A Picture and a Thousand Words: Visual Scaffolding for Mobile Communications in Developing Regions*

Robert Farrell
IBM T J Watson Research Center, USA

Catalina Danis
IBM T J Watson Research Center, USA

Thomas Erickson
IBM T J Watson Research Center, USA

Jason Ellis
IBM T J Watson Research Center, USA

Jim Christensen
IBM T J Watson Research Center, USA

Mark Bailey
IBM T J Watson Research Center, USA

Wendy A. Kellogg
IBM T J Watson Research Center, USA

ABSTRACT

Mobile communication is a key enabler for economic, social and political change in developing regions of the world. The authors argue that engaging citizens in developing regions in information creation and information sharing leverages peoples' existing social networks to facilitate transmission of critical information, exchange of ideas, and distributed problem solving. All of these activities can support economic development. They describe Picture Talk, a mobile social computing application framework designed to facilitate local information sharing in regions with sparse Internet connectivity, low literacy rates and having users with little prior experience with information technology. Picture Talk runs on today's internet-enabled smartphones as well as camera phones with multimedia messaging, now available in many of these regions.

**To Appear:* Farrell, R., Danis, C., Erickson, T., Ellis, J., Christensen, J., Bailey, M. and Kellogg, W.A. A Picture and a Thousand Words: Visual Scaffolding for Mobile Communications in Developing Regions. In D. Y. Jin (Ed.) *Global Media Convergence and Cultural Transformation: Emerging Social Patterns and Cultural Characteristics*, 165-178. IGI Global, 2011.

INTRODUCTION

We are interested in designing applications that enable people at the base of the economic pyramid (BoP) to create, share, and discuss information as is commonly done on the World-Wide Web today, but through mobile technologies. The BoPincludes over one billion people with little access to computer technology often living on less than $1US per day in some of the least developed countries in sub-SaharanAfrica, South/CentralAmerica and the Indian Sub-continent, as well as several other parts ofAsia.As others have recognized (Prahalad, 2004; Kumar et al, 2008), enabling connections among a wide spectrum of people can lead to the empowerment of the disenfranchised and enable people who have been largely excluded from modern technology and economic opportunity to express their entrepreneurial tendencies. This could result, for example, in the creation of broader markets for local goods and services. The global reach of mobileM communication networks offers, for the first time, a broad platform for delivering applications and software services that can realize this potential.

We have three design goals for the mobile applications we build. First, we want these applications to be usable by even the most disadvantaged users, many of whom are illiterate Second, we want to enable these users to document local needs, problems, and issues by creating, storing, and sharing digital artifacts (e.g., maps, photos, graphics, radio news reports, music, games, TV segments, informal news). Third, we want to enable these users to engage in conversation about these digital artifacts to offer solutions, share perspectives, or to engage in purely social exchanges.

Our initial implementation toward these goals is Picture Talk, a mobile social computing application framework that supports three major features: first, mobile phone users can start an inclusive discussion forum with other mobile phone users by asynchronously sharing short recordings of their voice over a wireless network; second, they can supplement and provide structure for the unfolding conversation by adding pictures; and third, they can discover and participate in discussions shared by other users, reaping the rewards of prior users' ideas, comments, and solutions. These features could enable Picture Talk to become a powerful tool to support existing social behavior and social connections between users (Danis et al., 2009).

This article first provides some background on some of the obstacles that BoP communities face in trying to access information technology, then introduces the problem of supporting information creation and dissemination in this environment, and then describes mobile social computing as a potential solution, including the Picture Talk application framework we have developed. We conclude with some future directions and conclusions.

BACKGROUND

BoP populations face a number of obstacles in becoming part of the global online community. In the economically developed world, access to information technology has been largely through Internet-connected computers. An important benefit of access to the Internet has been the potential for contact with the worldwide community of users. Through online communities, Internet users engage in discussions on topics of common interest, write comments that serve as a means of self-expression, and solve each others' problems. We argue here that through mobile phones, similar applications could be deployed to BoP communities to enable discussions on topics of local interest, provide a voice for individuals who would otherwise have no forum for their ideas, and enable solutions to communal problems through information exchange.

There remain at least three obstacles to the widespread use of information technology to support online communities and other applications in developing regions. First, despite initiatives such as One Laptop Per Child (OLPC, 2009), conventional computing technology continues to remain out of reach of the large majority of the BoP population. Lack of reliable networks to access the Internet further limits the ability of people in these regions to access information even where capable devices are available. BoP users are necessarily very cost conscious, driving a need for a low cost platform. Adapting applications designed for powerful computers with a keyboard and a large display for BoP users with mobile phones remains a challenge. Second, low literacy rates in some regions prevent significant portions of the BoP users from using the Internet's predominantly textual interaction mode. Despite vast quantities of online information, the plethora of native languages and alphabets in BoP regions render many of these resources useless even for people who are literate in their local language. Third, despite skills and experience in the social use of mobile phones, many BoP users may have little familiarity with or motivation to use the device to access information services, preferring face-to-face interaction. A radical rethinking of the human interface to information technology may be required.

