WWTW : The World Wide Telecom Web

Arun Kumar Nitendra Rajput Dipanjan Chakraborty Sheetal K. Agarwal Amit A. Nanavati
IBM India Research Laboratory
4, Block - C, Institutional Area,
Vasant Kunj, New Delhi - 110070, India.
{kkarun,rnitendra,cdipanjan,sheetaga,namit}@in.ibm.com

ABSTRACT

The World Wide Web (WWW) enabled quick and easy information dissemination and brought about fundamental changes to various aspects of our lives. However, a very large number of people, mostly in developing regions, are still untouched by this revolution. Compared to PCs, the primary access mechanism to WWW, mobile phones have made a phenomenal penetration into this population segment. Low cost of ownership, the simple user interface consisting of a small keyboard, limited menu and voice-based access contribute to the success of mobile phones with the less literate. However, apart from basic voice communication, these people are not being able to exploit the benefits of information and services available to WWW users.

In this paper, we present the World Wide Telecom Web (WWTW) — our vision of a voice-driven ecosystem parallel to that of the WWW. WWTW is a network of interconnected voice *sites* that are voice driven applications created by users and hosted in the network. It has the potential to enable the underprivileged population to become a part of the next generation converged networked world. We present a whole gamut of existing technology enablers for our vision as well as present research directions and open challenges that need to be solved to not only realize a WWTW but also to enable the two Webs to cross leverage each other.

Categories and Subject Descriptors

C.2.1 [**Computer Communication Networks**]: Network Architecture and Design

General Terms

Design, Human Factors

Keywords

World Wide Web, Developing Regions, Telecommunications, VoiceXML, Conversational Systems

NSDR'07, August 27, 2007, Kyoto, Japan.

1. MOTIVATION

The Internet is one of the most significant technologies that have changed our daily lives in the recent past. This has been made possible through the numerous information sources and applications available over the World Wide Web (WWW). However the impact of the WWW is still not at the level of some basic facilities such as the railroad, electricity or even communications. There is a significant percentage of population that is still untouched by this revolution and are either unaware of or are unable to catch the momentum [13].

Even today, barely 17% of the world's population has access to the Internet [6]. There are a variety of reasons that act as a hindrance for this technology to impact the remaining 83% section of the human population. Firstly, 53% of the world population lives below USD 2 per day [12] – so they cannot afford a PC or high end phones and hence cannot access the Internet. Secondly, a significant portion of the remaining 30% are illiterate and semi-literate people [14] who do not know how to operate a computer. Thirdly, most of the information and applications available on the Internet is hardly relevant to this section of the society. However, for this technology to become a *commodity* such as a road or electricity, a significant shift in paradigm is needed. Incremental improvements in terms of the sophisticated services, advanced user interfaces, easier application authoring techniques for the WWW do not appear to be helping in providing a shift of such a large scale.

Interestingly, the telecommunication network does not face some of the challenges of the Internet world – from an acceptance perspective. The cost of a phone is significantly lower than a PC and, the learning required to operate a phone is negligible as compared to a PC, especially when the phone is used as a device to communicate in free speech. Thus telecommunications have become a commodity for the common man and are a step closer towards achieving that status for the underprivileged as well.

In this paper, we present our vision of a World Wide Telecom Web (WWTW) that, we believe, has the potential to deliver to underprivileged, what WWW delivers to IT literate users today. Specifically, WWTW

- enables the underprivileged to create, host and share information and services produced by themselves,
- provides simple and affordable access mechanisms to let the masses exploit IT services and applications that are currently available to WWW users, and,
- provides a cost effective ecosystem that enables users to *create* and *sustain* a community parallel to the WWW.

Our vision aims to exploit the features of the telecommunication channel for creating an information and services ecosystem for the masses in developing regions, as it can lead us to a completely different and a much powerful digital superhighway. The vision requires novel technologies ranging from user interfaces to infrastructure to be developed in order to enable such a paradigm shift.

2. A SCENARIO

Ram is an electrician operating in the Shivadaspur area in the Varanasi town. His expertise lies in fixing household electrical problems of all kinds, except air coolers and air-conditioners. He cannot afford to have a shop of his own and his business depends on customers who know him by word-of-mouth. Recently, Ram bought a mobile phone, and started advertising his services in the Shivadaspur Yellow Pages. Since then, the business has started picking up. However, many a times, while on a home call on duty, he is unable to accept calls, and this often results in losing new customers and upsetting old ones. One day, he finds out about a Create-your-virtual-shop service offered by his Telecom operator, and decides to sign up.

