

Towards Knowledge Based Society

Abdulaziz S. Almazyad and Farooqui N.K.

Abstract— Knowledge Propagation is the challenge. Better society requires best knowledge management practices and use of latest tools and technologies. KM is dealing with the problem of different languages in big countries. Also Knowledge is present in different forms. It may be knowledge encoded in laws, decrees, standards, policies, rules, regulations and orders. All these forms of knowledge must be combined together so that it is readily available for general public in their own language. In addition there should be a proper knowledge propagation mechanism. Knowledge propagation becomes difficult if it is in the multiple languages. Non availability of computing resources makes it even more worst. Recently mobile phones have become widely used in third world country. Mobile phone has amazing features like continuous connectivity, location detection, facility of voice and data processing. This attracts KM researchers to consider Mobile as the medium of knowledge propagation. This paper proposes a framework for knowledge propagation through voiceXML in different languages using Mobile phone as the delivery medium. Voice Interface requires minimum resources and works well even in the remote regions of the country. This is what makes this idea effective in the third world countries where low cost mobile penetration is high.

Index Terms—Barriers for Knowledge Management, Knowledge Management, Knowledge Society, VoiceXML.

I. INTRODUCTION

A country like India, has 31 states with total population of 1 Billion plus. With 625 [5] districts India has 29 languages having more than 1 million native speakers in more than 1 Billion peoples [6], 61 % education [7], India has **7,84,84849** number of mobiles provided by different operators [approx figure]. The expected figures for mobile are attractive. India has only 3.7 % internet penetration, so knowledge propagation through internet is a difficult way.

Manuscript received May 9, 2009
A.S. Almazyad is the Dean of College of Computer Engineering and Sciences, King Saudi University, Riyadh, Saudi Arabia (phone:966-1-4677100; fax:966-1-4675630; email:mazyad@ccis.edu.sa)
Farooqui N.K. is with the College of Computer Engineering and Sciences, King Saudi University, Riyadh, Saudi Arabia
(email:farooquinajam@gmail.com)

Knowledge propagation is a Challenge in a country like India with 1 Billion Plus population. Following table shows the states with their population. With such huge population knowledge propagation is a difficult task.

Table I: State in India [9]

SrNo.	Name	Population
01	Uttar Pradesh	16,60,52,859
02	Maharashtra	9,67,52,247
03	Bihar	8,28,78,796
04	West Bengal	8,02,21,171
05	Andhra Pradesh	7,57,27,541
06	Tamil Nadu	6,21,10,839
07	Madhya Pradesh	6,03,85,118
08	Rajasthan	5,64,73,122
09	Karnataka	5,27,33,958
10	Gujarat	5,05,96,992
11	Orissa	3,67,06,920
12	Kerala	3,18,38,619
13	Jharkhand	2,69,09,428
14	Assam	2,66,38,407
15	Punjab	2,42,89,296
16	Haryana	2,10,82,989
17	Chhattisgarh	2,07,95,956
18	Jammu & Kashmir	1,00,69,917
19	Uttaranchal	84,79,562
20	Himachal Pradesh	60,77,248
21	Tripura	31,91,168
22	Manipur	23,88,634
23	Meghalaya	23,06,069
24	Nagaland	19,88,636
25	Goa	13,43,998
26	Arunachal Pradesh	10,91,117
27	Mizoram	8,91,058
28	Sikkim	5,40,493

Every state has its own local laws in addition to the central government laws. States have different tax structure; they have different policies for entertainment, education, law and order, construction, agriculture, food, traffic rules and regulation. Furthermore every state is liable to implement all the central government laws in all these sectors. Furthermore every state has its own languages for example Maharashtra has the primary official language as 'Marathi' while some states have their language as 'Hindi', other states have languages like Gujrati, Kashmiri, Oriya, Tamil, Kannad, Telgu or Assami. Education rate and technology penetration in all these states are very low. Hence it requires simple solution without the need of any high end gadgets.

A. Telecommunication Infrastructure

India has 91% growth rate in mobile industry. Companies like AirTel, BSNL, VodaFone and Idea are the major contributors to this enormous growth in the telecom sector. TRAI (Telecom regulatory authority of India) has played the major role in making this sector more competitive. As a result the growth in the sector is enormous with new facilities and cheapest services rates in the world. As the services are cheapest and they are affordable for the local economy, mobile penetration in last 5 years is very high. Almost every person and average income household has mobile phones. Although as India being a third world country has very low penetration of high end mobile phones. Most of the peoples have low end mobile phones with basic facility of voice and text communication.