The statistics for even basic access to electricity in developing regions are alarming. According to the Open Society Initiative for Southern Africa (OSISA, 2009), approximately 90% of Africa's one billion people have no regular access to electricity. Even where there are facilities, blackouts are common. The bulk of power plants and transmission facilities in West Africa, for example, are over 50 years old. With little maintenance or investment, the aging equipment fails or operates below capacity (Madamombe, 2005). Despite these obstacles, where power to homes is not available, people often travel to a centrally located solar-powered, wind-powered, or coin-operated charging station to maintain use of their mobile phone.

Global statistics for computer usage demonstrate huge differences between developed and developing countries. For example, the highest rates of access to the Internet in 2007 were in Sweden (82%), the US (81%), South Korea (81%), and other developed countries, whereas the lowest access rates included Tanzania (6%), Kenya (12%), and Uganda (11%). Similarly, computer ownership is lowest in Uganda and Tanzania, both at 2% (The Pew Research Center for the People and the Press, 2007).

The picture changes radically when considering mobile phones rather than computers. Arecent survey by the International Telecommunications Union (ITU) found that while only one quarter of the earth's population of 6.7 billion uses the Internet, nearly two thirds of the population uses mobile phones (ITU, 2009). Wireless phone use is exploding in the developing world: sixty-eight percent of mobile phone subscribers worldwide are outside of North America and Europe. In Africa, regular users of mobile phone services have jumped from 10 million to 400 million in the five years from 2003 to 2008 and the growth is still accelerating (ITU, 2009). The rate of mobile phone ownership in the Ivory Coast, Mali, Nigeria and South Africa is over 60%, higher than in Canada (The Pew Research Center for the People and the Press, 2007). Mobile phone use in economically developing regions crosses the barriers of gender, age, and education (Samuel et al., 2005).

The rapid growth of mobile phone networks in BoP markets is fueled by the need for communication in environments where there are few alternatives. The lack of traditional wired infrastructure creates an opportunity: much of the developing world is a "green field" where new technologies can be deployed without being hampered by existing business models, infrastructures, or user expectations. For example, in many parts of Africa, wireless networks have leapfrogged the public switched telephone network in terms of installed base. In 2007, the African continent had 280 million total telephone subscribers, but 260 million of these were mobile cellular subscribers. Building a new wireless network is faster, easier, more reliable, and less expensive than putting in a whole new wired infrastructure.

Despite the growth of wireless networks, few developing countries yet have data communication channels sufficient to provide rural populations with access to the public Internet. In 2008, only 5% of Africa (Appfrica, 2008) had access to the Internet. Still, it is likely that mobile data services (e.g., Global Packet Radio Service, GPRS) and mobile broadband (e.g. WiMax) will be the first contact that most Africans have with the Internet.

The capabilities of mobile phones are also increasing rapidly. The first deployments of camera phones occurred in 2001 and by 2004, 370 million mobile phones with digital cameras were sold (InfoTrend/CAP Ventures, 2004). The total number of mobile camera phones is expected to exceed 1 billion by 2011 with fastest growth in emerging markets (Lyra Research, 2007).In the late 1980s and early 1990s, mobile phones were used for voice communication only and users typed on a numeric keypad. Today's smartphones have high resolution touch screen displays, miniature keyboards, and other flexible input methods and are used for mobile computing, multimedia messaging, information access, and more. Worldwide smartphone sales increased 12.7 percent in the first quarter of 2009 and sales are anticipated to grow at more than a 30% compound annual growth rate over the next five years (Gartner Group, 2009).

We anticipate that smartphones connected to a wireless phone network will become a viable platform for delivering mobile ICT in Africa in the coming years.

Another major hurdle is literacy. Literacy is typically defined as the ability to read and write, however there is an inherent lack of precision that results from the methods of assessment and thus official figures often over-estimate functional literacy. For example, the commonly cited statistics, such as those compiled by UNESCO (2009), are based on census and other self-report methods, which are fundamentally inexact. Also, the definition of literacy can vary from 'the ability to write a simple sentence' to 'being able to freely communicate ideas in literate society.'Individuals as young as fifteen who may have been counted as literate because they were attending primary or secondary school may, because of limited continued use of written language, become functionally illiterate as adults (Seshagiri, Sagar & Joshi, 2007). According to UNESCO (2009), two-thirds of the world's 785 million illiterate adults are found in only eight countries (India, China, Bangladesh, Pakistan, Nigeria, Ethiopia, Indonesia, and Egypt). Low literacy rates are concentrated in South and WestAsia, sub-SaharanAfrica, and theArab states, with percentages averaging in the 60s, though some countries like Mali and Niger report rates for 15 to 20 year olds of less than 30%. Men typically have higher rates of literacy than women in traditional societies (UNESCO, 2009).