He calls up the advertised phone number and creates his virtual shop as a VoiceSite *[7] in a matter of minutes by talking to this voice driven system. He also specifies reference information about previous customers and links to their phone numbers. Now the customers trying to reach Ram land up at this virtual shop and schedule an appointment with him while he is serving other customers. In addition, he adds links to the virtual shops of his friends who can take up the job in the event that he is unavailable at the time specified by the customer.*

Seeing Ram's increasing customer base, the electrical shop owner in his area requests Ram to include a link to the electrical shop's tele-store (another virtual shop), where customers can place their orders which will be home delivered by the store. Customers can pay through their bank account or through one of the credit cards that have a tie up with store's telecom provider. The payment happens safely through a voice driven interaction with the bank's VoiceSite *during the phone call, much in the same way as online transactions happen on the Web. This adds another customer facing channel for the electrical shop and adds to the services offered by Ram. Ram gets a percentage of the profits for customers reaching through his virtual shop and thus both Ram and the local store thrive with the use of Virtual Shop service.*

WWTW enables many such services that can be created to deliver IT services to the underprivileged in developing regions.

3. THE WORLD WIDE TELECOM WEB

3.1 Definitions

We define the World Wide Telecom Web ("WWTW", "TelecomWeb" or simply "T-Web") as an information and services space in which the items of interest, referred to as *VoiceSites* [7], are identified by global identifiers called *VoiNumbers* and maybe interconnected through *VoiLinks*.

A *VoiNumber* is a virtual phone number that either maps onto a physical phone number or to other uniform resource identifiers such as a SIP URI [3].

A *VoiceSite* is a voice driven application that consists of one or more voice pages (e.g. VoiceXML files) that are hosted in the telecom infrastructure. *VoiceSites* are accessed by calling up the associated *VoiNumber* and interacting with its underlying application flow through a telephony interface.

A *VoiLink* is a link from one *VoiceSite* to another through which a caller interacting with the source *VoiceSite* can be transferred to the target *VoiceSite* in the context of the *VoiceSite* application.

The WWTW, therefore, could be visualized as a system that operates over the Telecom infrastructure and parallels can be drawn with the World Wide Web [1] that runs on the Internet infrastructure. The WWTW model can be represented as a graph G = (V, E) such that: $v \in V$ is a *VoiceSite*, and $e \in E$ is defined as an edge from v_1 to v_2 if there exists a *VoiLink* from v_1 to v_2, where $v_1, v_2 \in V$.

VoiceSites, *VoiNumbers* and *VoiLinks* form the basic building blocks of World Wide Telecom Web.

3.2 The Concept

As shown in Figure 1, WWTW comprises of a web of interconnected *VoiceSites* each of which represents a voice driven application. This basic concept has tremendous implications for the underprivileged people in developing countries. It has been observed that IT systems with a voice-based feedback have much more appeal for the illiterate and semi-literate population of these regions [11] as compared to GUI-based systems such as Kiosks [1]. Such kiosks require computer literacy, do not have a high density in villages and the information content is mostly global, which is not relevant to the target population.

VoiceSites can be created through a simple voice driven interface over a phone call [7] and therefore, provide the ability to individual subscribers to offer their own voice-based applications. This enables the underprivileged users to have an online identity and personification through which others can reach them and interact with them. This is currently taken for granted by the IT savvy users in the form of email and personal websites. *VoiceSites* can be augmented to connect to services in the converged network such as Web Services available in the IT infrastructure and services in the 3G/IMS infrastructure [7]. This facilitates creation of non-trivial applications that exploit services residing in the back-end.

Enabling search and browse functionalities for the WWTW would help to grow this ecosystem. Integrating such features into WWW would further help the underprivileged users to reach out to new markets for their services.

3.3 Relation with WWW

While the similarities of the T-Web with the WWW abound, we would like to highlight some of the key characteristics of T-Web, which are somewhat different from those of the WWW.