B. Information & Communication Technology in India

India is considered to be a software superpower. Indian software exports will cross 40 billion USD. India is the considered to be the favorite destination for outsourcing. The reasons are large English speaking population, large population with domain specific knowledge, cheap labor cost, time zone advantage, IT resources (Network specialists, hardware, software's, software developers) are easily available and favorable government rules and regulations for IT sector (Currently IT Sector is exempted from Income Tax & Sales Tax).

NIC (National Informatics Center) is the central body in India which runs several government websites including central and state government. It also runs websites for different educational organizations like state educational boards which conduct examinations for 10th and 12th Class in India. NIC also runs websites for displaying result for various exams. NIC handles the websites for major corporate entities belonging to various government departments like coal ministry, steel ministry, home ministry and education ministry.

Indian foreign ministry now accepts online applications for passport and related services in some selected cities. It is now mandatory for all major organization runs by central and state government in power, food, administration, law, public services to have websites and they should have all related information available to the citizens, particularly the rules, regulations and bidding process. For example Nuclear power Corporation of India, state electricity boards, food corporations in state, Steel authority of India, each and every ministry, major government offices and tourism offices. Most colleges in the cities now have their own websites. Almost all the higher education colleges and universities admit their students through website. Their admission process is either partial web based or full web based. Engineering and medical college admission processes are almost through the web. [10, 11].

All India fulfils all the requirements suggested by Saad Haj Bakry [12] in his STOPE view. The STOPE stands for Strategy, Technology, Organizations, Peoples and Environment. Indian Government and organizations have positive strategy towards the Information technology industry. Currently government does not impose any taxes on the business related to software as being a new and emerging business. Being a superpower in the Information Technology industry technology is easily available to Indian organization for Knowledge Management. Organizations have positive side towards the implementation of Information technology in their area. For example Indian Railways has the largest setup for their website www.irctc.nic.in which is highly loaded website for online railway reservations. Peoples working in government, semi government and private sector are highly educated with English as their primary language. With all these inputs the Environment in India for being a knowledge society is more prominent.

C. Knowledge Society

With the innovation of mobile communication, internet penetration, availability of higher education and every knowledge seeking young community India is going towards the knowledge enabled society. Students, universities, colleges and Indian young community are really the knowledge seeker. They want to acquire new knowledge, work with innovations in each and every field.

Following table shows queries hit through Google from India. It clearly shows that India has greater knowledge seekers in the world.

Table II: Google table for Queries hit from India [23]

astest Rising	Most Popular
youtube	orkut
orkut	gmail
katrina kaif	yahoo
cricket	google
irctc	youtube
facebook	yahoomail
genelia d'souza	indian railways
beijing 2008 olympic games	rediff
sixth pay commission	cricket
ipl	katrina kaif

op 'how to' searches	op searches on Mobile
how to reduce weight	orkut
how to kiss	yahoo
how to earn money	waptrick
how to get pregnant	gmail
how to learn English	games
how to gain weight	Katrina kaif
how to play guitar	rediffmail
how to create a website	yahoomail
how to impress a girl	hamitha
how to tie a tie	Google

India with 1 billion plus population has different industries, banks, organizations working in different sectors. Every organization is run by some rules and regulations. Peoples working in those organizations, taking services from those organizations want to know those rules and regulations and policies in order to take maximum advantages out of that. Hence India is emerging as knowledge based society. With the high penetration of internet, availability of broadband and presence of cheap mobile communication India is a true knowledge society today. Particularly specialized peoples working in specific job like journalists require on the move mobile knowledge management facility [1]. The problems before the Knowledge Management experts are to propagate the knowledge to them in some form which is understandable to common man. As they are not educated they are not able to browse through Internet, so mobile communication through voice is the best medium to do so.

