

THE NEW ERA OF BROWSING -VOICE BROWSING

Khushbu¹, Manika Kapoor², Ayesha Tafsir³

¹Student, CSE, Dronacharya College of Engineering, Gurgaon, India, khushi.bhardwaj1991@gmail.com

²Student, CSE, Dronacharya College of Engineering, Gurgaon, India, manikakapoor91@gmail.com

³Student, CSE, Dronacharya College of Engineering, Gurgaon, India, tafsirayasha@gmail.com

Correspondence Author: Khushbu¹Student, CSE, Dronacharya College of Engineering, Gurgaon, India
: khushi.bhardwaj1991@gmail.com

ABSTRACT:

The technology of voice browsing is rapidly evolving these days. It is because the use of cellphones is increasing at a very high rate, as compared to connected PCs. Listening and speaking are the natural modes of communication and information gathering. As a result we are now heading towards a more voice based approach of browsing rather than operating on textual mode. This paper concentrates on this new technique, voice browsing, which unites speech recognition and speech synthesis that can be very fruitful in the coming years.

1.INTRODUCTION:

A voice browser can simply be defined as an appliance or a gear which helps in interpreting a markup language(the markup language referred here is 'voice') and producing a voice output. It translates a given voice input into a voice output. It is the web browser which provide the users with an interactive voice user interface. It also provides the user with an interface to the PSTN(Public Switched Telephone Network). It is obvious from the first word of the name that the system deals with pages that specify voice dialogues, just as our visual web pages deals with HTML pages. But the question remains-how does a software system reciprocates to the user via

speech or voice browser? The software system procures its information from the internet. From a user's outlook, the goal is to provide to the devices which do not have full-browsers or even the screens to support them, a service which is similar to what the visual web browsers and the related technologies offer today.

2. STANDARDIZATION:

Standardization to voice browsing technique were given by:

The World Wide Web Consortium (W3C) develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential as a forum for information, commerce,

communication, and collective understanding. W3C which includes:

- 1 .Voice Browser Working Group
2. Speech Interface Framework

2.1 Voice Browser Working Group

It was established on 26 March 1999 and re-chartered through 31 January 2009. W3C voice browser working group made the speech interface framework possible . This framework allows developers to create speech enabled applications that are based on Web technologies.

The framework also provides developer with an environment that will be familiar to those

2.1(a) Aim:

The Aim of the W3C Working Group is to enable users to speak and listen to Web applications by making standard languages for developing Web-based speech applications. This Working Group concentrates on languages for capturing and producing speech and managing the conversation between user and computer system, while a related Group, the Multimodal Interaction Working Group, works on additional input modes including keyboard and mouse, ink and pen, etc.

2.1(b) W3C Recommendations: Its recommendations have been reviewed by w3c group Members, by software developers, and other interested parties, and are also endorsed by the Director as Web Standards.

2.2 Speech Interface Framework:

These framework includes:

2.2(a) VoiceXML: a language for creating audio dialogs that feature synthesized speech, digitized audio, recognition of spoken and DTMF key input, recording of spoken input,

who are familiar with Web development techniques. So, applications are written using parts of speech interface framework.

Thus speech applications are written in VoiceXML and are rendered through a Voice Browser. In much the same way as Web applications are written in html and run on a Web browser.

As per estimation, over 85% of Interactive Voice Response (IVR) applications for telephones (including mobile) use W3C's VoiceXML standard.

Voice Browser Working Group are coordinating their efforts to make the Web available on more devices and in more situations.

telephony, and mixed initiative conversations. Some of its versions are:

- VoiceXML 1.0: designed for creating audio dialogs.
- VoiceXML 2.0: uses form interpretation algorithm(FIA).
- VoiceXML 2.1: 8 additional elements.
- Voice XML 3.0: relationship between semantics and syntax.

2.2(b) Speech Recognition Grammar Specification (SRGS) 1.0: a document language that can be used by developers to specify the words and patterns of words to be listened for by a speech recognizer or other grammar processor.