- **Difference in Relation with Site Owner:** We believe that a *VoiceSite*, one of the key artifacts of T-Web, acts as a proxy for the owner in more ways than a website acts as a proxy in WWW. When one surfs a website, one doesn't typically expect to talk to a person or representative. Whereas, the owner of a *VoiceSite* can receive an incoming call personally or let

[1] http://en.wikipedia.org/wiki/World_Wide_Web

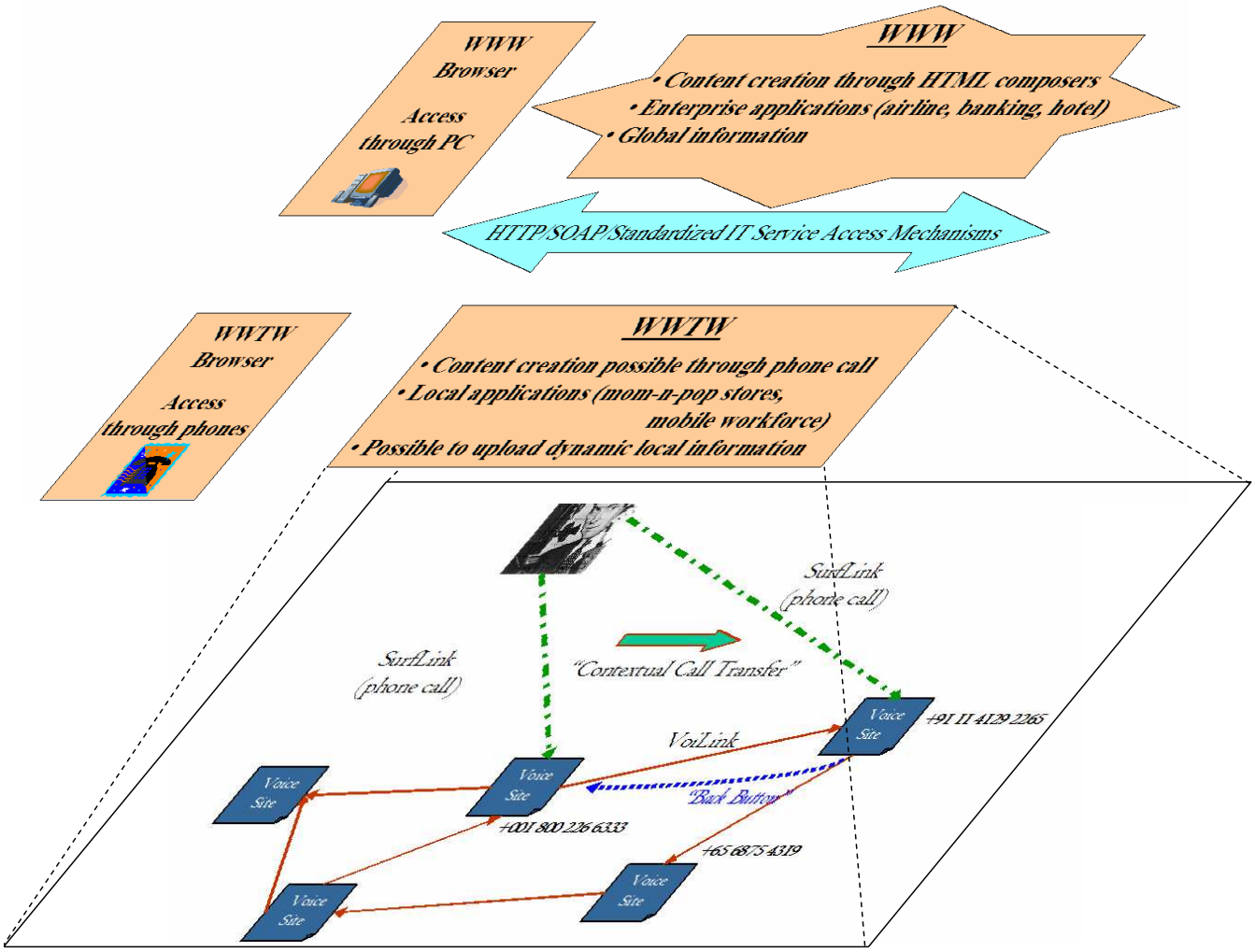

Figure 1: The WWTW ecosystem.

the caller interact with her *VoiceSite*. This makes the *VoiceSite* owner intrinsically connected to her *VoiceSite*. This could lead *VoiceSites* to have a profile comprising location, mood and several other real-time attributes that can appropriately represent the status of the person in real-life.

