Towards Inclusive Social Networks
for the Developing World

Christian Akpona[1], Rose Gohoue[1], Herve Ahouantchede[1],
Fatna Belqasmi[2(✉)], Roch Glitho[1,3], and Jules Degila[1]

[1] IMSP, Porto Novo, University of Abomey-Calavi, Cotonou, Republic of Benin
[2] Zayed University, Abu Dhabi, United Arab Emirates
`fatna.belqasmi@zu.ac.ae`
[3] Concordia University, Montreal, Canada

Abstract. Social networks (e.g. Facebook) are becoming ubiquitous, especially in the developed world. However, they rely on two fundamental assumptions (i.e., reliable Internet connectivity, literacy) that exclude scores of potential end-users, especially in the developing world. Including these potential end-users will require lifting those two assumptions. This paper is a first step towards inclusive social networks in the developing world. It proposes and validates a two-layer system architecture. The proposed architecture allows end-users to access the social network with low-end cellular phones, using voice and Short Message Service (SMS), in addition to traditional access via a computer or smart phone, using keyboards. We have built a proof of concept prototype in which poorly literate end-users are included in two simplified dedicated social networks (a small farmers' social network and a motorcycle taxis' social network).

Keywords: Social networks · Developing economies · Inclusive social networks · SMS · Low-end cellular phones · Text-to-speech

1 Introduction

Social networks are becoming ubiquitous, especially in the developed world. Facebook, for instance, announced during the release of its third quarter 2016 results that it had more than a billion daily active end-users. There is still no consensus on the fundamental definition of a social network, especially as most social networks embed a plethora of heteroclite applications. In this paper we use the definition provided by Ref. [1], i.e., that social network sites are web-based services that allow individuals to construct a public or semi-public profile within a bounded system, articulate a list of other users with whom they share a connection, view and traverse their list of connections and those made by other users within the system. In the definition, publishing a comment is part of the profile construction, and accessing publications by other individuals is part of viewing and traversing the list of one's connections.

Social networks remain quite exclusive despite their popularity, especially in the developing world. One of the primary reasons is that accessing a Web-based service requires an Internet connection. However, according to the ITU-T [2], only 31% of

people in the developing world accesses Internet, compared to 77% in the developed world. Another reason is that constructing profiles, articulating lists of users, viewing and traversing list of connections are usually done via computers/smart phones' keyboards. A certain level of literacy is indeed required. Nevertheless, according to UNESCO [3], while the average literacy rate is generally above 80% in the developed world, it is usually below 50% in the developing world.

There are several compelling scenarios for inclusive social networks, especially in the developing world. Dedicated social networks are the most convincing; small farmers' social networks and motorcycle taxis' social networks are illustrative examples. This paper is a first step towards inclusive social networks in the developing world. It proposes and validates a two-level system architecture for inclusive social networks. The following section provides a critical overview of the state of the art. The proposed system architecture is presented in the third section. This is followed by a description of the proof of concept prototype. We conclude in the last section.

2   A Critical Overview of the State of the Art

Motivating scenarios and requirements are presented first, followed by critical overviews of the work on social networks for areas with poor Internet connectivity and of the studies on the use of social networks by poorly literate end-users.

2.1   Motivating Scenarios and Requirements

One can envision several compelling scenarios of dedicated social networks in the developing world. Let us start with a small farmers' social network. Farming and its myriad issues are of vital importance in most of the developing world. Most small farmers are barely literate and live in rural areas with poor Internet coverage. Information on evolving farming techniques and marketing news are both vital to their survival. In several developing countries, there are government officials who act as advisors and pay visits to small farmers now and then. One could imagine an inclusive social network that allows government officials to spread timely information to small farmers while allowing farmers to exchange information on the problems they face, as well as facilitate the exchange of best practices.

Let us move to a more urban/village setting and propose motorcycle taxis' social networks. Motorcycle taxis are common in several African countries (e.g. "zemi djan" in the Republic of Benin, "Boda boda" in Uganda). Some of the drivers are illiterate while others are university graduates. In some countries, they are renowned as being highly politicized, as in the Republic of Benin. One could well imagine a dedicated social network where they could exchange ideas on on-going political events. This very social network could easily act as a dispatching system since there is usually no centralized taxi dispatching system for these motorcycle taxis. A driver dropping a client near a given landmark (e.g. church, mosque, or other sort of landmark) could publish a comment to signal the number of clients waiting for taxis at that point. Another driver who retrieves the publication might decide to go to the landmark in

order to pick up one of the clients. She/he would then respond to the comment to signal to the other drivers that she/he is going to that site.