II. THE PROBLEM

Every industry and organization coming under the Indian eGovernance paradigm has its own data, rules and regulations, policies, laws, way of working and financial power/implications. Every organization has to manage data

related to these things in proper way and propagate it to its employees and customers. The organizations working for development and education have to manage larger knowledge bases. For example there are organizations like 'Krushvi Vikas Kendra' (Agricultural Development Center) in each district. It is a kind of knowledge base for all kinds of agricultural activities and related industries (like animal husbandry). They provide expert advice to the farmers on different problems related to agriculture. Like they provide various problems related to allied industries like animal husbandry. Another example is the customer service centers of organizations like banks, railways, foreign office, power corporations, education departments, food corporation...etc. There in the customer service center of these organizations the customer executives have to face lot of questions on different problems. The executive may or may not have knowledge about the question being asked. A knowledgebase helps them in order to properly answer the queries of those customers. The queries about railways and bank loans may not be common so customer support executive always require the help of a knowledgebase. The presence of updated knowledgebase makes the answer to every question possible.

An uneducated/semi educated person who wants to know about the problem he has on his farm/field, cannot visit the help center for many reasons. Even if he is uneducated, he uses his mobile phone in English or in regional language for his day to day communication. Mobile phones in regional languages are available in India from various vendors like NOKIA, SAMSUNG, and LG. the Knowledge can be propagated through this small magical piece in the form of voice. There can be many such peoples seeking information about various eGovernance related problems. The government works like problems of ration cards, National ID cards, papers related to agricultural and non agricultural land, sale and purchase of properties, sale and purchase of 2/4 wheeler vehicles ...etc require awareness about the local rules and regulations. A Person who wants to get a two- wheeler or four-wheeler license required to get a license from Regional Transport officer. He wants information regarding the procedure. The knowledge of the procedure is required by both educated and uneducated persons. Educated person may visit website or can get information by some other means, but uneducated person cannot. Also if the person lives in the remote region, mobile is the best medium to propagate this knowledge. Village people use most basic handset. They have minimum functions available. Mostly the communication is voice and text based, the buttons are available for basic alphanumeric operations. There are various tools developed for knowledge management like RepTool, gIBIS [3, 4]. Some field specific tools also exist like NewsMate [1]. Newsmate was developed for journalists to manage their knowledge in multiple languages.

III. PROPOSED ARCHITECTURE

Fig 1 shows the proposed architecture to solve this multi faced problem. The figure shows that a query can be submitted through mobile devices (may be low end mobile or any mobile computing device). The query can be submitted for particular information forwarded to a particular number. Through mobile network it will convert the multilingual voice query to text based query. This multilingual text query should then be translated to English language. This English query should be fired to an already existing knowledge base. The knowledge base then responds with the multiple results. The best result must be selected with the help of some algorithm. Once the best result is selected then the results should be translated from English language to the native language of the mobile holder. Once the voice answer is ready, it can be fired to the mobile device with the help of the voice Browser.

A. Components in the proposed architecture

Component 1: Voice to Text Converter / Vice Versa

This is the software working on the mobile operator server, where mobile is directly connected after dialing a particular number to get the information about his/her query. Once the person dialed a number the system should activate this voice to text converter. Once this converter is activated, the person will ask his query. There should be some directive to put his/her query in a well formatted manner. This will make searching the knowledge base easy and reliable. There are many text-voice converters available. At & T has a demo available for converting text into voice [22].

Component 2: Multilingual converter from any language to the English language

Once the query is translated into the text, you have to activate the next component. This component will convert text query from native language to the English language. Once the text query is converted from native language to English language, it becomes searchable in the knowledge base. There are various converters available for this job. In case of implementation in India, where you have 28 different languages being spoken among 1 billion plus peoples, you will need a master translator to convert all 28 languages to English. Lot of work is already done in this area. Following are some of the experimental translation system already developed.

MEANING [19] (Developing Multilingual Web-scale technologies) – aiming at developing technologies to search by concepts and not by words; contributing in this was to sense disambiguation.

ONELOOK [20] - “search engine for words and phrases”, based on WorldNet. You enter the word and get the links to the dictionaries where you can find it.

TERMINAUTA [21] – a terminological metasearch engine developed for translation resources online, where you can combine different languages.

MUCHMORE [22] aiming at developing technologies that will result in a prototype system for cross-lingual information organization and access for the medical domain.

Component 3: Search in the given Knowledge base

The converted English language query can be fired to the stored knowledge base. This knowledge base must have all the information for the said field. For example the bank, agriculture or foreign office. If someone asks how to get a passport query, the knowledge base must have complete information about any query to any department. In case a query doesn't match any question, immediately it must be recorded and the data must be filled in to answer query next time.