- **Structural Difference:** The T-Web has a different graph model than that of the WWW graph. The WWW graph contains single type of vertices (primarily websites) and single type of directed edges (hyperlinks) [9]. The T-web graph contains two types of vertices: *VoiceSites* and Surfers (both are represented by their *VoiNumbers*), and two types of directed edges: *VoiLinks* (which have a source and target *VoiceSite* at each end) and edges extracted from the telecom call-graph (call them *SurfLinks*), which have a Surfer as the source and a *VoiceSite* as the target. This is likely to have major implications on the evolution of the T-Web structure and on the behaviour of the users (surfers).

What sets WWTW apart from WWW is the simplicity with which an ordinary phone subscriber can join the digital information revolution. This enables a significantly larger fraction of the human population to benefit from existing and envisioned services than what was made possible by WWW.

4. TECHNOLOGY ENABLERS FOR WWTW

In order for the WWTW to proliferate to the level that we envision, technologies that enable easy *creation* of *VoiceSites*, easy information retrieval and seamless integration with back-end services, are extremely critical. In this section, we describe the technology enablers that fulfill some of the these requirements.

4.1 Creation

The ease of creating *VoiceSites* is essential for our target population. To that effect, we have developed a voice driven generator of voice-based applications, called VoiGen [7]. Creating a *VoiceSite* with VoiGen is as simple as calling the system over phone and navigating through the custom options offered. VoiGen has two intertwined components – a user interface generator and an application composition system. It makes use of existing components (reusable dialogs as well as IT components such as databases, web services etc.) to compose custom applications. By virtue of

having a voice driven interface, the services get exposed to all telephony devices.

4.2 Browsing

As individuals and businesses create their *VoiceSites*, they would need to be linked together to get most benefit out of these deployments. We have developed *Hyperspeech Transfer Protocol (HSTP)* [2] to enable cross organizational transactions and navigation in speech. It enables creation of *VoiLinks* to link *VoiceSites* to each other thus resulting into a web of *VoiceSites*. A browser designed for WWTW could then allow a caller on a *VoiceSite* to navigate this web of *VoiceSites* reachable from the current *VoiceSite*. Navigational operations such as forward and backward during an interaction with a *VoiceSite* are supported by *HSTP*. It also enables secure telephonic transactions across multiple *VoiceSites* that may belong to different autonomous organizations. This makes it possible to enable payment systems such as credit cards, online banking to be seamlessly integrated into *VoiceSites* – further advancing the state of current tele-shopping and tele-banking applications.

4.3 Search

With wide scale creation and deployment of *VoiceSites*, effective searching of information available through these would be a key challenge. Searching in WWTW could be split up into search in information space and search in services space. Search in the former would use the meta-information of a *VoiceSite* in addition to the audio prompts and speech recognition grammars. Since primary content is voice, searching the audio data becomes an essential rather than optional feature for enabling enhanced search. Audio search could be made easier by restricting the search to within context of the meta-information.

The search in services space can be more structured, based upon the meta information that the services offer in their interfaces. For enabling search in services space, we have developed a matchmaking service [4] that goes beyond traditional parametric matchmaking and also takes into account real-time information. Such a component enables searching of services and service providers not only on the basis of their functional and non-functional parameters but their availability and presence information as well.

4.4 Integration

VoiceSites provide a window to several categories of services in the converged networked world. *VoiceSites* can readily use existing IP channels to harness information and services available in WWW. For example, they can easily integrate with local applications (like databases) as well as remote applications exposed through currently evolving standards such as web services. *VoiceSites* can also exploit data and services in the 3G IMS infrastructure. For example, *VoiceSites* and services available through them can be configured to adapt based upon presence and other information available from appropriate IMS components. More information on how this integration can be enabled is provided in [7]. WWW on the other hand can harness the information and services residing in WWTW using appropriate integration tools. For example, creation of a *VoiceSite* can lead to adding a profile in a web-based social networking site, thus potentially leading to a significant representation from developing regions.

5. APPLICATIONS POSSIBILITIES

Services in WWW took off and continue to thrive because WWW presents an open medium for all (IT literate) users to contribute. It contains globally relevant information as well as locally relevant information. In the context of underprivileged users there are plenty of efforts for delivering global information such as eChoupal [8] and Microsoft's Saksham[2]. However, what is missing for this set of population is a mechanism to generate and sustain an ecosystem of communities that can generate locally relevant content and services such as in eSAGU[3], Manobi[4] and aAQUA[5]. WWTW aims to fill that gap. For example, simple creation and deployment of applications through *VoiceSites* enables people in villages and small towns to build and contribute to local information portals.