While there are no generally accepted statistics on how much of the Internet is available in different languages, it is generally accepted that the dominant language on the Internet is English, making much of the Internet linguistically inaccessible to the large majority of the BoP. The large number of languages spoken in BoP countries is intertwined with literacy and access to written information. UNESCO (2009) estimates over 2000 languages are spoken on the African continent. An additional challenge is that individuals who may be literate in their native language may nevertheless be functionally illiterate if information is available only in one of the country's official languages (Plauche and Nallasamy, 2007).

A third major challenge in introducing technology in developing regions is being sensitive to the social and cultural context in which the technologies operate. To provide one example from the southeastern Indian state of Kerala, fishermen now use mobile phones to get market price information before deciding where to sell their fish (Abraham, 2007).About 40% report an increase in income and 50% report fewer losses due to unsold or spoiled fish when they start calling for prices. Interestingly, however, few of these fishermen consistently go to markets with the highest prices; instead many choose ports where their "commission agent" has a presence. Because commission agents invest in the fisherman's business (e.g., financing the purchase of a fishing vessel), the fisherman feels a social obligation to bow to the agent's wishes, even when doing so may prevent him from maximizing his income.

Several field reports illustrate specific ways in which trust in one's social network may provide a bridge to the acceptance of information from official, but unfamiliar sources of information through computing technology. For example, farmers in the southern state of Tamil Nadu use web-connected kiosks (telecenters) fielded by a local sugar factory to ask only "simple" (i.e., low-stakes) questions of a purported agricultural expert who is not known to them, saving "high-stakes" questions for successful farmers with whom they have some pre-existing relationship (Srinivasan, 2007). Gopakumar (2006) explains that local people play a critical intermediary role in the success of telecenters. For example, living in the same village led target users of the Akshaya telecenter to develop trust in the entrepreneurs and intermediaries who ran the centers. By extension, they also developed trust in the abstract systems of medicine and government that were the ultimate sources of the information.

To summarize, these studies demonstrate the power that access to information can have in improving people's lives, but also how the impact of information is gated by social factors like trust, accountability, and social and institutional pressures.

MOBILE SOCIAL COMPUTING FOR DEVELOPING REGIONS

Social computing makes the claim that people are fundamentally social; therefore their use of technologies cannot be understood separately from social considerations. In other words, people access, evaluate, and consume information in a social context. We believe that social computing is uniquely well suited to contribute to effective communication and information dissemination in developing regions.

Issues, Controversies, Problems

The widespread availability of mobile phone applications in developing regions is already having a broad impact at the BoP. It is giving individuals a way to share critical information in a timely fashion, for example. Furthermore, individual contributions can be aggregated to provide value to a broader audience. For example, Ushahidi, meaning "testimony" in Swahili, is a platform for crowdsourcing crisis information. Ushahidi allows anyone to transmit geo-coded data via Short Message Service (SMS) and the application aggregates this data (Ushahidi.com, 2009). We think that mobile social computing applications need to support an inclusive dialogue, enabling both the ability for citizens to report their thoughts and experiences and comment on the contributions of others. a

As we mentioned above, illiteracy is also a major challenge. Designers of applications geared towards illiterate users have focused on non-text modalities in order to design more generally accessible applications in the countries with low rates of literacy. For example, speech is a widely used modality. However, limitations on the generality of speech recognition technology in multi-lingual environments (Plauche and Nallasamy, 2007) demands the use of other modalities for a broad set of functions needed to complement spoken language interfaces. For example, Joshi, Welankar, Kanitkar and Sheikh (2008) developed and tested a phonebook they call Rangoli aimed at low literacy populations. Rather than entering phone numbers based on alphabetical order, users are able to use a combination of color, icon and spatial location. Similarly, Froehlich and colleagues (2009) proposed applying digital storytelling (for example, video or sequences of still photographs accompanied with spoken annotations) as ways of enabling low literacy individuals to participate in information creation and sharing. However, the mobile devices prevalent in developing regions of the world have limited support for images and almost no support for video.

