:

Runtime component	Description
Visual Composer Compiler	Validates a voice application
Voice Dialog Runtime	Provides access to runtime resources such as audio files and grammar files to the voice application at runtime
Voice Data Runtime	Provides access to the data at runtime. It interacts with the back-end system or the voice dialog runtime to provide the required data
Service Interface	Assists the voice data runtime component to access the back-end or Web Services data
Expression Evaluator	Evaluates expressions used in a voice application
VoiceObjects Infostore	Stores call data such as the duration of the call, number of callers and so on
SAP BI	Stores call data created in the VoiceObjects Infostore in a form that can be used by SAP Business Intelligence
SAP Business Intelligence	Generates BI reports based on the data stored in SAP BI

CONCLUSION

The potential for voice in apps is immense. An instance is including it in language learning apps to as an aid to handicapped users. In the future, wearable devices will fuel adoption of speech in apps. While it may be some time before speech will replace touch as primary input to smart phone apps, developers need to start considering whether and how to add speech control to their apps to stay competitive.

Reference

1. https://blogs.sap.com/2012/11/19/what-is-sap-netweaver/
2. http://searchsap.techtarget.com/definition/NetWeaver
3. https://en.wikipedia.org/wiki/SAP_NetWeaver
4. https://en.wikipedia.org/wiki/SAP_NetWeaver
5. https://www.guru99.com/what-is-netweaver.html

How to cite this article:

Jeevitha V and Kavitha E.S.2018, A Study About Voice Enabled Applications For Document Searching Service. *Int J Recent Sci Res.* 9(1), pp. 23349-23352. DOI: http://dx.doi.org/10.24327/ijrsr.2018.0901.1451

****