Once you fire the query to the knowledge base, there will be multiple answers. This is because the searching word may match to multiple files in the knowledge base. Here it will be difficult to select the file to be forwarded to mobile client. Now, question is how to select the best file to be converted and ultimately play on the mobile handset. Here we have to use another component called answer selector, this component should decide the best and most effective answer to be played on the mobile phone.

Component 4: Answer Selector

This component will select the best answer to be sent to the recipient's mobile handset. Here we can consider various algorithms to select the best answer.

Component 5: Voice Browser

This is the software which will be used on the mobile client to speak the text. VoiceXML (VXML) is the W3C's standard XML format for specifying interactive voice dialogues between a human and a computer. VoiceXML documents are interpreted by a *voice browser* [13]. There are various experimental systems developed depending on the VoiceXML. VoiceBox is a system for talking books [14]. This paper presents a framework through which you can access the written book through interactive telephone system. Furthermore the paper also discusses the navigation control in the hands of the user. "Mobile refining of web information using voice interface" is the exploratory study conducted by Robert G. Capra III & Manuel A. Pérez-Quiñones [15]. This

paper discusses the experimental system for refining of information on the web with the help of telephones. The making of semantic web browser is discussed by Deniis Quan and kargar[16]. This paper discusses the RDF (*Resource Description Framework*) to be interpreted by the semantic web browser. Volker Derballa & Key Pousttchi discusses extending of knowledge management to the workplace with mobile technology [17]. Another interesting paper discusses the implementation of Audio wikis through mobile collaboration [18]. They discuss how wikis in the audio format can be made, navigated and run on mobile handsets for fast and extensive knowledge management.

IX. CONCLUSION

Use of mobile technology for knowledge propagation is the need of the hour. As mobile services are cheap and widely available in the country like India, knowledge propagation through mobile is mandatory. With the presence of Telecommunication and Information Technology infrastructure India presents a best case of implementing such framework. Knowledge propagation puts two problems in front of KM researchers. First is to deal with multiple languages, second is the communication with audio rather than textual information. The framework has two components to deal with these problems, one is audio to text converter which converts and audio message into regional text and then to the English language query. The second component is the voice browser which can speak out the answer. This framework combines different/ emerging technologies from different areas of research. If implemented the architecture brings the Knowledge to every household in their own language.