Further, *VoiceSites* when linked together, open up numerous possibilities in the telephony world much as the hypertext did in the Internet world. *HSTP* enables voice-based t-commerce (tele-commerce) solutions similar in theme to e-commerce on the web today. For example, as suggested in Section 2, local micro-businesses can open up their shopping catalogues on their *VoiceSites* and enable automatic credit card payment by linking up with a credit card company voice application through *HSTP*. Further, efficient and low cost telebanking solutions can be built on top of this platform that would help advance the efforts of numerous micro-credit and micro-finance institutions existing today (such as Grameen Bank [15] and SKS Microfinance[6]).

WWTW not only enables the underprivileged to exploit the currently available back-end services but also prepares them to leverage the services being planned for next generation converged networks of tomorrow [7]. For example, mobile workforce (people who typically have to move from one place to another for their daily jobs) such as plumbers and electricians, can integrate their *VoiceSites* with location and presence [3] services residing in the telecom infrastructure to customize their *VoiceSites* (e.g. If I am in my shop, then *VoiceSite* says "I am working").

WWTW also enables users in WWW to reach out to users in WWW and vice versa. This increased reachability could help the underprivileged to create new customer bases and eventually move out of poverty. The WWW users, at the same time, reach out to a very large untapped market and a population largely unexplored. WWTW thus opens up new avenues for the research community to enable several other applications in the areas of Healthcare, Education, Finance - some of which are well established in WWW amongst the IT literate community.

6. KEY RESEARCH CHALLENGES

WWTW, as we envision it, presents plenty of research challenges for various disciplines. We attempt to summarize some of the important ones in this section.

- **Usability:**

  Given our target population it is important that the user interface offered to the callers is simple and strives

[2]http://www.microsoft.com/india/keynote.aspx
[3]http://www.esagu.in/esagu/
[4]http://www.manobi.net/fichier/Manobi.pdf
[5]http://aaqua.persistent.co.in/aaqua/forum/index
[6]http://www.sksindia.com/

to reduce the cognitive load on the caller. Also, given the numerous possibilities of integration with various services and use of real-time information [7], *VoiceSites* may need to be dynamically adaptable. We took a step in that direction by developing Conspeakuous [10] - an architecture for modeling, aggregating and using context in spoken language conversational systems to build intelligent user interfaces. It allows development of context-aware *VoiceSites* that adapt based upon real-time information (such as user's current location), and based on the user preferences learned over time.

Another aspect that can be explored further is the use of multi-modal applications for this set of population. It may be unaffordable today, but with advanced mobile phones becoming cheaper few years down the line, using iconic interface along with voice may help in leveraging the benefits of multi-modal interfaces.

- **Browsing:**

  The *VoiceSites* are accessed by calling the *VoiNumber* corresponding to the site. As the number of sites increase, it will be impractical to keep track of and call each individual site. The users will need an effective T-Web browser that will allow them to browse across VoiceSites, bookmark their favorite sites, view history of visited sites and possibly, browse multiple *VoiceSites* simultaneously. It will be challenging to design a T-web browser especially because it might not have any visual interface.

- **Scalable Search:**

  As WWTW gains in popularity and use, it can be foreseen that searching an increasing number of *VoiceSites* in WWTW will be a tremendous scalability challenge. For one, offline crawling does not support real-time information updates, and moreover, the key piece of information to be searched would be voice. This requires advanced automatic speech recognition techniques for indexing as well as methods to consider recognition errors. Another key research challenge is to get access to the source code of the *VoiceSites*. Since *VoiceSites* are accessed by *VoiNumbers*, the actual source code that corresponds to a particular *VoiNumber* is not available by crawling a *VoiceSite*, since crawling in T-Web world involves interacting with the *VoiceSite* over telephone. Therefore, for enabling search, means for accessing the *VoiceSite* source will have to be addressed.

- **Accessibility:**

  Though none of the conventional accessibility issues of WWW (e.g. economic viability, basic education, visual impairment) exist in WWTW, there are other accessibility challenges associated to WWTW. Underlying speech recognition and synthesis technologies in WWTW need to support multiple languages which is a big challenge. DTMF and numeric inputs can be used as an alternative but that limits the user interface. Moreover, for the population that has speaking or hearing disorders, alternative access channels to the WWTW will have to be developed.