A first requirement on the envisioned inclusive social network is obviously the ability to cope with poor Internet connectivity. The second one is to address the possibility of its use by poorly literate end-users. Beyond these two requirements, which are easily derived from the use cases, there are three additional requirements that need to be taken into account. The first is the possibility to access the social network with very low cost phones that support only voice and SMS. This reflects the fact that small farmers and the motorcyle taxi drivers are not at all likely to own high-end phones.

The second additional requirement is that the social network should offer the full set of social networking services to all end-users. The third additional requirement is re-use. The envisioned social network should re-use the infrastructure that has already been deployed whenever possible to reduce deployment cost. This deployed infrastructure includes the rich cellular network infrastructure as well as the exclusive social network infrastructure.

2.2   Social Networks for Areas with Poor Internet Coverage

Reference [4] proposes a social network that addresses the poor Internet connectivity issue. In the use case, Vijay, who has lost his goose, posts a comment on the community social network, in the hope that other members of the community will comment back on the whereabouts of the goose. The paper assumes that Vijay and the other members of the community are literate enough to understand highly intuitive graphical user interfaces. It is also assumed that they might eventually get help from trusted literate friends and family members for some of the social network operations.

The proposed architecture relies on the GSM infrastructure where it exists, and uses delay tolerant networking (DTN) to cater to areas where there is no GSM coverage [4]. DTN [5] is an overlay network that runs on top of transport layers and provides a store and forward mechanism to overcome intermittent Internet connections. In the architecture, this overlay is built on top of the very phones used by the members of the community. This means that the phones are expected to have short-range radio interfaces such as Bluetooth to take advantage of social encounters for communicating, acting as nodes of the DTN overlay. Low-cost phones with only voice and SMS support could not be used, and thus this system does not meet our requirements. The architecture does re-use the existing infrastructure as stipulated in our requirements. However, it does not re-use the existing infrastructure to the fullest extent since the social network infrastructure is not re-used at all.

We have previously proposed SNS4D, a social networking system for developing countries [6]. The system architecture proposed here actually builds on that system. SNS4D enables ubiquitous access by offsetting the poor Internet connectivity with the rich SMS access provided by cellular networks. It offers the same set of services to end users with SMS access and end-users with Web access. However, it assumes that all end-users are literate enough to use SMS, and this assumption does not hold in many parts of the developing world. As mentioned in the same paper [6], it is important to note that some existing social networks (e.g. Facebook) allow access by SMS. However, the set of services offered through this access is usually restricted. It is also

important to note, as mentioned in the very same paper [6], that social networks accessible exclusively via SMS have also been deployed, especially in India (e.g. Gupshup). It is obvious that neither the existing social networks that provide SMS access for a sub-set of functionality nor SMS-based social networks meet all our requirements.

Daknet [7] tackles the general problem of Internet connectivity in the developing world. It provides mobile ad hoc connectivity by following a model similar to the traditional postal system. Data is transmitted over short point-to-point links between kiosks and portable storage devices. Kiosks store the original content the end-users wish to transmit via the Internet. The portable storage devices mounted in vehicles such as buses and motorcycles retrieve the content from the kiosks and transport it "physically" to hubs that are wireless Internet access points. Daknet is content-agnostic and social network services may be built on top of it. However, the poor literacy issue remains to be solved.

2.3   Use of Social Networks by Poorly Literate End-Users

Voice is certainly the most natural modality that even barely literate end-users could use to interact with social networks. Several voice-enabled social networks have been proposed in the literature, although they do not usually target poorly literate end-users. Voice is rather used for convenience by literate end-users. Vehicular social networks are the archetypes. RoadSpeak [8] is one example. It enables voice chats between drivers on popular roadways and relies on an underlying social network overlay. The motivations are entertainment, utility and emergency. These voice-enabled social networks assume Internet connectivity, and are generally not accessible with low cost phones.

Spoken Web [9] provides a general architecture for a Web where sites are created and browsed with voice via ordinary phone calls. It re-uses the telecommunication infrastructure, provides Web access to poorly literate end-users, and low cost phones can be used. Its key concepts are voice sites, voice numbers and voice links. Voice sites are interconnected with voice links and are accessed by calling the associated voice numbers. The calls are mapped onto the underlying telecommunication infrastructure. The sites can be created via phone calls using simple voice-driven interfaces. While this architecture includes poorly literate end-users in Web browsing and voice site creation, it does not give any insights into how to design Web complex-based services such as inclusive social networks.