As noted previously, even for literate users in the population, the large number of languages in BoP countries makes it unlikely that the user's preferred local language will be used in the user interface. Thus the use of pictures to augment spoken language may allow many more people to have meaningful access to information. However, when speech is transposed into digital settings, a number of well-known problems arise. In the type of applications we are envisioning, conversations would be asynchronous, that is, carried out between people in different places speaking at different times. This means that conversations would lack some characteristics that are important for establishing and maintaining common ground (Clark & Brennan, 1991) - "the knowledge that the participants have in common, and they are aware that they have it in common" (Olson & Olson, 2000, pp. 157) - which is critical for enabling communication among participants in a discussion. For example, it would mean that speakers would not be able to see one another, or share visual cues like glances, gestures and shrugs that in collocated speech enable interlocutors to control the conversation's flow, easily refer to objects, and verify that they are being understood (e.g., Yankelovich et al., 2004). It also potentially means that many more people and people from a wider geographic area might be able to engage in a conversation, something that could be desirable but which also could exacerbate these problems.

Solutions and Recommendations

Picture Talk is a mobile social computing application framework intended to support a wide range of social interactions that can be accomplished through asynchronous communication, including conversations with remote participants, question and answer exchanges, and peer production of localized content. We begin by laying out the rationale that underlies Picture Talk by describing the scenarios and design sketches that marked the beginning of our design process. After presenting the initial vision, it goes on to describe a working prototype.

The concept of Picture Talk arose out of consideration of the issues mentioned above and how they might be addressed in the context of a mobile phone-based communication system. The crux of the solution was to augment speech with three types of visual component: 'comment proxies', pictorial contexts and visual controls. 'Comment proxies' are visual representations of digital speech that depict various types of meta-information, such as the identity of the speaker, the length of the comment, and the relationship of the comment to other comments (e.g., a reply); in the Picture Talk user interface they also provide direct access to the comment they represent, thus mitigating the difficulty of navigating voice posts. Pictorial contexts are diagrams or photographs that provide a background for a particular conversation; pictorial contexts serve both to represent the conversation as a whole, and allow comment proxies to take on additional meaning by virtue of their location with respect to the pictorial background. Finally, visual controls are a variety of visual user interface components for controlling the system, for example, a message play button.

Figure I. Three design sketches of the Picture Talk concept for applications set in rural India: (a) Rice Talk, for farmers to discuss problems with their Rice plants; (b) Health Talk, for villagers to discuss health problems; and (c) TinkerTalk, for people in a region to indicate that they need the services of a traveling tinker. Background images of rice plant © Ivan Kopylov I Dreamstime.com, of human body © Dannyphoto80 I Dreamstime.com, and of map © Robert Adrian Hillman I Dreamstime.com. Used with permission.

Figure 1 shows three early 'design sketches' of Picture Talk developed in the context of a scenario set in rural India. (By 'design sketches' we mean provisional concepts that are intended as conversation starters with stakeholders, rather than as depictions of well-considered solutions.) The first sketch, Rice Talk, envisions an asynchronous conversation among farmers about pests and diseases affecting their rice plants. It consists of (1) a diagram of a rice plant (the pictorial context); (2) a series of shaded bars (the comment proxies) that represent spoken comments, showing their durations, which of them have been made by the same speaker, and the part of the plant to which the comments refer; and (3) a floating 'talk' button (the visual control). The second sketch shows a health-oriented conversation with circles (the comment proxies) superimposed over a diagram of the human body (the pictorial context), the circles' positions indicating what aspect of the body or health they refer to and how they are related to other comments. The third sketch shows a conversation between a traveling tinker (i.e., a mender of pots) and potential customers, the pictorial context being a map of the region, and comment proxies (the balloons) indicating where the speaker is located.

Besides communicating the basic idea behind Picture Talk - using pictures, and simple visual representations of voice comments to provide scaffolding for asynchronous speech-based communication - the sketches serve other purposes. First of all, they illustrate the flexibility of the basic concepts. The pictorial contexts, and similarly the comment proxies, can represent a large range of topics, and even when depicted as simple geometric shapes, they can represent a considerable array of meta-information. Perhaps more importantly, the sketches are useful in raising a number of questions both within the design team, and with other audiences. How do the pictures get into the system? What sort of meta-information should comment proxies depict? Do different conversations benefit from the display of different comment meta-information? What sort of visual representations will be understandable by the envisioned user populations? How do users find their ways to particular conversations? As the aim of this paper is not to trace the trajectory of the design, it will not detail its evolution, but will instead move on to describe the user experience of the resulting working prototype.

Our implementation of Picture Talk consists of a client application running on the T-Mobile G1™ mobile phone, an Internet-enabled smartphone running the Google Android™ operating system, and a centralized data server running an application-specific Web service in the Ruby on Rails™ Web application server environment.

When users launch the client application on their mobile phone, their phone number is used to retrieve their user profile from the Web service. If this is the first time the user has accessed the service, he is prompted to record his name and take a picture of himself. The user is then presented with a menu that has four options: take a picture, view the gallery of shared pictures, view the profiles of other users, or update one's own profile.