- **Infrastructure:**

  WWTW aims to deliver IT to a significantly large population by enabling them to openly participate in and contribute to the ecosystem. It requires support for easy means of introducing new services, enabling quick and easy deployment as well as seamless integration with back-end services in the infrastructure. To achieve that, significant challenges need to be overcome since the telecom world is much more controlled than the Internet. There is a lack of open standards and network operators are not comfortable opening up their infrastructure to third parties. This might restrict the growth and usability of WWTW.

7. SUMMARY

T-Web is an attempt to envision a service for the underprivileged communities, similar in theme to what WWW is to the IT literate users today. It enables masses to access information and services through voice driven channels. Information and services could be community created as well as leveraged from existing Internet infrastructure. We have summarized key technology enablers for this vision to be successful as well as attempted to articulate interesting research problems in this vision.

Enabling voice-driven front-ends to websites and WWW services would only enable the underprivileged to access global information [5]. Telecom Web builds a vision of a service for users in developing regions that harnesses WWW services as well as the ones in the converged networks — under one umbrella. Further, it provides the means to create and sustain an ecosystem of local (and global) services, information and communities *relevant* to these underprivileged users.

8. REFERENCES

[1] Common Services Centre (CSC). *Department of Information Technology, Government of India, http://www.mit.gov.in/default.aspx?id=825.*

[2] S. Agarwal, D. Chakraborty, A. Kumar, A. A. Nanavati, and N. Rajput. HSTP: Hyperspeech Transfer Protocol. In *Hypertext 2007 Proceedings of the ACM Conference on Hypertext and Hypermedia. To appear*, UK, September 2007.

[3] G. Camarillo, M. Angel, and G. Martin. The 3G IP multimedia subsystem (IMS): Merging the internet and the cellular worlds. John Wiley and Sons Publishers.

[4] D. Chakraborty, K. Dasgupta, S. Mittal, A. Misra, C. Oberle, A. Gupta, and E. Newmark. Businessfinder: Harnessing presence to enable live yellow pages for small, medium and micro mobile businesses. In *IEEE Communications, Issue on Networking Technologies in Emerging Economies*. January 2007.

[5] R. Hernandez and Y. Mugica. What Works: PRODEM FFP's Multilingual Smart ATMs for Microfinance. *http://www.digitaldividend.org/pdf/prodem.pdf*, August 2003.

[6] Internet World Stats. World Internet Users and Population Stats. *http://www.internetworldstats.com/stats.htm*, March 2007.

[7] A. Kumar, N. Rajput, D. Chakraborty, S. Agarwal, and A. A. Nanavati. Voiserv: Creation and delivery of converged services through voice for emerging economies. In *WoWMoM'07 Proceedings of the 2007 International Symposium on a World of Wireless, Mobile and Multimedia Networks. To appear*, Finland, June 2007.

[8] R. Kumar. E-Choupals: A Study on the Financial Sustainability of Village Internet Centers in Rural Madhya Pradesh. *Information Technologies and International Development*, 1:45–73, 2004.

[9] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal. The Web as a Graph. In *Proceedings of 19th ACM SIGACT-SIGMOD-AIGART Symposium on Principles of Database Systems, PODS*, 2000.

[10] A. Nair, A. A. Nanavati, and N. Rajput. Conspeakuous: Contextualising conversational systems. In *Proceedings of the 12th International Conference on Human-Computer Interaction. To appear*, China, July 2007.

[11] T. Parikh and E. Lazowska. Designing an Architecture for Delivering Mobile Information Services to the Rural Developing World. In *Proc. Intl. Conf. on World Wide Web (WWW), Scotland*, May 2006.

[12] Population Reference Bureau. 2006 World Population Data Sheet. Technical Report ISSN 0085-8315, August 2006.

[13] C. K. Prahalad. The Fortune at the Bottom of the Pyramid: Eradicating Poverty Through Profits. In *Wharton School Publishing*, 2004.

[14] R. Wedgeworth. State of Adult Literacy. *Proliteracy Resources, http://www.proliteracy.org*, September 2004.

[15] M. Yunus. The Grameen Bank. *Scientific American Magazine*, November 1999.