REFERENCES

- [1] Henrik Fagrell, Kerstin Forsberg, Johan Sanneblad, "FieldWise: a mobile knowledge management architecture", Proceedings of the 2000 ACM conference on Computer supported cooperative work, Philadelphia, PennsylvaniaPages: 211 - 220 (2000).
- [2]. Jordan, B., R. Goldman and A. Eichler (1998) "A Technology for Supporting Knowledge Work: The RepTool," In Information Technology for Knowledge management, U.M. Borghoff and R. Pareschi (eds.), pp. 79-96. Berlin: Springer-Verlag.
- [3]. Conklin, J. and M.L. Begeman (1988) "gIBIS: a hypertext tool for exploratory policy discussion," ACM Transactions on Office Information Systems, 6 (4), pp.303-331.
- [4]. Jordan, B., R. Goldman and A. Eichler (1998) "A Technology for Supporting Knowledge Work: The RepTool," In Information Technology for Knowledge management, U.M. Borghoff and R. Pareschi (eds.), pp. 79-96. Berlin: Springer-Verlag.
- [5] Districts of India [online] available at <http://districts.nic.in/dstats.asp>.
- [6] "List of Indian Languages & their Speakers", [online], available at http://en.wikipedia.org/wiki/List_of_Indian_languages_by_total_speakers
- [7] "Literacy in India", [online] available at http://en.wikipedia.org/wiki/Literacy_in_India.
- [8] "Internet Usage Stats and Telecommunications Market Report", [online], available at <http://www.internetworldstats.com/asia/in.htm>.
- [9] "List of Indian States with their populations", [online] available at http://www.haryana-online.com/states_of_india.htm.
- [10] "Top Engineering Colleges in India", [online] available at <http://www.hindustanlink.com/careertex/top-institutes-india/top-tenengineering-colleges-india.htm>.
- [11] "Top Medical Colleges in India", [online], available at <http://www.hindustanlink.com/careertex/top-institutes-india/top-ten-medical-colleges-india.htm>.
- [12] Saad Haj Bakry Home Page, [online], available at <http://faculty.ksu.edu.sa/shb/default.aspx>.
- [13] "VoiceXML, W3C's standard XML format for specifying interactive voice dialogues between a human and a computer", [online] at <http://en.wikipedia.org/wiki/VoiceXML>.
- [14] Anchal Jain & Gopal Gupta", VoxBox: A System for automatic generation of interactive books." **Assets '06**: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility.
- [15] Robert G. Capra & Manuel A. Perez- Quinones, "Mobile refining of web information using voice interface". Proceedings of 2005 Latin American conference on Human computer interaction, October 2005.
- [16] D.A. Quan & R. Karger, "How to make a semantic web browser" (2004), Proceedings of the 13th international conference on World Wide Web, New York, NY, USA, May 2004.
- [17] Volker Derballa, Key Pousttchi (2004), "Extending Knowledge Management to Mobile Workplaces", Proceedings of the 6th international conference on Electronic commerce, Delft, Netherlands.
- [18] Lei Wang, Paul Roe, Binh Pham, Dian Tjondronegoro, "An audio wiki supporting mobile collaboration". Proceedings of the 2008 ACM symposium on Applied computing, March 2008.
- [19] Baeza-Yates, Ricardo & Berthier Ribeiro-Neto. (1999), *Modern Information Retrieval*, New York: Addison-Wesley Publishing Company, USA.
- [20] Baker, Tom, (ed.) (s)d) *Managing a Vocabulary for the Semantic Web -- Best Practice*. Draft under development for W3C Semantic Web Best Practices and Deployment Working Group. [Online], available at http://esw.w3.org/topic/VocabManagementNote_
- [21] Francesca Bertagna1, Alessandro Lenci2, Monica Monachini1, Nicoletta Calzolari1, "Content Interoperability of Lexical Resources: Open Problems and "MILE" Perspectives". In *Proceedings of LREC 2004 Fourth International* (2004).
- [22] Online Demo for AT&T text to speech s technology. [online], available at <http://www.research.att.com/~ttsweb/tts/demo.php>.
- [23] Google zeitgeist (Google's product showing summary of queries hit from different countries.), [online], available at <http://www.google.com/intl/en/press/zeitgeist2008/#top>.

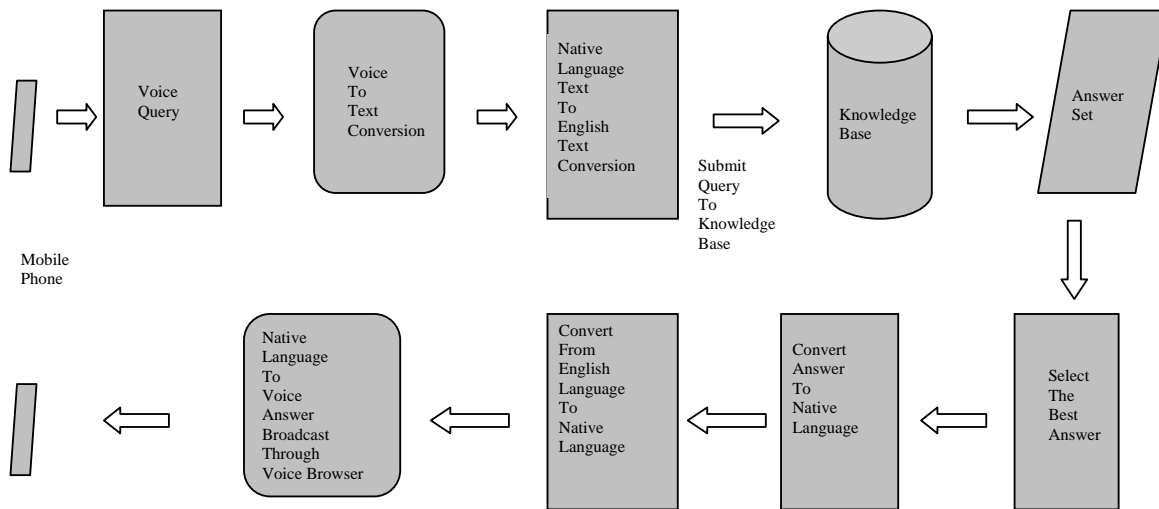


Fig 1: Architecture of Knowledge Based Society