Users can start a discussion by simply taking a picture using the phone's built-in camera and tapping anywhere on the photo. The system stores the picture in the gallery of shared pictures and records various metadata (e.g., who started the discussion, the time and date).Additional metadata could be stored, such as the location where the picture was taken. Subsequently, users can join an ongoing discussion by finding the picture in the gallery and tapping on it, and being led to the discussion screen.

The discussion screen (see Figure 2) implements the three user interface elements identified earlier: the pictorial context (in this case, a picture of a plant), 'comment proxies' (graphics on the upper right depicting spoken comments about the picture and associated speaker photo below), and visual controls (buttons at the bottom of the screen to control audio recording and playback). In the spoken comments area, to the right of the picture, each graphic represents a single comment from a user. Clicking on a graphic element displays the associated speaker's photo.

Clicking on a speaker's photo leads to the speaker's profile. The speaker's profile can help discussion participants determine their trust in the information provided. For example, a speaker may be a friend who is instantly recognizable from her photo or may be someone not known to the discussants but nonetheless reputable.

We designed the audio controls to allow the user to compose and review a recording before posting it to the discussion for others to hear. Before or after recording audio, we will allow the user to tap on the picture to point out something of interest in the picture, for example, the diseased part of a rice plant. The resulting visual annotation will consist of the comment in association with the particular location the speaker indicated on the picture. When a user posts a comment, the bar graphic is posted to the discussion area to the right of the picture. The bar graphic provides a visual "residue" of the comment recording (Hollan, Hutchins, & Kirsh, 2000) for subsequent users. The length of the bar reflects the length of the recording. Pressing anywhere on the bar graphic starts playing the recorded audio and displays any corresponding visual annotations on the picture. The same set of controls is used for both recording and playback, much like a music player. Users

Figure 2. A Picture Talk discussion

can pause the playback or replay the audio from the beginning, but they can only listen to one recorded comment at a time. The bar graphics in the spoken comments area are listed chronologically from top to bottom in a scrolling window with the most recent always visible.

Posting a comment stores the audio, and any visual annotations, with the discussion so that subsequent users accessing the picture can access the comment's audio and visual elements. It also stores metadata (who made the comment, the date and time of their comment). Picture Talk can also notify other participants in the discussion that there is a new comment. Individuals can block these notifications.

Picture Talk users can also browse through a scrolling gallery of all user photos. Touching a user's photo leads to that user's profile. The user profile includes a photo, contact information (telephone number) and a scrolling gallery of the pictures that anchor discussions the user has started. Touching a picture leads to a discussion screen with the given picture as the context. In this way, users can quickly find and engage in discussions started by other participants. This could be useful, for example, if a user has come across a farmer who has posted useful information about, for example, rice fungi and wants to see what other advice the farmer may have provided on similar topics. As users engage in discussions on a topic, their network of co-discussants grows. Photos of co-discussants can cue memory for relevant discussion contexts and serve as a visual index to organize the pictures anchoring discussions.

Picture Talk implements a concept we call 'blended synchrony' (Erickson et al., 2006), meaning that the same application supports (near) synchronous and asynchronous interaction among participants. Picture Talk discussions persist over time, with remarks separated by seconds, minutes, days, or even months. Some discussions will feel quite immediate and rapid-fire, whereas others may be slower paced, or might be more like announcements than a true conversation. It just depends on the pattern of participation. Blended synchrony is useful in environments where communication needs to be close to real-time in some cases but can be asynchronous in others. Cultural and societal as well as pragmatic factors may come into play in deciding when and how to communicate through ICT applications (Hudson, Christensen, Kellogg, & Erickson, 2002).

Picture Talk is architected as a client-server application (see Figure 3). The Picture Talk client is installed as a third-party Android application on the G1 mobile phone. The G1 has suitable

Figure 3. Picture Talk architecture

hardware for running the Picture Talk client: a 3.2-inch touch-screen display, wireless networking, a microphone, built-in speakers, a camera, and gigabytes of storage on a removable flash memory card. A number of smartphones provide similar functionality, but Picture Talk takes advantage of the Android operating system's (OS's) capability of accessing the phone's hardware, including detecting the presence of wireless network services, recording and playback of audio, controlling the built-in camera and storing pictures on the phone and in external storage.

The Picture Talk data server provides a persistent data model for the application's objects (discussions, pictures, comments, audio clips, users, etc.). To minimize the data exchanged between clients and server (and hence conserve wireless bandwidth), the server assigns version numbers to the data objects so that both client and server know when data object updates are needed to synchronize the data model. The server uses Rails' active record support to store and access the data objects in a MySQL® database. Pictures (images) and voice recordings (audio clips) are stored in files.