Reference [10] describes an architecture that enables Internet access with voice (and SMS). It is known as SiteOnMobile. Low cost phones can be used, just as we envision in our architecture. The architecture re-uses the telecommunication infrastructure, as with the spoken Web architecture. It uses concepts such as TaskLet, which enables a Web of page to become a Web of tasks. A smart gateway is used to convert SMS patterns into TaskLet invocations. Text to speech conversions are performed on the web content accessed by the end-user. Just like SpokenWeb, SiteOnMobile includes poorly literate end-users in Web browsing, but does not give any insights into how to design complex Web-based services such as inclusive social networks.

3   Proposed System Architecture

Figure 1 shows the proposed system architecture. It is comprised of two layers: a front-end layer and a back-end layer. The front end layer is made up of an access sub-layer and a mediation sub-layer. The architectural assumptions and principles are presented first, followed by a description of the functional entities and the interfaces. The last sub-section discusses the procedures.
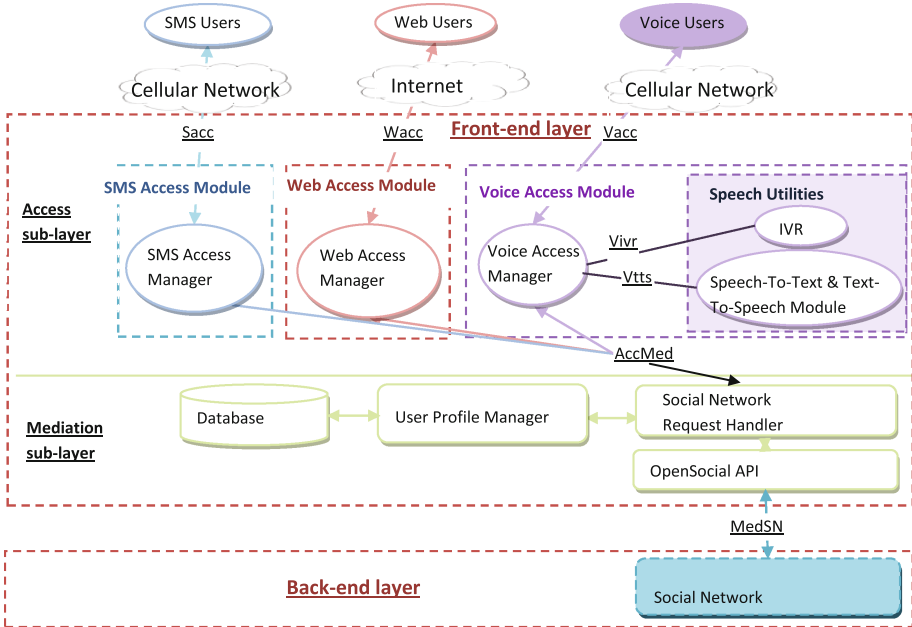

Fig. 1. Overall architecture

3.1   Architectural Assumptions and Principles

Our key assumption is that there is cellular network coverage wherever there is poor Internet connectivity. This is not farfetched at all, since according to an ITU-T report [2], cellular network coverage has now reached 89% in the developing world. Furthermore, this assumption helps us avoid more tenuous assumptions, such as the one made in Ref. [4] (i.e., the support of short-range radio connection by all phones for DTN connectivity) that are much less likely to hold in many parts of the developing world.

  The first architectural principle is that we build on the deployed social network infrastructure in addition to the cellular network infrastructure. We use the OpenSocial standard [12] for that purpose. It is a set of programming interfaces that enable the development of social applications that are portable and inter-operable across the social networks that support the standard. The back-end layer of our proposed architecture could be any social network that supports OpenSocial (e.g. Google+ social network). The front end can be considered as a social application running on top of OpenSocial.

The second architectural principle is that text is used as the common denominator for requests (e.g. account creation, publication, and others). The key reason is that the existing social networks that could be used as a back-end layer via OpenSocial only support text as the interaction modality – voice is not supported. Requests made by voice users in our architecture are thus always translated into text, and textual comments retrieved from the back-end are translated into voice for voice users.