2.2(c) Speech Synthesis Markup Language (SSML): a markup language for rendering a combination of prerecorded speech, synthetic speech, and music. Some of its versions are:

- Speech Synthesis Markup Language (SSML) 1.0
- Speech Synthesis Markup Language (SSML) 1.1

2.2(d) Semantic Interpretation (SISR): document format that represents annotations to grammar rules for extracting the semantic results from recognition.

Eg: Semantic Interpretation(SISR)1.0 version

2.2(e) Pronunciation Lexicon Specification (PLS) a representation of phonetic information for use in speech recognition and synthesis .

Eg: Pronunciation Lexicon Specification (PLS) 1.0 version

3. LEVELS OF VOICE BROWSING

In order to understand it better in its current form, voice browsing can be examined under three levels.

3.1 voice browsing

This is the first level in examining the technique of voice browsing. The simplest way of understanding the voice browser and the voice web would be to take the web itself into consideration. We know that people all over the world visit websites and get visual feedback. Now, the voice web contains voice sites where the feedback is delivered through dialogues. The most basic example would be calling up our cellphone operator portal, where speech recognition software provides the caller with a series of options like recharging your account, talking to the customer care executive or listen to the new offers. The technology of voice browsing has brought down the cost considerably, since earlier it would cost a rupee per minute for human operators to talk to customers while 10paise per minute for an automated call. Voice browsing also has its use in the corporate sector specially in banks and airlines.

3.2 voice browsing the world wide web.

This is the second level through which we can examine the voice browsing technique. We can browse the websites through voice which offer voice portals. In order to maintain customer loyalty these portals offer voice browsing of personal content. To have a claim in this sector or we can say market, AOL bought quack.com.

Almost everywhere the voice portal market is being expanded, be it Europe, the US or Asia.

3.3 The Voice Web

The level three is the voice web. Many companies have introduced forums for programmers for setting up voice sites so as to enhance the interest in voice browsing and speech recognition. The forums further become voice webs which sprout voice auction sites and voice based chat rooms.

4. Future Of Voice Technology

The speech technology is supposed to grow rapidly. The voice portal market is going to reach billions in just a few years. It is estimated by the kelsey group that voice browsing market will reach 6.5 billion dollars, while OVUM estimates a world market of 26 billion dollars. Anyone may guess the actual growth of the industry of voice technology due to variations in these figures. It is very difficult to navigate on a WAP to scroll through many lists. Hands-free interaction enables us to develop an easy communication between the user and the system.

voice browsing can be used to access three kinds of information:

(a)**Business:** information like automated telephone ordering services, support desks, order tracking, airline arrival and departure information services, cinema and theatre booking services, home banking services,

etc can be retrieved using voice browsing very easily.

(b)**public**: voice browser can be used to access services like local , national and international news alongwith community information such as weather forecasting, traffic conditions, school closure and events. it can also be used to gather information on national and international stock market information and also business and e-commerce transactions.

(c) **personal use** : It is used in accessing personal information like voice mails, personal horoscope ,personal newsletter, calendars, address and telephone lists etc.

In future it is expected that voice browsing will become visual i.e MULTI MODAL. But greatest achievement would be when voice browsing is integrated with all types of operating system .This success would surly make voice browsing available to each and every application.

5. Conclusion:

In order to make technology more familiar to the user its access should be made more easier. As we know that visual internet access experiences various limitations such as people who are physically handicapped (specially blind users) cannot use keypads or touch screens for giving instructions. Above all these limitations today's generation demands to use internet independent of PC's and also hands free access to it. For this VOICE BROWSING is an intelligent idea. This allows user to access web even in situations like driving etc where user operate web just by listening and speaking rather then typing. Thus at last we conclude that Voice browsing provides a natural way of accessing webs. Now it is up to the developers to take up some inventory measures in order to bring this technology to us in a more colourful way.