Communication with the Ruby on Rails server takes place over a wireless network. The Picture Talk client sends pictures captured with the phone's camera and audio captured with the phone's microphone to the Picture Talk server and automatically updates the currently displayed discussion. When the phone is not able to connect to the wireless network, new pictures from the camera are stored on its Secure Digital (SD) card, if available, or on the phone's local storage, thus enabling users to start discussions, make audio postings, listen to previously accessed audio postings, and update personal information. In this disconnected mode, data objects are stored and retrieved from a database local to the phone using Android's SQLite software library.

Given the current technology trajectory in developing nations, we expect to see increased adoption of smartphones in the next three to five years. But in order to get early feedback on our designs and provide benefit to the BoP, we are interested in deploying Picture Talk as widely as possible in the near term as well. Thus, we have developed a voice-only version of Picture Talk in order to make the application available to people with lower end mobile phones. We have also created a web version, suitable for kiosk or telecenter use.

The Picture Talk voice version (the "voice client ") allows people using basic mobile phones, commonly found in BoP environments, to listen to and record discussion comments. Users with camera phones can also exchange pictures with the Picture Talk data server via MMS. A Picture Talk server-side component, built using the open source Asterisk® Public Branch Exchange (PBX) telephony toolkit, provides voice and telephone keypad Interactive Voice Response (IVR) interfaces, and uses the Picture Talk data server to access application objects.

Both web and voice clients access the same data as the Android client, but display that data in a suitable way for the target platform. For example, the rice plant anchoring the discussion in Figure 1 could be sent to the Picture Talk data server using MMS. Subsequently when another user wants to participate in the discussion, the server can send the picture to their phone in another MMS message. Users listen to voice comments over a normal voice channel. We had to develop additional server-side functions to transform the audio and image objects into formats usable by and optimized for both mobile phones and desktop computers.

Searching and browsing discussions is difficult using basic mobile phones, so we have explored how to perform these operations using voice only. The first feature we implemented of this nature is 'voice tagging'. When users create a discussion, they are asked to add optional 'voice tags.' A voice tag is a spoken word, phrase, or other short descriptive audio recording that is then associated with the discussion. Voice tags play a similar role to textual tags on the Web (e.g., Youtube, Flickr). Users can press touchtone keys to listen to tags, navigate to more specific tags, and add a new tag. After choosing a voice tag, the user can listen to the description of discussions and select discussions using a touchtone key. We have also implemented a form of 'voice bookmarking'. After listening to a voice post, the user can use touchtone to bookmark the post. Like social bookmarks on the Web, users share bookmarks with one another. We are just starting to explore how to provide visual proxies for voice tags and bookmarks on smartphones.

While the user experience on basic mobile phones is necessarily more restrictive than with smartphone or web browser clients, having the option of using these inexpensive mobile devices makes Picture Talk discussions potentially available to a broad range of BoPusers. Further research is needed on methods of effectively finding and navigating relevant content, sharing information, and tapping into discussion databases on mobile devices. We hope that future mobile phone users at the BoPwill have access to information services that have heretofore been usable only from desktop computers in the hands of experienced users.

FUTURE RESEARCH DIRECTIONS

Many of Picture Talk's features represent general capabilities that could be applied in a variety of contexts. In future work, we would like to apply the concepts to a wider range of mobile social computing applications.

Another area of future work is improving user identification. Like many social software applications, Picture Talk helps users share photos, contact information, and so on, with each other. However, in many developing regions of the world, it is common for mobile phones to be shared amongst members of a family or even an entire village. Mobile social computing applications in many BoP contexts may need to allow users to identify themselves to the system explicitly and in innovative ways (e.g., by selecting their picture or identifying a vocal sample) in order to work in contexts of shared phone use.

One additional area of research is the propagation of information by social networks. Picture Talk currently promotes participation and inclusion by providing an open forum for discussion. Furthermore, the ability to notify and invite people to participate, even people who are not currently registered users of the system, supports the possibility of "viral" growth of the Picture Talk user population. In developing regions where the idea of using technology to access information beyond one's social network may be unfamiliar, this kind of "viral spread" can be a key bridging mechanism. However, supporting the application at a larger scale involves creating mechanisms whereby users can find each other easily. We are exploring how to organize information about people in ways that can be easily appropriated by the population, such as by village or family.

Another problem that has arisen in our work is efficient and effective navigation through large amounts of audiovisual content. As a conversational system, Picture Talk's audio postings could quickly grow to an unmanageable size as more users access the system. Interfaces for time-varying multimedia e.g., audio and video, do not offer the same navigational affordances as visual interfaces (Muller, Farrell, Cebulka & Smith, 1992). In Picture Talk, we have mitigated this problem by anchoring aural information to metadata that is made explicit through photos and graphics. Ultimately we would like users to be able to easily switch between visual or voice menus organized by author, topic, time period, location, and so on. A more complete solution will no doubt ultimately be needed.