3.2   Functional Entities and Interfaces

The access sub-layer has four functional entities: SMS access manager, Web access manager, Voice access manager and speech utilities. The SMS access manager, the Web access manager and the voice access manager entities offer the Sacc, the Wacc and the Vacc interfaces, respectively, to enable access by SMS users, Web users and voice users, respectively. The Sacc and Vacc interfaces bridge cellular networks and our proposed inclusive social network, and their implementation could rely on cellular modems. Wacc is nothing more than the usual interface for accessing Web-based services, and it can be implemented as a simple Web page. The speech utilities include the Interactive Voice Response (IVR) system and the speech to text (STT) and text to speech (TTS) systems. They are accessible via the Vivr and Vtts interfaces, respectively.

Vivr and Vtts are modelled according to the Representational state transfer (REST) principles. REST is an architectural style for designing distributed client-server applications. In REST, each resource is identified by a unique Uniform Resource Identifier (URI) and is accessed via a subset of HTTP methods. The most common of these methods are GET, POST, PUT, and DELETE, which can be used to read, create, update and delete a resource, respectively. Reference [12] provides an overview. The list of resources defined by Vivr interface are shown in Table 1 for illustration purposes, along with the URI of each resource and the HTTP methods that are supported. The call-back resource is used by the IVR to send back a notification to the voice access manager about the termination of an ongoing communication session. The same applies to sending back the notification about the termination of a TTS or an STT request. The call back resources are used because the recording and the translation requests may take time and therefore the initiating requests may timeout before the processing has been completed.

The mediation sub-layer mediates between the access sub-layer and the back-end layer. For instance, when a comment is retrieved from the social network in the back-end layer, it ensures that the comment is dispatched to the correct access manager. It comprises a database, a user profile manager and a social network request handler. The database contains information such as the access manager to which end-users are connected at any given point in time. It is updated by the user profile manager. The social network request handler processes requests from the access sub-layer and interacts with either the user profile manager (e.g. updating the access manager used by an end-user) or the OpenSocial API (e.g. publication of a comment). Every end-user has two accounts: one in the social network of the back-end and the other in the inclusive social network. The first account is transparent to the end-user and is automatically created by our system. The access sub-layer and the mediation layer interact via the AccMed interface. AccMed is also modeled according to REST principles.

Table 1. VIVR rest interface

Resources	Operation	HTTP method: resource URI
List of sessions	Create: establish a new communication session between an end-user and the IVR	POST: http://www.ivrSNS.com/
A specific session	Read: get information about a specific communication session	GET: http://www.ivrSNS.com/{idCom}
	Delete: end a specific communication session	DELETE: http://www.ivrSNS.com/{idCom}
Recording	Create: record a specific session	POST: http://www.ivrSNS.com/{idCom}/record/
A specific recording	Delete: delete a specific recorded communication session	DELETE: http://www.ivrSNS.com/{idCom}/record/{idRecord}
Call-back	Create: send back a notification to the voice access manager	POST: http://www.voiceAccessManager.com/ivr

3.3   Procedures

New account creation and information publication procedures are described below for the purpose of illustration.

New Account Creation: To create a new account using voice commands, the following steps are followed:

- The end-user calls the phone number for the voice access manager, which transfers the call to the IVR.
- The end-user then chooses the option to create a new account via the IVR menu. He/she follows the voice steps offered by the IVR to enter the required information (e.g. the pseudo name, the first and last names, the password). The communication is saved in a voice file on the IVR side.
- At the end of the communication, the IVR informs the voice access manager and sends it the URI of the recorded voice file. The voice access manager then calls the appropriate API on the STT to translate the voice commands into the corresponding text request, and then sends the request to the social network request handler for processing.
- The social network request handler instructs the user profile manager to create a new user profile in the local database and then issues a new account creation request which it sends to the back-end social network. The local profile includes information such as the request-ID and the access manager from which the request was received. When an account creation confirmation is received from the back-end social network, the request handler transfers the response to the originating manager. If the end-user has asked for a confirmation, the voice access manager uses the TTS to create a speech file that corresponds to the received response (the output file will be stored on the TTS side), creates a voice call between the IVR and the end-user and instructs the IVR to play the output file to the user by giving the IVR the URI of the appropriate file.

To create a new account via SMS, the end-user sends an SMS message to the phone number associated to the SMS access manager, which transfers the request to the social network request handler. The process then proceeds in the same way as for the voice request, creating a new account on the back-end side. When the response is sent back to the SMS access manager, it creates the appropriate SMS response and sends it to the end-user, if the latter has asked for a confirmation. The same procedure is applied for account creation via the Web, except that the end-user communicates with the system via the Web.