CONCLUSION

We are at an exciting point in the history of mobile computing. For the first time, the billions of people in some of the world's most economically disadvantaged countries have the promise of participating in the information revolution through mobile phones. If successful, this could bring about positive social, political, and economic change in regions struggling with illiteracy, disease, poverty, natural disasters, oppression, and other challenges. Enabling ordinary citizens to become both producers and consumers of information could facilitate the spread of critical information from one person to the next during crises, encourage broad exchange of ideas, connect experts with those needing help, strengthen social networks, and enable people at the base of the economic pyramid to become full participants in society and world economic markets.

We introduced Picture Talk, a mobile social computing application framework we designed for use in environments with low literacy rates, limited Internet connectivity, and user populations with little familiarity with information services. Picture Talk can be accessed over a voice channel from mobile camera phones available to many BoP populations, but we have also implemented a smartphone version, anticipating further penetration of mobile touch screen devices in developing regions, and a Web version, suitable for kiosk or telecenter use. The limitations of using mobile devices to access rich structured multimedia content by users with limited literacy skills exposes human-computer interaction challenges that are key to enabling broad access to information by people in BoP populations.

REFERENCES

Abraham, R. (2007). Mobile phones and economic development: Evidence from the fishing industry in India. *MlT Press Journal*, *4*(1), 5-17.

Appfrica (2008). *The current state of lnternet penetration in Africa.* Retrieved June 15, 2009, from http://appfrica.net/blog/archives/248

Clark, H. H., & Brennan, S. E. (1991). Grounding in Communication. In L. Resnick, J. Levine & S. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (127-149). Hyattsville, MD: American Psychological Association.

Danis, C. Bailey, M., Christensen, J., Ellis, J., Erickson, T., Farrell, R., & Kellogg, W. A. (2009) Social Computing Applications for the Next Billion Users. In *Designing Future Mobile Software for Underserved Users Workshop* at CSCW 2008.

Erickson, T., Kellogg, W. A., Laff, M., Sussman, J., Wolf, T. V., Halverson, C. A., & Edwards, D. (2006). A persistent chat space for work groups: the design, evaluation and deployment of loops. In *Proceedings of the 6th Conference on Designing lnteractive Systems 06* (pp. 331-340). New York: ACM Press.

Froehlich, D. M., Rachovides, D., Riga, K., Bhat, R., Frank, M., Edirisinghe, E., et al. (2009). StoryBank: mobile digital storytelling in a development context. In *Proceedings of CHl 2009* (pp. 1761-1770). New York: ACM.

Gartner Group. (2009). *Gartner Says Worldwide Mobile Phone Sales Declined 8.6 Per Cent and Smartphones Grew 12.7 Per Cent in First Quarter of 2009.* Press released dated May 20, 2009. Retrieved June 15, 2009, from http://www.gartner.com/it/pa ge.jsp?id=985912

Gopakumar, K. (2006). E-governance services through telecentres: Role of human intermediary and issues of trust. *lnformation Technologies and Development*, *4*(1), 19-35.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human lnteraction*, *7*(2), 174-196. doi:10.1145/353485.353487

Hudson, J. H., Christensen, J., Kellogg, W. A., & Erickson, T. (2002). I'd be overwhelmed, but it's just one more thing to do. In *Proceedings of the S!GCH! conference on Human Factors in Computing Systems*, (pp. 97-104). New York: ACM Press.

Infotrend/CAP Ventures. (2004). W*orldwide Camera Phone and Photo Messaging Forecast: 2004-2009.*London: Kluwer. Retrieved June 15, 2009, from http://store.infotrendsresearch.com/ PhotoGall ery.asp?ProductCode=MobileImagin gStudy10106

ITU. (2009). *New !TU !CT Development !ndex compares 154 countries.* (press release dated March 2, 2009).http://www.itu.int/newsroom / press_releases/2009/07.html

Joshi, A., Welankar, N., Bl, N., Kanitkar, K., & Sheikh, R. (2008). Rangoli: A Visual Phonebook for Low-literate Users. *MobileHC! 2008*, Amsterdam, The Netherlands.

Kumar, A., Rajput, N., Agarwal, S., Chakraborty, D., & Nanavati, A. A. (2008). Organizing the unorganized: Employing IT to empower the under-privileged. In *Proceedings of the !nternational World-Wide Web Conference* (pp. 935-944). Beijing, China: ACM Press.

Lyra Research. (2007). *The Worldwide Camera Phone Forecast.* Retrieved March 3, 2010, from http://lyra.ecnext.com/coms2/ summary_0290-801_ITM

Madamombe, I. (2005). Energy key to Africa's prosperity: Challenges in West Africa's quest for electricity. *Africa Renewal*, *18*(4), 6.

Muller, M. J., Farrell, R., Cebulka, K. D., & Smith, J. G. (1992). Issues in the usability of time-varying multimedia. In Blattner, M. M., & Dannenberg, R. B. (Eds.), *Multimedia interface design* (pp. 7-38). New York: ACM Press.

OLPC. (2009). *One Laptop Per Child.* Retrieved June 15, 2009, from http://laptop.org

Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human-Computer !nteraction*, *15*(2), 139-178. doi:10.1207/S15327051HCI1523_4

OSISA. (2009). *Electricity for Africa?* Retrieved June 15, 2009, from http://www.osisa.org/ node/4164

Plauche, M., & Nallasamy, U. (2007). Speech interfaces for equitable access to information technology. *!nformation Technologies and !nternational Development*, *4*(1), 69-86. doi:10.1162/ itid.2007.4.1.69

Prahalad, C. K. (2004). *The Fortune at the Bottom of the Pyramid: Eradicating Poverty through Profits*. Upper Saddle River, NJ: Wharton School Publishing.

Samuel, J., Shah, N., & Hadingham, W. (2005). Mobile communications in SouthAfrica, Tanzania and Egypt: Results from Community and Business Surveys. *Africa: The !mpact of Mobile Phones. The Vodafone Policy Paper Series*, *2*(March), 44-52.

Seshagiri, S., Sagar, A., & Joshi, D. (2007). Connecting the 'Bottom of the Pyramid' - An Exploratory Case Study of India's Rural Communication Environment. *WWW 2007, May 8-12, 2007*, Alberta, Canada.

Srinivasan, J. (2007). The role of trustworthiness in information service usage: The case of Parry information kiosks in Tamil Nadu, India. In *Proceedings of the !nternational Conference on !nformation and Communication Technologies for Development (!CTD),* (pp. 345-352). New York: ACM Press.

The Pew Research Center for the People and the Press. (2007). *The Pew Global Attitudes Pro}ect*. Retrieved October 4, 2007, from http://www. pewglobal.org

UNESCO. (2009). *UNESCO Institute for Statistics.* Retrieved June 15, 2009, from http://www.uls.unesco.org

Ushahidi.com. (2009). *Crowdsourcing Crisis Information (FOSS).* Retrieved June 15, 2009, from http://www.ushahidi.com/

Yankelovich, W. W., Roberts, P., Wessler, M., Kaplan, J., & Provino, J. (2004). Meeting Central: Making distributed meetings more effective. In *Proceedings of the Conference on Computer-supported Cooperative Work (pp.* 419-428). New York: ACM Press.

ADDITIONAL READING

Reeks, R. (2008). ICT4D 2.0: The Next Phase of Applying ICT for International Development. *Computer*, *41*(6), 26-33. doi:10.1109/MC.2008.192

Parameswaran, M., & Whinston, A. B. (2007). Social Computing: An Overview. *Communications of the Association for Information Systems*, *19*, 762-780.

Zaslavsky, A. B., & Tari, Z. (1998). Mobile computing: Overview and current status. *Australian Computer Journal*, *30*(2), 42-52.

KEY TERMS AND DEFINITIONS

Visual Scaffolding: Images, words, or other visible elements with the sole purpose of supporting some primary user interaction e.g., tactile and/or aural user interfaces.

BoP: Stands for 'Base of the Pyramid' aka 'Bottom of the Pyramid'; a term referring to the largest but poorest economic group in the world with billions of people living on a small number of dollars per day (e.g., less than $1US/day).

Developing Regions: Parts of the world with a relatively low average standard of living when compared with advanced economies including parts of sub-Saharan Africa, South/Central America, the Indian Sub-continent, and southeast Asia.

Social Computing: An area of computer science concerned with intersection of social behavior and computation. This includes social behavior using computer systems and computing through the social behavior of large numbers of people.

MMS: Stands for Multimedia Messaging Services; a standard way to send messages that include multimedia content to and from mobile phones. Apopular use is to send photographs from camera-equipped handsets.

Smartphone: A mobile phone offering capabilities beyond text message and voice communication, often including functions often found on personal digital assistants (PDAs) or personal computers (PCs), for example, a way to connect to the Internet to send e-mail or documents.

ICT: Stands for Information and Communication Technology; hardware, software, or networks in services of the creation, transfer, or storage of information in all its forms (e.g., text, audio, images, video) including over voice over telecommunications networks; often used as an extension of IT (Information Technology) and

ENDNOTES

1. Android is a trademark of Google, Inc.
2. G1 is a trademark of T-Mobile USA, Inc.
3. Asterisk is a registered trademark of Digium, Inc.
4. MySQL is a registered trademark of MySQL AB in the United States, the European Union and other countries.