Publication: The publication procedure is similar to that of account creation. We therefore summarize the main steps for voice publication and focus on the differences. To publish new information, the end-user calls the voice access manager and he/she is put in contact with the IVR. After choosing the appropriate option from the IVR menu, the end-user dictates the information to be published to the IVR, which stores it in a voice file. The access manager transcribes the posted information into text before issuing a publication request towards the back-end. It also retrieves the end-user contact list from the back-end and sends a notification to the SMS members that requested to be informed about new publications. The Web and voice end-users will be informed when they connect to the system and explicitly ask for new publications.

4 A Proof of Concept Prototype

Figure 2 depicts the overall prototype setup. We first introduce the implemented scenarios. This is followed by a short description of how the TTS/STT tools have been customized for Fon, a vernacular language of the Republic of Benin. We end with a description of the prototype itself including its setup.
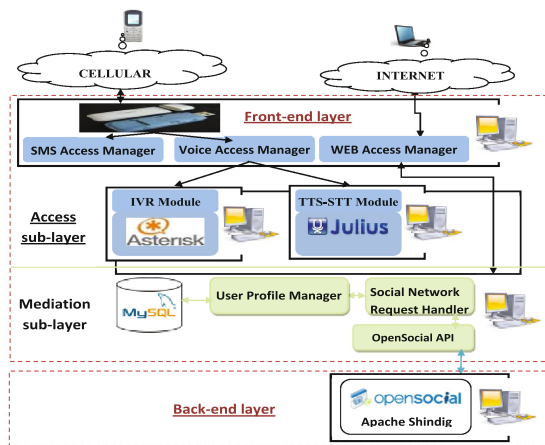

Fig. 2. Prototype setup

4.1   Implemented Scenarios

We have implemented two scenarios. The first relates to the small farmer social network. The second deals with the dispatching functionality of the motorcycle taxis' social network we discussed earlier. In the first scenario, we consider a simplified poultry epidemic situation in a given area. Sika (a farmer) notices the death of two chickens on her farm. She is a voice user and publishes the information on the social network. This triggers a series of publications by other farmers in her network. Dossou, a voice user, indicates that five of his chickens have died, and Ablawa, who is an SMS user, subsequently signals the death of six chickens on her farm.

Jean, the veterinarian who is also a member of the small farmer social network, is a Web user. Upon receipt of this series of publications, he realizes that there is a poultry epidemic. He then publishes a message through his Web interface to tell the farmers that preventive pills are available (for sale) at a specific veterinarian pharmacy in the area. We assume that there are several veterinarian pharmacists in the area and that Jean knows the type of medicine each one sells.

In the second scenario, Bossou, a motorcycle taxi driver who is a voice user, notices that there are two clients waiting for a taxi at a specific place. He publishes this to the other drivers who are looking for clients. Pierre is one of these drivers and is an SMS user. Upon receipt of the publication, he decides to go to that place to pick up one of the clients. He notifies the other drivers via a publication. Kokou is another driver and an SMS user. After receipt of the two publications, he realizes that there might still be one client waiting. He decides to go to pick her/him up and notifies the group via a publication.

4.2   Customization of the STT and TTS Engines

There is no STT/TTS engine readily usable for the Fon language. This has led us to the customization of Julius [13] for STT and Festival [14] for TTS. This choice is motivated by the flexibility of these engines – both Julius and Festival have been customized in the past for a wide range of languages. We have considered a very limited subset of Fon, and this subset has made the implementation of the two scenarios possible, within well-restricted boundaries.

The sub-set comprises the following concepts: number, entity, state, and place. Farmers, motorcycle drivers, and preventive pills are entities. "Standing in", "going to", "dead in", and "available in" are the states.

All sentences are expressed as "A number of a given entities are in a given state in a given place". For example, "3 chickens are dead in Jonkey". Jonkey is a well-known neighborhood in Cotonou, the capital city of Benin. Another example is "A motorcycle client is going to Jonkey". The number can actually be omitted in some cases such as "Preventive pills are available in Jonkey". We explain briefly below how we have customized Julius with the subset.

Julius offers speech recognition by combining a language model and an acoustic model. The language models supported are N-gram, rule based grammar, and isolated word recognition. The acoustic model can be monophone or triphone, or a combination. Reference [14] can be consulted for tutorial-level information. We have used

isolated word recognition and a very simple grammar, with two sentence structures. The first is when the number is not omitted and the second is when it is. A specificity of FON that was factored in is that "number" (when not omitted) comes after "entity". This leads to the following sentence structure "entity number state place". A combination of monophones and triphones were used. The numbers were limited to the digits from 1 to 10, the places to 5 well-known Cotonou neighborhoods.

4.3   Prototype

The prototype setup is shown in Fig. 2. The modules inside the same rectangle with solid borders are deployed on the same machine. The SMS and voice end-users communicate with the front-end layer via their cellular network operator. The users' SMSs and voice calls are routed to the number of a SIM card on a GSM modem (i.e. HUAWEI E153) connected to the machine running the front-end layer modules. The voice calls are automatically routed to the voice access manager. The SMS access manager uses a Java library (i.e. SMSLib API) to receive and send SMS messages via the GSM modem.

The IVR module is implemented using Asterisk, a free and open source software framework for building communications systems. Asterisk supports a number of VoIP signalling protocols. In our prototype, we use the session initiation protocol (SIP). We use an Oktell SIP-GSM gateway [15] to bridge between the GSM and SIP domains on both the signalling and the media planes.

The back-end social network is implemented using Apache Shindig, the reference implementation of OpenSocial API specifications. It is open source and it assists in the building and hosting of new OpenSocial applications.

5   Conclusion

Social networks currently exclude scores of potential end-users de facto, especially in the developing world. This is due to the two fundamental assumptions on which they rely on: Internet connectivity and literacy. This paper has made a first step towards inclusive social networks by proposing and validating a new system architecture. It is a first step; there are several research directions possible in this area. We have assumed good cellular network connectivity wherever there is no Internet connectivity. Although the assumption is not farfetched, it should be possible to omit it for some situations. However, omitting that assumption should not be replaced by a less realistic assumption (i.e. the support of short-range radio connection by all phones in order to enable opportunistic connections). More research is needed. We envision inclusive social networks as dedicated social networks or as closed groups of general purpose social networks. Ontologies need to be developed to formalize the vocabularies used in the dedicated areas. This is a key to the optimal customization of the TTS and STT tools. Social networks are certainly the most complex social media that exist today. However, there are other social media (e.g. blogs, podcasts) that have become as exclusive as social networks. Architectures are also needed to make them more inclusive.

References

1. Boyd, D.M., Ellison, N.B.: Social networks: definition, history and scholarship. J. Comput. Mediat. Commun. **13**, 210–230 (2008)
2. ITU-T, ICT facts and figures. www.itu.int/en/ITU-D/Statistics/Documents/…/ICTFactsFigures2013.pdf
3. UNESCO: The official source of literacy data. http://www.uis.unesco.org/literacy/Pages/default.aspx
4. Vallina-Rodriguez, N., Hui, P., Crowcroft, J.: Has anyone seen my goose? Social network services in developing regions. In: 2009 International Conference on Computational Science and Engineering (2009)
5. Fall, K.: A delay-tolerant network architecture for challenged internets. In: Proceedings of SIGCOMM (2003)
6. Ahouantchede, H., Belqasmi, F., Glitho, R.: SNS4D: an Online social network system for developing countries. In: Fourth International IEEE EAI Conference on e-Infrastructure and e-Services for Developing Countries, Africomm 2012, November 2012
7. Pentland, A., et al.: DakNet: rethinking connectivity in developing nations. IEEE Comput. **37**(1), 78–83 (2004)
8. Smaldone, S., et al.: RoadSpeak: enabling voice chat on roadways using vehicular social networks. In: SocialNets 2008, April 2008
9. Agarwal,. S.K., et al.: The Spoken web: a web for the underprivileged. In: ACM SIGWEB Newsletter, Summer 2010
10. Manjunath, G., et al.: Delivering mobile eGovernance on low end phones. In: 2012 IEEE International Conference on Mobile Data Management (2012)
11. Hasel, M.: Opensocial: an enabler for social applications on the web. Commun. ACM **54**(1), 139–144 (2011)
12. Belqasmi, F., Fu, C., Glitho, R.: RESTful web services for service provisioning in next generation networks: a survey. IEEE Commun. Mag. **49**(12), 66–73 (2011)
13. Lee, A., et al.: Recent development of open source speech recognition engine Julius. In: APISIPA ASC (2009)
14. Beutnagel, M., et al.: The AT&T next-gen TTS system, Joint Meeting of ASA (1999)
15. http://www.oktellpbx.com/gateways/, June 2001